GOVERNMENT OF TAMILNADU

HIGHER SECONDARY SECOND YEAR

BUSINESS MATHEMATICS AND STATISTICS

VOLUME - II

A publication under Free Textbook Programme of Government of Tamil Nadu

Department of School Education

Untouchability is Inhuman and a Crime

Government of Tamil Nadu

First Edition - 2019

(Published under new syllabus)

NOT FOR SALE

Content Creation

State Council of Educational
Research and Training

Printing & Publishing

Tamil NaduTextbook and Educational
Services Corporation

HOW TO USE THE BOOK?

Career Options	List of Further Studies & Professions.
Learning Objectives:	Learning objectives are brief statements that describe what students will be expected to learn by the end of school year, course, unit, lesson or class period.
	Additional information about the concept.
DO YOU KNOW?	Amazing facts, Rhetorical questions to lead students to Mathematical inquiry
Exercise	Assess students' critical thinking and their understanding
ICT	To enhance digital skills among students
Web links	List of digital resources
QR GUIDE	To motivate the students to further explore the content digitally and take them in to virtual world
Miscellaneous Problems	Additional problems for the students
Glossary	Tamil translation of Mathematical terms
Book for References	List of related books for further studies of the topic

QR GUIDE

Lets use the QR code in the text books ! How ?

- Download the QR code scanner from the Google PlayStore/ Apple App Store into your smartphone
- Open the QR code scanner application
- Once the scanner button in the application is clicked, camera opens and then bring it closer to the QR code in the text book.
- Once the camera detects the QR code, a url appears in the screen.Click the url and goto the content page.

CAREER OPTIONS IN BUSINESS MATHEMATICS and STATISTICS

Higher Secondary students who have taken commerce with Business Mathematics and Statiscs can take up careers in BCA, B.Com., and B.Sc. Statistics. Students who have taken up commerce stream, have a good future in banking and financial institutions.

A lot of students choose to do B.Com with a specialization in computers. Higher Secondary Commerce students planning for further studies can take up careers in professional fields such as Company Secretary , Chartered Accountant (CA), ICAI and so on. Others can take up bachelor's degree in commerce (B.Com), followed by M.Com, Ph.D and M.Phil. There are wide range of career opportunities for B.Com graduates.

After graduation in commerce, one can choose MBA, MA Economics, MA Operational and Research Statistics at Postgraduate level. Apart from these, there are several diploma, certificate and vocational courses which provide entry level jobs in the field of commerce.

Career chart for Higher Secondary students who have taken commerce with Business Mathematics and Statistics.

Courses	Institutions	Scope for further studies
B.Com., B.B.A., B.B.M., B.C.A., B.Com (Computer), B.A.	• Government Arts & Science Colleges, Aided Colleges, Self financing Colleges. • Shri Ram College of Commerce (SRCC), Delhi • Symbiosis Society's College of Arts & Commerce, Pune. • St. Joseph's College, Bangalore	**C.A., I.C.W.A, C.S.**
B.Sc Statistics	• Presidency College, Chepauk, Chennai. • Dr. Ambedkar Govt. Arts College, Vyasarpadi, Chennai - 39. • Govt. Arts College, Tindivanam, Villupram & Nagercoil. • Madras Christian College, Tambaram • Loyola College, Chennai. • D.R.B.C.C Hindu College, Pattabiram, Chennai.	**M.Sc., Statistics**
B.B.A., LLB, B.A., LLB, B.Com., LL.B. (Five years integrated Course)	• Government Law College. • School of excellence, Affiliated to Dr.Ambethkar Law University	**M.L.**
M.A. Economics (Integrated Five Year course) – Admission based on All India Entrance Examination	• Madras School of Economics, Kotturpuram, Chennai.	**Ph.D.,**
B.S.W.	• School of Social studies, Egmore, Chennai	M.S.W

CONTENTS

As the Statistics component of this text book involves problems based on numerical calculations, Business Mathematics and Statistics students are advised to use calculator

VPV976 2N6NJ4 B284_12_BUS_MAT_EM

SYLLABUS

6.	**RANDOM VARIABLE AND MATHEMATICAL EXPECTATION**	**(21 periods)**
	Random Variable: Definition – Discrete random variable, probability mass function and cumulative distribution function – Continuous random variable, probability density function and cumulative distribution function. **Mathematical Expectation:** Definition, Mean and Variance – Properties of expectation and variance (without proof).	
7.	**PROBABILITY DISTRIBUTIONS**	**(21 periods)**
	Distributions: Binomial distribution: Definition, Mean and Variance – Poisson Distribution: Definition, Mean and Variance – Normal Distribution: Definition, Properties and Standard Normal Variate	
8.	**SAMPLING TECHNIQUES AND STATISTICAL INFERENCE**	**(21 periods)**
	Sampling: Meaning, types of sampling methods - Simple random sampling, Stratified Sampling and Systematic Sampling – Sampling and Non sampling errors – Sampling Distributions: Definition of Sampling Distributions and Standard Error – Computing Standard error in simple cases. Estimation: Point and interval estimation (concept only). **Hypothesis testing:** Meaning, Null hypothesis and Alternative hypothesis, Level of significance and types of errors. Testing procedure, Large sample theory and test of significance for single mean.	
9.	**APPLIED STATISTICS**	**(21 periods)**
	Time Series Analysis: Meaning, Uses and Basic Components – Estimating trends: Semi Average method – Moving Average method (three yearly, four yearly) – Method of least squares – Seasonal Variation by method of simple averages. **Index Numbers:** Meaning, Classifications and Uses – Weighted Averages by Laspyre's, Paasche's and Fisher's ideal index numbers – Time reversal and Factor reversal test – Construction of cost of living index. **Statistical Quality control:** Meaning – Causes for Variation - Assignable cause and Chance cause – Process control and product control – Construction of $\overline{X}$ and R chart	
10.	**OPERATIONS RESEARCH**	**(21 periods)**
	Transportation Problem: Definition and Formulation – Methods of finding initial basic feasible solutions, North west corner rule, least cost method and Vogel's approximation method. **Assignment Problems:** Definition and Formulation – Solution of assignment problems. **Decision theory:** Meaning –Situations: Certainty and uncertainty – Maximin and Minimax strategy.	

6 Random Variable and Mathematical Expectation

Simeon-Denis Poisson

(June 21, 1781 – April 25, 1840)

Introduction

Between Sylvestre-François Lacroix in 1816 and Louis Bachelier in 1914, the concepts random variable and mean of a random variable were invented. Some authors contend that Simeon-Denis Poisson invented the concepts random variable and expected value. Simeon-Denis Poisson, a French mathematician known for his work on probability theory. Poisson's research was mainly on Probability. In 1838, Poisson published his ideas on probability theory, which also included what is now known as Poisson distribution. He published between 300 and 400 mathematical works including applications to electricity and magnetism, and astronomy.

Learning Objectives

After studying this chapter students are able to understand

- Why do we use random variable?
- Why do we need to define the random variable?
- the types of Random variables.
- the probability function.
- the distribution function.
- the nature of problems.
- the methods for studying random experiments with outcomes that can be described numerically.
- the real probabilistic computation.
- the concept of mathematical expectation for discrete and continuous random variables.
- the properties of mathematical expectation for discrete and continuous random variables.
- the determination of mean and variance for discrete and continuous random variables.

- and the practical applications of mathematical expectations for discrete and continuous random variables.

6.1. Random variable

Introduction:

Let the random experiment be the toss of a coin. When 'n' coins are tossed, one may be interested in knowing the number of heads obtained. When a pair of dice is thrown, one may seek information about the sum of sample points. Thus, we associate a real number with each outcome of an experiment. In other words, we are considering a function whose domain is the set of possible outcomes and whose range is subset of the set of real numbers. Such a function is called random variable.

In algebra, you learned about different variables like X or Y or any other letter in a particular problem. Thus in basic mathematics, a variable is an alphabetical character that represents an unknown number. A random variable is a variable that is subject to randomness, which means it could take on different values. In statistics, it is quite general to use X to denote a random variable and it takes on different values depending on the situation.

Some of the examples of random variable:

(i) Number of heads, if a coin is tossed 8 times.

(ii) The return on an investment in one-year period.

(iii) Faces on rolling a die.

(iv) Number of customers who arrive at a bank in the regular interval of one hour between 9.00 a.m and 4.30 p.m from Monday to Friday.

(v) The sale volume of a store on a particular day.

For instance, the random experiment 'E' consists of three tosses of a coin and the outcomes of this experiment forms the sample space is 'S'. Let X denotes the number of heads obtained. Here X is a real number connected with the outcome of a random experiment E. The details given below

Tossing a Coin

Fig.6.1

Outcome (ω) : (HHH) (HHT) (HTH) (THH) (HTT) (THT) (TTH) (TTT)

Values of $X = x$: 3 2 2 2 1 1 1 0

i.e., $R_X = \{0, 1, 2, 3\}$

From the above said example, for each outcomes ω, there corresponds a real number $X(\omega)$Since the points of the sample space 'S' corresponds to outcomes is defined for each $\omega \in S$.

6.1.1 Definition of a random variable

Definition 6.1

A random variable (r.v.) is a real valued function defined on a sample space S and taking values in $(-\infty, \infty)$ or whose possible values are numerical outcomes of a random experiment.

Note

(i) If x is a real number, the set of all ω in S such that $X(\omega)=x$ is, denoted by $X=x$. Thus $P(X = x) = P\{\omega:X(\omega) = x\}$.

(ii) $P(X \leq a) = P\{\omega:X(\omega) \in (-\infty, a]\}$ and $P(a < X \leq b) = P\{\omega : X(\omega) \in (a, b]\}$.

(iii) One-dimensional random variables will be denoted by capital letters, $X, Y, Z, ...$, etc. A typical outcome of the experiment will be denoted by ω. Thusb $X(\omega)$ represents the real number which the random variable X associates with the outcome ω. The values which $X, Y, Z, ...$, etc, can assume are denoted by lower case letters, $viz., x, y, z, ...$, etc.

DO YOU KNOW?

(i) If X_1 and X_2 are random variables and C is a constant, then $CX_1, X_1 + X_2, X_1X_2, X_1 - X_2$ are also random variables.

(ii)If X is a random variable, then $(i)\frac{1}{X}$ and $(ii)|X|$ are also random variables.

Types of Random Variable:

Random variables are classified into two types namely discrete and continuous random variables. These are important for practical applications in the field of Mathematics and Statistics. The above types of random variable are defined with examples as follows.

6.1.2 Discrete random variable

Definition 6.2

A variable which can assume finite number of possible values or an infinite sequence of countable real numbers is called a discrete random variable.

Examples of discrete random variable:

- Marks obtained in a test.
- Number of red marbles in a jar.
- Number of telephone calls at a particular time.
- Number of cars sold by a car dealer in one month, etc.,

For instance, three responsible persons say, P_1, P_2, and P_3 are asked about their opinion in favour of building a model school in a certain district. Each person's response is recorded as Yes (Y) or No (N). Determine the random variable that could be of interest in this regard. The possibilities of the response are as follows

Table 6.1

Possibilities	P_1	P_2	P_3	Values of Random variable (Number of Yes those who are given)
1.	Y	Y	Y	3
2.	Y	Y	N	2
3.	Y	N	Y	2
4.	Y	N	N	1
5.	N	Y	Y	2
6.	N	Y	N	1
7.	N	N	Y	1
8.	N	N	N	0

Form the above table, the discrete random variable take values 0, 1, 2 and 3.

Probability Mass function

Definition 6.3

If X is a discrete random variable with distinct values $x_1, x_2, ..., x_n, ...$, then the function, denoted by $P_X(x)$ and defined by

$$P_X(x) = \text{p}(\text{x}) = \begin{cases} \text{P}(\text{X} = \text{x}_\text{i}) = p_i = p(x_i) & \text{if } x = x_i, i = 1, 2, ..., n, ... \\ 0 & \text{if } x \neq x_i \end{cases}$$

This is defined to be the probability mass function or discrete probability function of X.

The probability mass function $p(x)$ must satisfy the following conditions

(i) $p(x_i) \geq 0 \, \forall i$, (ii) $\sum_{i=1}^{\infty} p(x_i) = 1$

Example 6.1

The number of cars in a household is given below.

No. of cars	0	1	2	3	4
No. of Household	30	320	380	190	80

Estimate the probability mass function. Verify $p(x_i)$ is a probability mass function.

Solution:

Let X be the number of cars

Table 6.2

$X = x_i$	Number of Household	$p(x_i)$
0	30	0.03
1	320	0.32
2	380	0.38
3	190	0.19
4	80	0.08
Total	1000	1.00

(i) $p(x_i) \geq 0 \; \forall \; i$ and

(ii) $\sum_{i=1}^{\infty} p(x_i) = p(0)+p(1)+p(2)+p(3)+p(4)$

$= 0.03+0.32+0.38+0.19+0.08 = 1$

Hence $p(x_i)$ is a probability mass function.

Note

For $X=0$, the probability 0.03, comes from 30/1000, the other probabilities are estimated similarly.

Example 6.2

A random variable X has the following probability function

Values of X	0	1	2	3	4	5	6	7
$p(x)$	0	a	$2a$	$2a$	$3a$	a^2	$2a^2$	$7a^2 + a$

(i) Find a , Evaluate (ii) $P(X<3)$, (iii) $P(X>2)$, and (iv) $P(2<X\leq 5)$.

Solution:

(i) Since the condition of probability mass function

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

$$\sum_{i=0}^{7} p(x_i) = 1$$

$$0+a+2a+2a+3a+a^2+2a^2+7a^2+a = 1$$

$$10a^2+9a-1 = 0$$

$$(10a-1)(a+1) = 0$$

$$a = \frac{1}{10} \text{ and } -1$$

Since $p(x)$ cannot be negative, $a = -1$ is not applicable. Hence, $a = \dfrac{1}{10}$

(ii) $P(X<3)=P(X=0)+P(X=1)+P(X=2)$

$$= 0+a+2a$$

$$=3a$$

$$= \frac{3}{10} \qquad \left(\because a=\frac{1}{10}\right)$$

(iii) $P(X>2)=1-P(X\leq 2)$

$$=1-\left[P(X=0)+P(X=1)+P(X=2)\right]$$

$$= 1-\frac{3}{10}$$

$$= \frac{7}{10}$$

(iv) $P(2<X\leq 5)=P(X=3)+P(X=4)+P(X=5)$

$$= 2a+3a+a^2$$

$$= 5a+a^2$$

$$= \frac{5}{10}+\frac{1}{100}$$

$$= \frac{51}{100}$$

Example 6.3

$$\text{If } p(x) = \begin{cases} \frac{x}{20}, & x = 0, 1, 2, 3, 4, 5 \\ 0, & otherwise \end{cases}$$

Find (i) $P(X<3)$ and (ii) $P(2<X\leq 4)$

Solution:

$$\begin{aligned} P(X<3) &= P(X=0)+P(X=1)+P(X=2) \\ &= 0+\frac{1}{20}+\frac{2}{20} \\ &= \frac{3}{20} \\ P(2<X\leq 4) &= P(X=3)+P(X=4) \\ &= \frac{3}{20}+\frac{4}{20} \\ &= \frac{7}{20} \end{aligned}$$

Example 6.4

If you toss a fair coin three times, the outcome of an experiment consider as random variable which counts the number of heads on the upturned faces. Find out the probability mass function and check the properties of the probability mass function.

Solution:

Let X is the random variable which counts the number of heads on the upturned faces. The outcomes are stated below

Outcomes	(HHH)	(HHT)	(HTH)	(THH)	(THT)	(TTH)	(HTT)	(TTT)
Values of X	3	2	2	2	1	1	1	0

These values are summarized in the following probability table.

Table 6.3

Value of X	0	1	2	3	Total
$p(x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\sum_{i=0}^{3} p(x_i)=1$

(i) $p(x_i) \geq 0 \ \forall \ i$ and

(ii) $\sum_{i=0}^{3} p(x_i) = 1$

Hence, $p(x_i)$ is a probability mass function.

Example 6.5

Two unbiased dice are thrown simultaneously and sum of the upturned faces considered as random variable. Construct a probability mass function.

Solution:

$$\text{Sample space } (S) = \left\{ \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix} \right\}$$

Fig.6.2

Total outcomes : $n(S) = 36$

Table 6.4

Out comes	(1,1)	(1,2) (2,1)	(1,3) (2,2) (3,1)	(1,4) (2,3) (3,2) (4,1)	(1,5) (2,4) (3,3) (4,2) (5,1)	(1,6) (2,5) (3,4) (4,3) (5,2) (6,1)	(2,6) (3,5) (4,4) (5,3) (6,2)	(3,6) (4,5) (5,4) (6,3)	(4,6) (5,5) (6,4)	(5,6) (6,5)	(6,6)
Sum 0f the upturned faces (X)	2	3	4	5	6	7	8	9	10	11	12
$P_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Discrete distribution function

Definition 6.4

The discrete cumulative distribution function or distribution function of a real valued discrete random variable X takes the countable number of points $x_1, x_2, \ldots$ with corresponding probabilities $p(x_1),\ p(x_2), \ldots$ and then the cumulative distribution function is defined by

$$F_X(x) = P(X \le x),\ \text{for all } x \in R$$

$$\text{i.e., } F_X(x) = \sum_{x_i \le x} p(x_i)$$

For instance, suppose we have a family of two children. The sample space

S = {bb, bg, gb, gg}, where b =boy and g = girl

Let X be the random variable which counts the number of boys. Then, the values (X) corresponding to the sample space are 2, 1, 1, and 0.

Hence, the probability mass function of X is

$X = x$	0	1	2
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Then, we can form a cumulative distribution function of X is

$X = x$	0	1	2
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
$F_X(x) = P(X \le x)$	$\frac{1}{4}$	$\frac{1}{4}+\frac{1}{2}=\frac{3}{4}$	$\frac{3}{4}+\frac{1}{4}=1$

Example 6.6

A coin is tossed thrice. Let X be the number of observed heads. Find the cumulative distribution function of X .

Solution:

The sample space (S) = {(HHH), (HHT), (HTH), (HTT), (THH), (THT), (TTH), (TTT)}

X takes the values: 3, 2, 2, 1, 2, 1, 1, and 0

Range of $X(Rx)$	0	1	2	3
$P_x(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$F_x(x)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{7}{8}$	1

Thus, we have

$$F_X(x)=\begin{cases} 0, & for\ x<0 \\ \frac{1}{8}, & for\ 0\le x<1 \\ \frac{4}{8}, & for\ 1\le x<2 \\ \frac{7}{8}, & for\ 2\le x<3 \\ 1, & for\ x\le 3 \end{cases}$$

Example 6.7

Construct the distribution function for the discrete random variable X whose probability distribution is given below. Also draw a graph of p(x) and F(x).

$X = x$	1	2	3	4	5	6	7
$P(x)$	0.10	0.12	0.20	0.30	0.15	0.08	0.05

Solution

From the values of $p(x)$ given in the probability distribution, we obtain

$F(1) = P(x \le 1) = P(1) = 0.10$

$F(2) = P(x \le 2) = P(1) + P(2) = 0.10 + 0.12 = 0.22$

$F(3) = P(x \le 3) = P(1) + P(2) + P(3)$

$= F(2) + P(3)$

$= 0.22 + 0.20$

$= 0.42$

$F(4) = F(3) + P(4)$

$= 0.42 + 0.30$

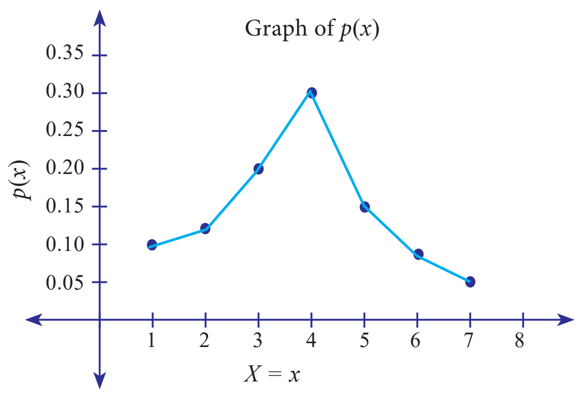


Fig.6.3

$$= 0.72$$

$$F(5) = F(4) + P(5)$$

$$= 0.72 + 0.15$$

$$= 0.87$$

$$F(6) = F(5) + P(6)$$

$$= 0.87 + 0.08$$

$$= 0.95$$

$$F(7) = F(6) + P(7)$$

$$= 0.95 + 0.05$$

$$= 1.00$$

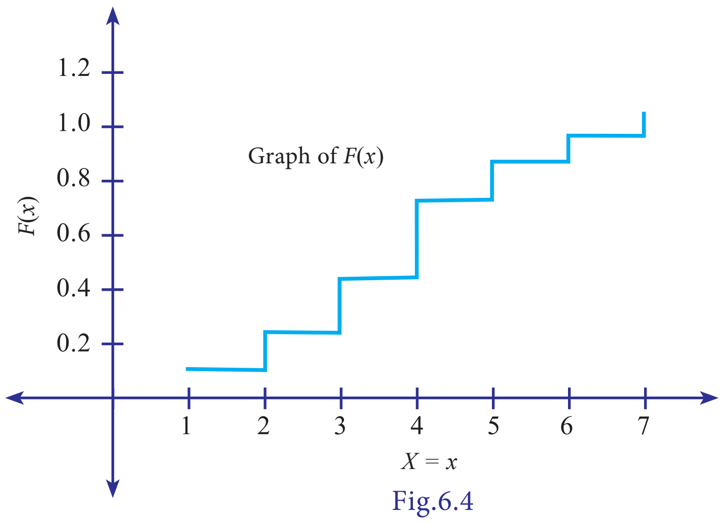


Fig.6.4

F(x) is

$$F_X(x) = \begin{cases} 0, & \text{if } x < 1 \\ 0.10, & \text{if } x \le 1 \\ 0.22, & \text{if } x \le 2 \\ 0.42, & \text{if } x \le 3 \\ 0.72, & \text{if } x \le 4 \\ 0.87, & \text{if } x \le 5 \\ 0.95, & \text{if } x \le 6 \\ 1, & \text{if } x \le 7 \end{cases}$$

6.1.3 Continuous random variable

Definition 6.5

A random variable X which can take on any value (integral as well as fraction) in the interval is called continuous random variable.

Examples of continuous random variable

- The amount of water in a 10 ounce bottle.
- The speed of a car.
- Electricity consumption in kilowatt hours.
- Height of people in a population.

- Weight of students in a class.
- The length of time taken by a truck driver to go from Chennai to Madurai, etc.

Probability density function

Definition 6.6

The probability that a random variable X takes a value in the interval $[t_1, t_2]$ (open or closed) is given by the integral of a function called the probability density function $f_X(x)$:

$$P(t_1 \le X \le t_2) = \int_{t_1}^{t_2} f_X(x)dx .$$

Other names that are used instead of probability density function include density function, continuous probability function, integrating density function.

The probability density functions $f_X(x)$ or simply by $f(x)$ must satisfy the following conditions.

(i) $f(x) \ge 0 \; \forall \; x$ and

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

Example 6.8

A continuous random variable X has the following p.d.f

$f(x) = ax \quad , \quad 0 \le x \le 1$

Determine the constant a and also find $P\left[X \le \frac{1}{2}\right]$

Solution:

We know that

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

$$\int_0^1 ax\, dx \Rightarrow a\int_0^1 xdx = 1$$

$$\Rightarrow a\left(\frac{x^2}{2}\right)_0^1 = 1$$

$$\Rightarrow \frac{a}{2}(1-0) = 1$$

$$\Rightarrow a = 2$$

$$P\left[x \leq \frac{1}{2}\right] = \int_{-\infty}^{\frac{1}{2}} f(x)dx$$

$$= \int_{0}^{\frac{1}{2}} axdx$$

$$= \int_{0}^{\frac{1}{2}} 2xdx$$

$$= \frac{1}{4}$$

Example 6.9

A continuous random variable X has p.d.f

$f(x) = 5x^4, 0 \leq x \leq 1$

Find a_1 and a_2 such that (i) $P[X \leq a_1] = P[X > a_1]$ (ii) $P[X > a_2] = 0.05$

Solution

(i) Since $P[X \leq a_1] = P[X > a_1]$

$$P[X \leq a_1] = \frac{1}{2}$$

i.e., $\displaystyle\int_{0}^{a_1} f(x)dx = \frac{1}{2}$

i.e., $\displaystyle\int_{0}^{a_1} 5x^4 dx = \frac{1}{2}$

$$5\left[\frac{x^5}{5}\right]_0^{a_1} = \frac{1}{2}$$

$$a_1 = (0.5)^{\frac{1}{5}}$$

(ii) $P[X > a_2] = 0.05$

$$\int_{a_2}^{1} f(x)dx = 0.05$$

$$\int_{a_2}^{1} 5x^4 dx = 0.05$$

$$5\left[\frac{x^5}{5}\right]_{a_2}^{1} = 0.05$$

$$a_2 = [0.95]^{\frac{1}{5}}$$

Continuous distribution function

Definition 6.7

If X is a continuous random variable with the probability density function $f_X(x)$, then the function $F_X(x)$ is defined by

$F_X(x) = P[X \leq x] = \int_{-\infty}^{x} f(t)dt, -\infty < x < \infty$ is called the distribution function (d.f) or sometimes the cumulative distribution function (c.d.f) of the continuous random variable X.

Properties of cumulative distribution function

The function $F_X(x)$ or simply $F(x)$ has the following properties

(i) $0 \leq F(x) \leq 1, -\infty < x < \infty$

(ii) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ and $F(+\infty) = \lim_{x \to \infty} F(x) = 1.$

(iii) $F(\cdot)$ is a monotone, non-decreasing function; that is, $F(a) \leq F(b)$ for $a < b$.

(iv) $F(\cdot)$ is continuous from the right; that is, $\lim_{h \to 0} F(x+h) = F(x)$.

(v) $F'(x) = \frac{d}{dx}F(x) = f(x) \geq 0$

(vi) $F'(x) = \frac{d}{dx}F(x) = f(x) \Rightarrow dF(x) = f(x)dx$

$dF(x)$ is known as probability differential of X.

(vii)
$$\begin{aligned} P(a \leq x \leq b) &= \int_a^b f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx \\ &= P(X \leq b) - P(X \leq a) \\ &= F(b) - F(a) \end{aligned}$$

Example 6.10

Suppose, the life in hours of a radio tube has the following p.d.f

$$f(x) = \begin{cases} \frac{100}{x^2}, when\, x \geq 100 \\ 0, \quad when\, x < 100 \end{cases}$$

Find the distribution function.

Solution:

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

$$= \int_{100}^{x} \frac{100}{t^2} dt, x \geq 100$$

$$= \left[\frac{100}{-t}\right]_{100}^{x}, x \geq 100$$

$$F(x) = \left[1 - \frac{100}{x}\right], x \geq 100$$

Example 6.11

The amount of bread (in hundreds of pounds) x that a certain bakery is able to sell in a day is found to be a numerical valued random phenomenon, with a probability function specified by the probability density function $f(x)$ is given by

$$f(x) = \begin{cases} Ax, \textit{ for } 0 \leq x < 10 \\ A(20-x), \textit{ for } 10 \leq x < 20 \\ 0, \qquad \textit{otherwise} \end{cases}$$

(a) Find the value of A.

(b) What is the probability that the number of pounds of bread that will be sold tomorrow is

(i) More than 10 pounds,

(ii) Less than 10 pounds, and

(iii) Between 5 and 15 pounds?

Solution:

(a) We know that

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

$$\int_{0}^{10} Axdx + \int_{10}^{20} A(20-x)dx = 1$$

$$A\left\{\left[\frac{x^2}{2}\right]_0^{10} + \left[20x - \frac{x^2}{2}\right]_{10}^{20}\right\} = 1$$

$$A\left[(50-0)+(400-200)-(200-50)\right]=1$$

$$A=\frac{1}{100}$$

(b) (i) The probability that the number of pounds of bread that will be sold tomorrow is more than 10 pounds is given by

$$P(10\le X\le 20)=\int_{10}^{20}\frac{1}{100}(20-x)\,dx$$

$$=\frac{1}{100}\left[20x-\frac{x^2}{2}\right]_{10}^{20}$$

$$=\frac{1}{100}\left[(400-200)-(200-50)\right]$$

$$=0.5$$

(ii) The probability that the number of pounds of bread that will be sold tomorrow is less than 10 pounds, is given by

$$P(0\le X<10)=\int_{0}^{10}\frac{1}{100}x\,dx$$

$$=\frac{1}{100}\left[\frac{x^2}{2}\right]_{0}^{10}$$

$$=\frac{1}{100}(50-0)$$

$$=0.5$$

(ii) The probability that the number of pounds of bread that will be sold tomorrow is between 5 and 15 pounds is

$$P(5\le X\le 15)=\int_{5}^{10}\frac{1}{100}x\,dx+\int_{10}^{15}\frac{1}{100}(20-x)dx$$

$$=\frac{1}{100}\left[\frac{x^2}{2}\right]_{5}^{10}+\frac{1}{100}\left[20x-\frac{x^2}{2}\right]_{10}^{15}$$

$$=0.75$$

Exercise 6.1

1. Construct cumulative distribution function for the given probability distribution.

X	0	1	2	3
$P(X=x)$	0.3	0.2	0.4	0.1

2. Let X be a discrete random variable with the following p.m.f

$$p(x)=\begin{cases}0.3 & for\ x=3\\ 0.2 & for\ x=5\\ 0.3 & for\ x=8\\ 0.2 & for\ x=10\\ 0 & otherwise\end{cases}$$

Find and plot the c.d.f. of X.

3. The discrete random variable X has the following probability function

$$P(X=x)=\begin{cases}kx & x=2,4,6\\ k(x-2) & x=8\\ 0 & otherwise\end{cases}$$

where k is a constant. Show that $k=\frac{1}{18}$

4. The discrete random variable X has the probability function

X	1	2	3	4
$P(X=x)$	k	$2k$	$3k$	$4k$

Show that $k = 0{\cdot}1$.

5. Two coins are tossed simultaneously. Getting a head is termed as success. Find the probability distribution of the number of successes.

6. A continuous random variable X has the following probability function

Value of $X=x$	0	1	2	3	4	5	6	7
$P(x)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

(i) Find k

(ii) Ealuate $p(x<6)$, $p(x \geq 6)$ and $p(0<x<5)$

(iii) If $P(X\leq x) > \frac{1}{2}$, then find the minimum value of x.

7. The distribution of a continuous random variable X in range (–3, 3) is given by p.d.f.

$$f(x)=\begin{cases}\frac{1}{16}(3+x)^2, & -3\le x\le -1\\ \frac{1}{16}(6-2x^2), & -1\le x\le 1\\ \frac{1}{16}(3-x)^2, & 1\le x\le 3\end{cases}$$

Verify that the area under the curve is unity.

8. A continuous random variable X has the following distribution function:

$$F(x)=\begin{cases}0 & ,\text{if } x\le 1\\ k(x-1)^4 & ,\text{if } 1<x\le 3\\ 1 & ,\text{if } x>3\end{cases}$$

Find (i) k and (ii) the probability density function.

9. The length of time (in minutes) that a certain person speaks on the telephone is found to be random phenomenon, with a probability function specified by the probability density function $f(x)$ as $f(x)=\begin{cases}Ae^{-x/5}, & \text{for } x\ge 0\\ 0 & ,\text{otherwise}\end{cases}$

(a) Find the value of A that makes $f(x)$ a p.d.f.

(b) What is the probability that the number of minutes that person will talk over the phone is (i) more than 10 minutes, (ii) less than 5 minutes and (iii) between 5 and 10 minutes.

10. Suppose that the time in minutes that a person has to wait at a certain station for a train is found to be a random phenomenon with a probability function specified by the distribution function

$$F(x)=\begin{cases}0, & \text{for } x\le 0\\ \frac{x}{2}, & \text{for } 0\le x<1\\ \frac{1}{2}, & \text{for } 1\le x<2\\ \frac{x}{4}, & \text{for } 2\le x<4\\ 1, & \text{for } x\ge 4\end{cases}$$

(a) Is the distribution function continuous? If so, give its probability density function?

(b) What is the probability that a person will have to wait (i) more than 3 minutes, (ii) less than 3 minutes and (iii) between 1 and 3 minutes?

11. Define random variable.

12. Explain what are the types of random variable?

13. Define discrete random variable.

14. What do you understand by continuous random variable?

15. Describe what is meant by a random variable.

16. Distinguish between discrete and continuous random variable.

17. Explain the distribution function of a random variable.

18. Explain the terms (i) probability mass function, (ii) probability density function and (iii) probability distribution function.

19. What are the properties of (i) discrete random variable and (ii) continuous random variable?

20. State the properties of distribution function.

6.2. Mathematical Expectation

Introduction

An extremely useful concept in problems involving random variables or distributions is that of expectation. Random variables can be characterized and dealt with effectively for practical purposes by consideration of quantities called their expectation. The concept of mathematical expectation arose in connection with games of chance. For example, a gambler might be interested in his average winnings at a game, a businessman in his average profits on a product, and so on. The average value of a random phenomenon is also termed as its Mathematical expectation or expected value. In the following sections, we will define and study the concept of mathematical expectation for both discrete and continuous random variables, which will be used in the following subsection.

6.2.1 Expected value and Variance

Expected value

The expected value is a weighted average of the values of a random variable may assume. The weights are the probabilities.

Definition 6.8

Let X be a discrete random variable with probability mass function (p.m.f.) $p(x)$. Then, its expected value is defined by

$$E(X) = \sum_{x} x\,p(x) \quad \text{... (1)}$$

If X is a continuous random variable and $f(x)$ is the value of its probability density function at x, the expected value of X is

$$E(X) = \int_{-\infty}^{\infty} x\,f(x)\,dx \quad \text{... (2)}$$

Note

- In (1), $E(X)$ is defined to be the indicated series provided that the series is absolutely convergent; otherwise, we say that the mean does not exist.
- In (1), $E(X)$ is an "average" of the values that the random variable takes on, where each value is weighted by the probability that the random variable is equal to that value. Values that are more probable receive more weight.
- In (2), $E(X)$ is defined to be the indicated integral if the integral exists; otherwise, we say that the mean does not exist.
- In (2), $E(X)$ is an "average" of the values that the random variable takes on, where each value x is multiplied by the approximate probability that X equals the value x, namely $f_X(x)dx$ and then integrated over all values.

DO YOU KNOW?

- $E(X)$ is the center of gravity or centroid of the unit mass that is determined by the density function of X. So the mean of X is measure of where the values of the random variable X are "centered".
- The mean of X, denoted by μ_x or $E(X)$.

Variance

The variance is a weighted average of the squared deviations of a random variable from its mean. The weights are the probabilities. The mean of a random variable X, defined in (1) and (2), was a measure of central location of the density of X. The variance of a random variable X will be a measure of the spread or dispersion of the density of X or simply the variability in the values of a random variable.

Definition 6.9

The variance of X is defined by

$$Var(X) = \sum \left[x - E(X)\right]^2 p(x) \qquad \text{... (3)}$$

if X is discrete random variable with probability mass function $p(x)$.

$$Var(X) = \int_{-\infty}^{\infty} \left[x - E(X)\right]^2 f_X(x)\,dx \qquad \text{... (4)}$$

if X is continuous random variable with probability density function $f_X(x)$.

Definition 6.10

Expected value of $[X - E(X)]^2$ is called the variance of the random variable.

i.e., $$Var(X) = E\left[X - E(X)\right]^2 = E\left(X^2\right) - \left[E(X)\right]^2 \qquad \text{... (5)}$$

where $$E\left(X^2\right) = \begin{cases} \sum\limits_x x^2\, p(x), \text{ if } X \text{ is Discrete Random Variable} \\ \int\limits_{-\infty}^{\infty} x^2 f(x)\,dx, \text{ if } X \text{ is Continuous Random Variable} \end{cases}$$

Note

- In the following examples, variance will be found using definition 6.10.
- The variances are defined only if the series in (3) is convergent or if the integrals in (4) exist.
- If X is a random variable, the standard deviation of X (S.D(X)), denoted by σ_X, is defined as $+\sqrt{Var[X]}$.
- The variance of X, denoted by $\sigma_X{}^2$ or $Var(X)$ or $V(X)$

DO YOU KNOW?

- Mean is the center of gravity of a density; similarly, variance represents the moment of inertia of the same density with respect to a perpendicular axis through the center of gravity.

6.2.2 Properties of Mathematical expectation

(i) $E(a) = a$, where 'a' is a constant

(ii) $E(aX) = aE(X)$

(iii) $E(aX+b)=aE(X)+b$, where 'a' and 'b' are constants.

(iv) If $X \geq 0, then\, E(X) \geq 0$

(v) $V(a)=0$

(vi) If X is random variable, then $V(aX+b)=a^2V(X)$

Concept of moments

The moments (or raw moments) of a random variable or of a distribution are the expectations of the powers of the random variable which has the given distribution.

Definition 6.11

If X is a random variable, then the r^{th} moment of X, usually denoted by μ'_r, is defined as

$$\mu'_r = E\left(X^r\right) = \begin{cases} \sum_x x^r p(x), & \text{for discrete random variable} \\ \int_{-\infty}^{\infty} x^r f(x)dx, & \text{for continuous random variable} \end{cases}$$

provided the expectation exists.

Definition 6.12

If X is a random variable, the r^{th} central moment of X about a is defined as $E\left[(X-a)^r\right]$. If $a = \mu_x$, we have the r^{th} central moment of X about μ_x, denoted by μ_r, which is

$$\mu_r = E[(X-\mu_X)^r]$$

Note

- $\mu'_1 = E(X) = \mu_X$, the mean of X.
- $\mu_1 = E[X-\mu_X] = 0$.
- $\mu_2 = E[(X-\mu_X)^2]$, the variance of X.
- All odd moments of X about μ_X are 0 if the density function of X is symmetrical about μ_X, provided such moments exist.

Example 6.12

Determine the mean and variance of the random variable X having the following probability distribution.

$X = x$	1	2	3	4	5	6	7	8	9	10
$P(x)$	0.15	0.10	0.10	0.01	0.08	0.01	0.05	0.02	0.28	0.20

Solution:

Mean of the random variable $X = E(X) = \sum_x x P_X(x)$

$= (1 \times 0.15) + (2 \times 0.10) + (3 \times 0.10) + (4 \times 0.01) + (5 \times 0.08) + (6 \times 0.01) +$

$(7 \times 0.05) + (8 \times 0.02) + (9 \times 0.28) + (10 \times 0.20)$

$E(X) = 6.56$

$$E\left(X^2\right) = \sum_x x^2 P_X(x)$$

$= (1^2 \times 0.15) + (2^2 \times 0.10) + (3^2 \times 0.10) + (4^2 \times 0.01) +$

$(5^2 \times 0.08) + (6^2 \times 0.01) + (7^2 \times 0.05) + (8^2 \times 0.02) +$

$(9^2 \times 0.28) + (10^2 \times 0.20).$

$= 50.38$

Variance of the Random Variagble $X = V(X) = E\left(X^2\right) - [E(X)]^2$

$= 50.38 - (6.56)^2$

$= 7.35$

Therefore, the mean and variance of the given discrete distribution are 6.56 and 7.35 respectively.

Example 6.13

Six men and five women apply for an executive position in a small company. Two of the applicants are selected for an interview. Let X denote the number of women in the interview pool. We have found the probability mass function of X.

$X = x$	0	1	2
$P(x)$	$\frac{2}{11}$	$\frac{5}{11}$	$\frac{4}{11}$

How many women do you expect in the interview pool?

Solution:

Expected number of women in the interview pool is

$$E(X) = \sum_x x\,P_X(x)$$

$$= \left[\left(0\times\frac{2}{11}\right)+\left(1\times\frac{5}{11}\right)+\left(2\times\frac{4}{11}\right)\right]$$

$$= \frac{13}{11}$$

Example 6.14

Determine the mean and variance of a discrete random variable, given its distribution as follows.

$X = x$	1	2	3	4	5	6
$F_x(x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	1

Solution

From the given data, you first calculate the probability distribution of the random variable. Then using it you calculate mean and variance.

X	$p(x)$
1	$F(1) = \frac{1}{6}$
2	$F(2) - F(1) = \frac{2}{6}-\frac{1}{6}=\frac{1}{6}$
3	$F(3) - F(2) = \frac{3}{6}-\frac{2}{6}=\frac{1}{6}$
4	$F(4) - F(3) = \frac{4}{6}-\frac{3}{6}=\frac{1}{6}$
5	$F(5) - F(4) = \frac{5}{6}-\frac{4}{6}=\frac{1}{6}$
6	$F(6) - F(5) = 1-\frac{5}{6}=\frac{1}{6}$

The probability mass function is

$X = x$	1	2	3	4	5	6
$P(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Mean of the random variable $X = E(X) = \sum_x x P_X(x)$

$$= \left(1\times\frac{1}{6}\right)+\left(2\times\frac{1}{6}\right)+\left(3\times\frac{1}{6}\right)+\left(4\times\frac{1}{6}\right)+\left(5\times\frac{1}{6}\right)+\left(6\times\frac{1}{6}\right)$$

$$= \frac{1}{6}(1+2+3+4+5+6)$$

$$= \frac{7}{2}$$

$$E\left(X^2\right) = \sum_x x^2 P_X(x)$$

$$= \left(1^2\times\frac{1}{6}\right)+\left(2^2\times\frac{1}{6}\right)+\left(3^2\times\frac{1}{6}\right)+\left(4^2\times\frac{1}{6}\right)+\left(5^2\times\frac{1}{6}\right)+\left(6^2\times\frac{1}{6}\right)$$

$$= \frac{1}{6}\left(1^2+2^2+3^2+4^2+5^2+6^2\right)$$

$$= \frac{91}{6}$$

Variance of the Random Variable $X = V(X) = E\left(X^2\right) - [E(X)]^2$

$$= \frac{91}{6} - \left(\frac{7}{2}\right)^2$$

$$= \frac{35}{12}$$

Example 6.15

The following information is the probability distribution of successes.

No. of Successes	0	1	2
Probability	$\frac{6}{11}$	$\frac{9}{22}$	$\frac{1}{22}$

Determine the expected number of success.

Solution:

Expected number of success is

$$E(X) = \sum_x x P_X(x)$$

$$= \left(0\times\frac{6}{11}\right)+\left(1\times\frac{9}{22}\right)+\left(2\times\frac{1}{22}\right)$$

$$= \frac{11}{22}$$

$$= 0.5$$

Therefore, the expected number of success is 0.5. Approximately one success.

Example 6.16

An urn contains four balls of red, black, green and blue colours. There is an equal probability of getting any coloured ball. What is the expected value of getting a blue ball out of 30 experiments with replacement?

Solution:

Probability of getting a blue ball = $(p) = \frac{1}{4} = 0.25$

$$
\begin{aligned}
\text{Total experiments (N)} &= 30 \\
\text{Expected value} &= \text{Number of experiments} \times \text{Probability} \\
&= N \times p \\
&= 30 \times 0.25 \\
&= 7.50
\end{aligned}
$$

Therefore, the expected value of getting blue ball is approximately 8.

Example 6.17

A fair die is thrown. Find out the expected value of its outcomes.

Solution:

If the random variable X is the top face of a tossed, fair, six sided die, then the probability mass function of X is

$$P_X(x) = \frac{1}{6}, \text{ for } x = 1,2,3,4,5 \text{ and } 6$$

The average toss, that is, the expected value of X is

$$E(X) = \sum_x x\,P_X(x)$$

$$
\begin{aligned}
\mathrm{E(X)} &= \left(1\times\frac{1}{6}\right)+\left(2\times\frac{1}{6}\right)+\left(3\times\frac{1}{6}\right)+\left(4\times\frac{1}{6}\right)+\left(5\times\frac{1}{6}\right)+\left(6\times\frac{1}{6}\right) \\
&= \frac{1}{6}(1+2+3+4+5+6) \\
&= \frac{7}{2} \\
&= 3.5
\end{aligned}
$$

Therefore, the expected toss of a fair six sided die is 3.5.

Example 6.18

Suppose the probability mass function of the discrete random variable is

$X = x$	0	1	2	3
$p(x)$	0.2	0.1	0.4	0.3

What is the value of $E(3X + 2X^2)$?

Solution:

$$E(X) = \sum_x x P_X(x)$$

$$= (0 \times 0.2) + (1 \times 0.1) + (2 \times 0.4) + (3 \times 0.3)$$

$$= 1.8$$

$$E\left(X^2\right) = \sum_x x^2 P_X(x)$$

$$= (0^2 \times 0.2) + (1^2 \times 0.1) + (2^2 \times 0.4) + (3^2 \times 0.3)$$

$$= 4.4$$

$$E(3X + 2X^2) = 3E(X) + 2E(X^2)$$

$$= (3 \times 1.8) + (2 \times 4.4)$$

$$= 14.2$$

Example 6.19

Consider a random variable X with probability density function

$$f(x) = \begin{cases} 4x^3, if\ 0 < x < 1 \\ 0,\ otherwise \end{cases}$$

Find $E(X)$ and $V(X)$.

Solution:

We know that,

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$= \int_0^1 x\, 4x^3\, dx$$

$$= 4\left[\frac{x^5}{5}\right]_0^1$$

$$E(X) = \frac{4}{5}$$

$$E\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$$

$$= \int_0^1 x^2\, 4x^3\, dx$$

$$= 4\left[\frac{x^6}{6}\right]_0^1$$

$$= \frac{4}{6}$$

$$V(X) = E\left(X^2\right) - \left[E(X)\right]^2$$

$$= \frac{4}{6} - \left[\frac{4}{5}\right]^2$$

$$= \frac{2}{75}$$

Example 6.20

If $f(x)$ is defined by $f(x) = ke^{-2x}, 0 \le x < \infty$

is a density function. Determine the constant k and also find mean.

Solution:

We know that

$$\int_{-\infty}^{\infty} f(x)\, dx = 1, \text{ since } f(x) \text{ is a density function.}$$

$$\int_0^{\infty} k\, e^{-2x}\, dx = 1$$

$$k\int_0^{\infty} e^{-2x}\, dx = 1$$

$$k\left[\frac{e^{-2x}}{-2}\right]_0^{\infty} = 1$$

$$k = 2$$

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$= \int_{0}^{\infty} xk\, e^{-2x}\, dx$$

$$= 2\int_{0}^{\infty} xe^{-2x}\, dx$$

$$= 2\left\{\left[\frac{xe^{-2x}}{-2}\right]_0^{\infty} - \int_{0}^{\infty} \frac{e^{-2x}}{-2} dx\right\} \quad \left(\because \int udv = uv - \int vdu\right)$$

$$= \int_{0}^{\infty} e^{-2x}\, dx$$

$$= \frac{1}{2}$$

Example 6.21

The time to failure in thousands of hours of an important piece of electronic equipment used in a manufactured DVD player has the density function.

$$f(x) = \begin{cases} 3e^{-3x}, & x > 0 \\ 0, & otherwise \end{cases}$$

Find the expected life of the piece of equipment.

Solution:

We know that,

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$= \int_{0}^{\infty} x\, 3e^{-3x} dx$$

$$= 3\int_{0}^{\infty} x\, e^{-3x} dx$$

$$= 3\left\{\left[x\frac{e^{-3x}}{-3}\right]_0^{\infty} - \int_{0}^{\infty}\left(\frac{e^{-3x}}{-3}\right)dx\right\} \left(\because \int u\, dv = u\, v - \int v\, du\right)$$

$$= \int_{0}^{\infty} e^{-3x} dx$$

$$= \frac{1}{3}$$

Therefore, the expected life of the piece of equipment is $\frac{1}{3}$ hrs (in thousands).

Example 6.22

A commuter train arrives punctually at a station every 25 minutes. Each morning, a commuter leaves his house and casually walks to the train station. Let X denote the amount of time, in minutes, that commuter waits for the train from the time he reaches the train station. It is known that the probability density function of X is

$$f(x)=\begin{cases}\frac{1}{25}, & for\, 0<x<25\\ 0, & otherwise.\end{cases}$$

Obtain and interpret the expected value of the random variable X.

Solution:

Expected value of the random variable is

$$\begin{aligned}E(X)&=\int_{-\infty}^{\infty} x\, f(x)\, dx\\ &=\int_0^{25} x\frac{1}{25}dx\\ &=\frac{1}{25}\int_0^{25} x\, dx\\ &=\frac{1}{25}\left[\frac{x^2}{2}\right]_0^{25}\\ &=12.5\end{aligned}$$

Therefore, the expected waiting time of the commuter is 12.5 minutes.

Example 6.23

Suppose the life in hours of a radio tube has the probability density function

$$f(x)=\begin{cases}e^{-\frac{x}{100}}, & when\, x\geq 100\\ 0, & when\, x<100\end{cases}$$

Find the mean of the life of a radio tube.

Solution:

We know that, the expected random variable

$$E(X)=\int_{-\infty}^{\infty} x\,f(x)\,dx$$

$$=\int_{100}^{\infty} x\,e^{-\frac{x}{100}}\,dx$$

$$=\left\{\left[x\left(\frac{e^{-\frac{x}{100}}}{-\frac{1}{100}}\right)\right]_{100}^{\infty}-\int_{100}^{\infty}\left(\frac{e^{-\frac{x}{100}}}{-\frac{1}{100}}\right)dx\right\} \qquad \left(\because \int u\,dv=uv-\int v\,du\right)$$

$$=\left[(10000)\left(e^{-1}\right)+(10000)\left(e^{-1}\right)\right]$$

$$=\left[(10000)(0.3679)+(10000)(0.3679)\right]$$

$$=7358\,hours$$

Therefore, the mean life of a radio tube is 7,358 hours.

Example 6.24

The probability density function of a random variable X is

$$f(x)=ke^{-|x|},-\infty<x<\infty$$

Find the value of k and also find mean and variance for the random variable.

Solution

We know that,

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

$$\int_{-\infty}^{\infty} k\,e^{-|x|}dx=1$$

$$k\int_{-\infty}^{\infty} e^{-|x|}dx=1$$

$$2k\int_{0}^{\infty} e^{-x}dx=1 \qquad (\because\ x^2e^{-|x|} \text{ is an even function})$$

$$2k\left[\frac{e^{-x}}{-1}\right]_{0}^{\infty}=1$$

$$k=\frac{1}{2}$$

Mean of the random variable is

$$E(X)=\int_{-\infty}^{\infty} x\,f(x)\,dx$$

$$E(X)=\int_{-\infty}^{\infty} x\,k\,e^{-|x|}dx \qquad (\because\ xe^{-|x|} \text{ is an odd function of } x)$$

$$=\frac{1}{2}\int_{-\infty}^{\infty} x\,e^{-|x|}dx$$

$$=0$$

$$E\left(X^2\right)=\int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$=\int_{-\infty}^{\infty} x^2 k\,e^{-|x|}dx$$

$$=\frac{1}{2}\int_{-\infty}^{\infty} x^2 e^{-|x|}dx$$

$$=\int_{0}^{\infty} x^2 e^{-x}dx \qquad (\because\ x^2e^{-|x|} \text{ is an even function})$$

$$=\Gamma 3 \qquad \left(\because \Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}e^{-x}dx, \alpha>0; \Gamma n=(n-1)!\right)$$

$$=2$$

$$V(X)=E\left(X^2\right)-\left[E\left(X\right)\right]^2$$

$$=2-[0]^2$$

$$=2$$

Exercise 6.2

1. Find the expected value for the random variable of an unbiased die

2. Let X be a random variable defining number of students getting A grade. Find the expected value of X from the given table

$X=x$	0	1	2	3
$P(X=x)$	0.2	0.1	0.4	0.3

3. The following table is describing about the probability mass function of the random variable X

x	3	4	5
$P(x)$	0.1	0.1	0.2

Find the standard deviation of x.

4. Let X be a continuous random variable with probability density function

$$f_X(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & otherwise \end{cases}$$

Find the expected value of X.

5. Let X be a continuous random variable with probability density function

$$f(x) = \begin{cases} \frac{3}{x^4}, x \ge 1 \\ 0, \quad otherwise \end{cases}$$

Find the mean and variance of X.

6. In an investment, a man can make a profit of ₹ 5,000 with a probability of 0.62 or a loss of ₹ 8,000 with a probability of $0 \cdot 38$. Find the expected gain.
7. What are the properties of Mathematical expectation?
8. What do you understand by Mathematical expectation?
9. How do you define variance in terms of Mathematical expectation?
10. Define Mathematical expectation in terms of discrete random variable.
11. State the definition of Mathematical expectation using continuous random variable.
12. In a business venture a man can make a profit of ₹ 2,000 with a probability of 0.4 or have a loss of ₹ 1,000 with a probability of $0 \cdot 6$. What is his expected, variance and standard deviation of profit?
13. The number of miles an automobile tire lasts before it reaches a critical point in tread wear can be represented by a p.d.f.

$$f(x) = \begin{cases} \frac{1}{30} e^{-\frac{x}{30}}, & for\ x > 0 \\ 0 & , for\ x \le 0 \end{cases}$$

Find the expected number of miles (in thousands) a tire would last until it reaches the critical tread wear point.

14. A person tosses a coin and is to receive ₹ 4 for a head and is to pay ₹ 2 for a tail. Find the expectation and variance of his gains.

15. Let X be a random variable and $Y = 2X + 1$. What is the variance of Y if variance of X is 5 ?

Exercise 6.3

Choose the correct Answer

1. Value which is obtained by multiplying possible values of random variable with probability of occurrence and is equal to weighted average is called

   (a) Discrete value (b) Weighted value

   (c) Expected value (d) Cumulative value

2. Demand of products per day for three days are 21, 19, 22 units and their respective probabilities are $0\cdot 29$, $0\cdot 40$, $0\cdot 35$. Pofit per unit is $0\cdot 50$ paisa then expected profits for three days are

   (a) 21, 19, 22 (b) 21.5, 19.5, 22.5 (c) 0.29, 0.40, 0.35 (d) 3.045, 3.8, 3.85

3. Probability which explains x is equal to or less than particular value is classified as

   (a) discrete probability (b) cumulative probability

   (c) marginal probability (d) continuous probability

4. Given $E(X) = 5$ and $E(Y) = -2$, then $E(X - Y)$ is

   (a) 3 (b) 5 (c) 7 (d) -2

5. A variable that can assume any possible value between two points is called

   (a) discrete random variable (b) continuous random variable

   (c) discrete sample space (d) random variable

6. A formula or equation used to represent the probability distribution of a continuous random variable is called

   (a) probability distribution (b) distribution function

   (c) probability density function (d) mathematical expectation

7. If X is a discrete random variable and $p(x)$ is the probability of X, then the expected value of this random variable is equal to

   (a) $\sum f(x)$ (b) $\sum\left[x+f(x)\right]$ (c) $\sum f(x)+x$ (d) $\sum xp(x)$

8. Which of the following is not possible in probability distribution?

(a) $\sum p(x) \geq 0$ (b) $\sum p(x) = 1$ (c) $\sum x\, p(x) = 2$ (d) $p(x) = -0.5$

9. If c is a constant, then $E(c)$ is

(a) 0 (b) 1 (c) $c\, f(c)$ (d) c

10. A discrete probability distribution may be represented by

(a) table (b) graph

(c) mathematical equation (d) all of these

EJSI8Z

11. A probability density function may be represented by:

(a) table (b) graph

(c) mathematical equation (d) both (b) and (c)

12. If c is a constant in a continuous probability distribution, then $p(x = c)$ is always equal to

(a) zero (b) one (c) negative (d) does not exist

13. $E\left[X - E(X)\right]$ is equal to

(a) $E(X)$ (b) $V(X)$ (c) 0 (d) $E(X) - X$

14. $E\left[X - E(X)\right]^2$ is

(a) $E(X)$ (b) $E(X^2)$ (c) $V(X)$ (d) $S.D(X)$

15. If the random variable takes negative values, then the negative values will have

(a) positive probabilities (b) negative probabilities

(c) constant probabilities (d) difficult to tell

16. If we have $f(x) = 2x, 0 \leq x \leq 1$, then $f(x)$ is a

(a) probability distribution (b) probability density function

(c) distribution function (d) continuous random variable

17. $\int_{-\infty}^{\infty} f(x)dx$ is always equal to

(a) zero (b) one (c) $E(X)$ (d) $f(x)+1$

18. A listing of all the outcomes of an experiment and the probability associated with each outcome is called

(a) probability distribution (b) probability density function

(c) attributes (d) distribution function

19. Which one is not an example of random experiment?

(a) A coin is tossed and the outcome is either a head or a tail

(b) A six-sided die is rolled

(c) Some number of persons will be admitted to a hospital emergency room during any hour.

(d) All medical insurance claims received by a company in a given year.

20. A set of numerical values assigned to a sample space is called

(a) random sample (b) random variable

(c) random numbers (d) random experiment

21. A variable which can assume finite or countably infinite number of values is known as

(a) continuous (b) discrete (c) qualitative (d) none of them

22. The probability function of a random variable is defined as

$X = x$	-1	-2	0	1	2
$P(x)$	k	2k	3k	4k	5k

Then k is equal to

(a) zero (b) $\frac{1}{4}$ (c) $\frac{1}{15}$ (d) one

23. If $p(x)=\frac{1}{10}$, $x = 10$, then $E(X)$ is

(a) zero (b) $\frac{6}{8}$ (c) 1 (d) -1

24. A discrete probability function $p(x)$ is always

(a) non-negative (b) negative (c) one (d) zero

25. In a discrete probability distribution the sum of all the probabilities is always equal to

(a) zero (b) one (c) minimum (d) maximum

26. An expected value of a random variable is equal to it's

(a) variance (b) standard deviation (c) mean (d) covariance

27. A discrete probability function $p(x)$ is always non-negative and always lies between

(a) 0 and ∞ (b) 0 and 1 (c) -1 and $+1$ (d) $-\infty$ and $+\infty$

28. The probability density function $p(x)$ cannot exceed

(a) zero (b) one (c) mean (d) infinity

29. The height of persons in a country is a random variable of the type

(a) discrete random variable (b) continuous random variable

(c) both (a) and (b) (d) neither (a) nor (b)

30. The distribution function $F(x)$ is equal to

(a) $P(X=x)$ (b) $P(X \le x)$ (c) $P(X \ge x)$ (d) all of these

Miscellaneous Problems

1. The probability function of a random variable X is given by

$$p(x)=\begin{cases} \frac{1}{4}, & for\ x=-2 \\ \frac{1}{4}, & for\ x=0 \\ \frac{1}{2}, & for\ x=10 \\ 0, & elsewhere \end{cases}$$

Evaluate the following probabilities.

(i) $P(X \le 0)$, (ii) $P(X<0)$, (iii) $P(|X| \le 2)$ and (iv) $P(0 \le X \le 10)$

2. Let X be a random variable with cumulative distribution function

$$F(x)=\begin{cases} 0, \; if \; x<0 \\ \frac{x}{8}, if \; 0\le x<1 \\ \frac{1}{4}+\frac{x}{8}, if \; 1\le x<2 \\ \frac{3}{4}+\frac{x}{12}, if \; 2\le x<3 \\ 1, \; for \; 3\le x. \end{cases}$$

(a) Compute: (i) $P(1\le X\le 2)$ and (ii) $P(X=3)$.

(b) Is X a discrete random variable? Justify your answer.

3. The p.d.f. of X is defined as

$$f(x)=\begin{cases} k, & for \; 0<x\le 4 \\ 0, & otherwise \end{cases}$$

Find the value of k and also find $P(2\le X\le 4)$.

4. The probability distribution function of a discrete random variable X is

$$f(x)=\begin{cases} 2k, & x=1 \\ 3k, & x=3 \\ 4k, & x=5 \\ 0, & otherwise \end{cases}$$

where k is some constant. Find (a) k and (b) $P(X>2)$.

5. The probability density function of a continuous random variable X is

$$f(x)=\begin{cases} a+bx^2, 0\le x\le 1; \\ 0, \; otherwise. \end{cases}$$

where a and b are some constants. Find (i) a and b if $E(X)=\frac{3}{5}$ (ii) $Var(X)$.

6. Prove that if $E(X) = 0$, then $V(X) = E\left(X^2\right)$.

7. What is the expected value of a game that works as follows: I flip a coin and, if tails pay you ₹ 2; if heads pay you ₹ 1. In either case I also pay you ₹ 50.

8. Prove that, (i) $V(aX)=a^2V(X)$, and (ii) $V(X+b)=V(X)$

9. Consider a random variable X with p.d.f

$$f(x)=\begin{cases}3x^2, \text{ if } 0<x<1\\ 0, \text{ otherwise}\end{cases}$$

Find $E(X)$ and $V(3X-2)$.

10. The time to failure in thousands of hours of an important piece of electronic equipment used in a manufactured DVD player has the density function

$$f(x)=\begin{cases}2e^{-2x}, \ x>0\\ 0, \ \text{otherwise}\end{cases}$$

Find the expected life of this piece of equipment.

Summary

- A variable which can assume finite number of possible values or an infinite sequence of countable real numbers is called a discrete random variable.
- Probability mass function (p.m.f.)

$$P_X(x)=\text{p}(\text{x})=\begin{cases}\text{P}(\text{X}=\text{x}_i)=p_i=p(x_i) & \text{if } x=x_i, i=1,2,\ldots,n,\ldots\\ 0 & \text{if } x\neq x_i\end{cases}$$

Conditions:

  - $p(x_i)\geq 0\,\forall i$ and
  - $\sum_{i=1}^{\infty} p(x_i)=1$

- Discrete distribution function (d.f.):

$$F_X(x)=P\left(X\leq x\right) \text{ for all } x\in R$$

$$\text{i.e., } F_X(x)=\sum_{x_i\leq x} p(x_i)$$

- A random variable X which can take on any value (integral as well as fraction) in the interval is called continuous random variable.
- Probability density function (p.d.f.)

The probability that a random variable X takes a value in the (open or closed) interval $[t_1,t_2]$ is given by the integral of a function called the probability density function $f_X(x)$

$$P\left(t_1\leq X\leq t_2\right)=\int_{t_1}^{t_2} f_X(x)dx$$

Other names that are used instead of probability density function include density function, continuous probability function, integrating density function.

Conditions:

- $f(x) \geq 0 \, \forall \, x$
- $\int_{-\infty}^{\infty} f(x)dx = 1$

- Continuous distribution function

If X is a continuous random variable with the probability density function $f_X(x)$, then the function $F_X(x)$ is defined by

$$F_X(x) = P[X \leq x] = \int_{-\infty}^{x} f_X(t)dt, -\infty < x < \infty$$

is called the distribution function (d.f) or sometimes the cumulative distribution function (c.d.f) of the random variable X.

- Properties of cumulative distribution function (c.d.f.)

The function $F_X(X)$ or simply $F(X)$ has the following properties

(i) $0 \leq F(x) \leq 1, -\infty < x < \infty$

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

(iii) $F(\cdot)$ is a monotone, non-decreasing function; that is, $F(a) \leq F(b)$ for $a < b$.

(iv) $F(\cdot)$ is continuous from the right; that is, $\lim_{h \to 0} F(x+h) = F(x)$.

(v) $F'(x) = \dfrac{d}{dx}F(x) = f(x) \geq 0$

(vi) $F'(x) = \dfrac{d}{dx}F(x) = f(x) \Rightarrow dF(x) = f(x)dx$

(vii) $dF(x)$ is known as probability differential of X.

(viii) $$\begin{aligned} P(a \leq x \leq b) &= \int_{a}^{b} f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx \\ &= P(X \leq b) - P(X \leq a) \\ &= F(b) - F(a) \end{aligned}$$

- Mathematical Expectation

  The expected value is a weighted average of the values of a random variable may assume.

- Discrete random variable with probability mass function (p.m.f.)

  $$E(X) = \sum_x x\, p(x)$$

- Continuous random variable with probability density function

  $$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

- The mean or expected value of X, denoted by μ_X or E(X).
- The variance is a weighted average of the squared deviations of a random variable from its mean.
- $Var(X) = \sum \left[x - E(X)\right]^2 p(x)$

  if X is discrete random variable with probability mass function $p(x)$.

- $Var(X) = \int_{-\infty}^{\infty} \left[X - E(X)\right]^2 f_X(x)\, dx$

  if X is continuous random variable with probability density function $f_X(x)$.

- Expected value of $[X - E(X)]^2$ is called the variance of the random variable.

  i.e., $Var(X) = E\left[X - E(X)\right]^2 = E\left(X^2\right) - \left[E(X)\right]^2$

  where $E\left(X^2\right) = \begin{cases} \sum_x x^2\, p(x),\ if\ X\ is\ Discrete\ Random\ Variable \\ \int_{-\infty}^{\infty} x^2 f(x)\, dx,\ if\ X\ is\ Continuous\ Random\ Variable \end{cases}$

- If X is a random variable, the standard deviation of X, denoted by σ_X, is defined as $+\sqrt{Var[X]}$.
- The variance of X, denoted by $\sigma_X{}^2$ or $Var(X)$ or $V(X)$.

- Properties of Mathematical expectation

  (i) $E(a) = a$, where 'a' is a constant

  (ii) $E(aX) = aE(X)$

  (iii) $E(aX + b) = aE(X) + b$, where 'a' and 'b' are constants.

  (iv) If $X \geq 0, then\ E(X) \geq 0$

  (v) $V(a) = 0$

  (vi) If X is random variable, then $V(aX + b) = a^2 V(X)$

- Raw moments

$$\mu'_r = E\left(X^r\right) = \begin{cases} \sum_x x^r p(x), & \textit{for discrete} \\ \int_{-\infty}^{\infty} x^r f(x)dx, & \textit{for continuous} \end{cases}$$

- Central Moments

$\mu_r = E[(X-\mu_X)^r]$

$\mu'_1 = E(X) = \mu_X$, the mean of X.

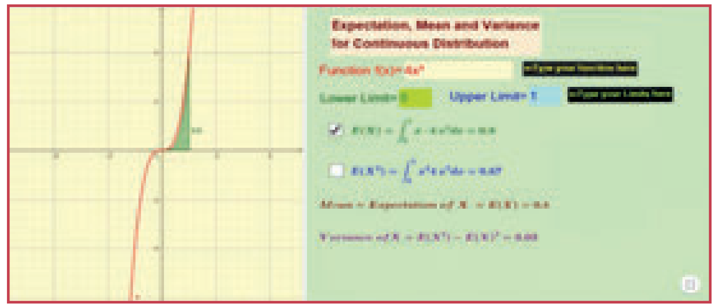
$\mu_1 = E[X-\mu_X] = 0.$

$\mu_2 = E[(X-\mu_X)^2]$, the variance of X.

GLOSSARY	
Absolutely Convergent	முற்றிலும் குவிதல் இயல்புடைய
Biased	நடுநிலையற்ற
Central moments	மையநிலை விலக்கப் பெருக்கம்
Continuous distribution function	தொடர்ச்சியான பரவல் சார்பு
Continuous random variable	தொடர்ச்சியான சமவாய்ப்பு மாறி
Cumulative	திரள், குவிந்த
Discrete distribution function	தொடர்ச்சியற்ற பரவல் சார்பு
Discrete random variable	தொடர்ச்சியற்ற சமவாய்ப்பு மாறி
Distribution function	பரவல் சார்பு
Event	நிகழ்வு, நிகழ்ச்சி
Expectation	எதிர்பார்த்தல்
Expected value	எதிர்பார்க்கத்தக்க மதிப்பு / எதிர்பார்த்தல் மதிப்பு
Mathematical expectation	கணக்கியல் எதிர்பார்த்தல்
Mean	சராசரி
Moments	விலக்கப் பெருக்கங்கள்
Probability function	நிகழ்தகவுச் சார்பு
Probability mass function	நிகழ்தகவு நிறைச் சார்பு
Random variable	சமவாய்ப்பு மாறி
Standard deviation	திட்ட விலக்கம், நியமச்சாய்வு
Unbiased	நடுநிலையான
Urn	குடுவை, கலசம்
Variance	மாறுபாட்டு அளவை / பரவர்படி
Weighted average	நிறைச் சராசரி

ICT Corner

Expected Result is shown in this picture

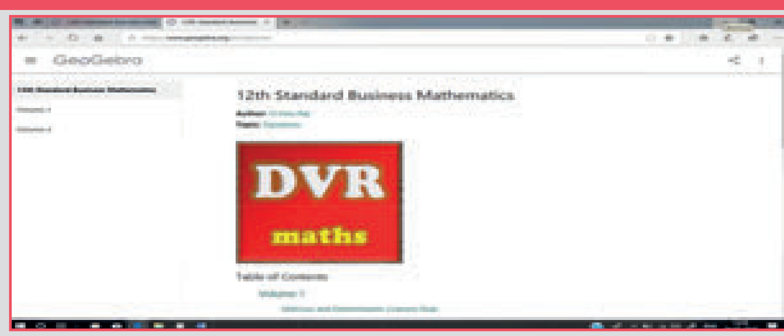


Step – 1 : Open the Browser, type the URL Link given below (or) Scan the QR Code. GeoGebra work Book named "12th Standard Business Mathematics " will open. In the work book there are two Volumes. Select "Volume-2".

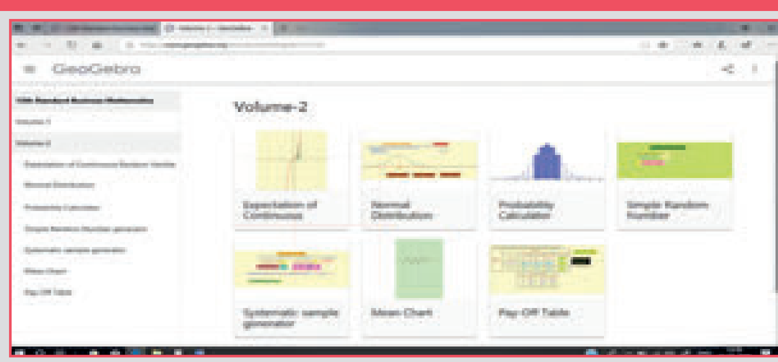
Step - 2 : Select the worksheet named "Expectation of continuous Random Variable"

You can type the function and the limits in the respective boxes to get E(X), Mean and Variance for the given Probability Distribution. Calculate yourself and check your answer.

Step 1



Step 2



Scan the QR Code.

B284_12_BUS_MAT_EM

7 Probability Distributions

Johann Carl Friedrich Gauss
(April 30, 1777 – Feb. 23, 1855)

Introduction

The frequency distributions are of two types namely Observed frequency distribution and Theoretical frequency distribution. The distributions which are based on actual data or experimentation are called the Observed Frequency distribution. On the other hand, the distributions based on expectations on the basis of past experience are known as Theoretical Frequency distribution or Probability distribution.

Johann Carl Friedrich Gauss (30 April 1777 – 23 February 1855) was a German mathematician and physicist who made significant contributions to many fields in mathematics and sciences. Gauss had an exceptional influence in many fields of mathematics and science, and is ranked among history's most influential mathematicians

In discrete probability distribution we will discuss Binomial and Poisson distribution and the Normal Distribution is a continuous probability distribution

Learning Objectives

After studying this chapter students are able to understand

- Concept of Bernoulli trial
- Binomial , poisson and normal density function
- Mean and variance of binomial and poisson distribution
- Properties of normal probability curve

KI3XKX

7.1 Distribution

The following are the two types of Theoretical distributions :

1. Discrete distribution   2. Continous distribution

Discrete distribution

The binomial and Poisson distributions are the most useful theoretical distributions for discrete variables.

7.1.1 Binomial distribution

Binomial distribution was discovered by James Bernoulli(1654_1705) in the year 1700 and was first published posthumously in 1713 , eight years after his death.

A random experiment whose outcomes are of two types namely success S and failure F, occurring with probabilities p and q respectively, is called a Bernoulli trial.

Some examples of Bernoulli trials are :

(i) Tossing of a coin (Head or tail)

(ii) Throwing of a die (getting even or odd number)

Consider a set of n independent Bernoullian trails (n being finite) in which the probability 'p' of success in any trial is constant , then $q = 1-p$, is the probability of failure. The probability of x successes and consequently $(n-x)$ failures in n independent trials, in a specified order (say) SSFSFFFS….FSF is given in the compound probability theorem by the expression

P(SSFSFFFS…..FSF) = P(S)P(S)P(F)P(S)x…….xP(F)P(S)P(F)

$$p.p.qp\ldots\ldots\ldots\ldots\ldots\ldots q.p.q$$

$$\underbrace{p.p.\,p.p.}_{\{x \text{ factors}\}} \qquad \underbrace{\ldots\ldots q.q.q.q.q.q}_{\{(n-x) \text{ factors}\}}$$

$$p^x q^{(n-x)}$$

x successes in n trials can occur in ${}^nC_x$ ways and the probability for each of these ways is same namely $p^x q^{n-x}$.

The probability distribution of the number of successes, so obtained is called the binomial probability distribution and the binomial expansion is $(q + p)^n$

Definition 7.1

A random variable X is said to follow binomial distribution with parameter n and p, if it assumes only non- negative value and its probability mass function in given by

$$P\,(x = x) = p(x) = \begin{cases} {}^nC_x p^x q^{n-x} x = 0,1,2,\ldots..n; q = 1-p \\ \qquad 0, otherwise \end{cases}$$

Note

Any random variable which follows binomial distribution is known as binomial variate i.e $X \sim B(n,p)$ is a binomial variate.

The Binomial distribution can be used under the following conditions :

1. The number of trials 'n' finite
2. The trials are independent of each other.
3. The probability of success 'p' is constant for each trial.
4. In every trial there are only two possible outcomes – success or failure.

Derivation of the Mean and Variance of Binomial distribution :

The mean of the binomial distribution $E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$

$$= p \sum_{x=1}^{n} x.\left(\frac{n}{x}\right)\binom{n-1}{x-1} p^{x-1} q^{n-x}$$

$$= np\ (q+p)n{-}1 \qquad [\text{since } p+q = 1]$$

$$= np$$

$$E(X) = np$$

$\therefore$ The mean of the binomial distribution is np

$$\text{Var}(X) = E(X^2) - E(X)^2$$

Here $E(X^2) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x}$

$$\sum_{x=0}^{n} \{x(x-1)+x\} \binom{n}{x} p^x q^{n-x}$$

$$\sum_{x=0}^{n} \{x(x-1)\} \binom{n}{x} p^x q^{n-x} + \sum x \binom{n}{x} p^x q^{n-x}$$

$$\sum_{x=2}^{n} \{x(x-1)\} \left(\frac{n(n-1)}{x(x-1)}\right)\binom{n-2}{x-2} p^{x-2} q^{n-x}\} + \sum x \binom{n}{x} p^x q^{n-x}$$

$$= n(n\text{-}1)\ p^2\ \{ \sum \binom{n-2}{x-2} p^{x-2} q^{n-x}\} \ + np$$

$$= n(n-1)p^2(q+p)(n-2) + np$$

$$= n(n-1)p^2 + np$$

Note

Mean and variance in terms of raw moments and central moments are denoted as μ'_1, and μ_2 respectively.

$$\therefore \quad \text{Variance} = E(X^2) - E(X)^2$$
$$= n^2p^2 - np^2 + np - n^2p^2$$
$$= np(1-p) = npq$$

Hence, mean of the BD is np and the Variance is npq.

Properties of Binomial distribution

1. Binomial distribution is symmetrical if $p = q = 0.5$. It is skew symmetric if $p \neq q$. It is positively skewed if $p < 0.5$ and it is negatively skewed if $p > 0.5$

2. For Binomial distribution, variance is less than mean

$$\text{Variance } npq = (np)q < np$$

Example 7.1

A and B play a game in which their chance of winning are in the ratio 3:2 Find A's chance of winning atleast three games out of five games played.

Soltion:

Let 'p' be the probability that 'A' wins the game. Then we are given $n = 5$, $p = 3/5$, $q = 1-\frac{3}{5} = \frac{2}{5}$ (since $q = 1-p$)

Hence by binomial probability law, the probability that out of the 5 games played, A wins 'x' games is given by

$$P(X{=}x) = p(x) = 5Cx\left(\frac{3}{5}\right)^x\left(\frac{2}{5}\right)^{5-x}$$

The required probability that 'A' wins atleast three games is given by

$$P(X \geq 3) = P(\mathrm{X} = 3) + P(\mathrm{X} = 4) + P(\mathrm{X} = 5)$$

$$= 5C3\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^2 + 5C4\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right)^1 + 5C5\left(\frac{3}{5}\right)^5\left(\frac{2}{5}\right)^o$$

$$= 0.6826$$

Example 7.2

A fair coin is tossed 6 times. Find the probability that exactly 2 heads occurs.

Solution :

Let X be a random variable follows binomial distribution with probability value $p = 1/2$ and $q = 1/2$

Prob: that exactcy 2 heads occur are as follows

$$P(X=2) = \binom{6}{x} p^x q^{n-x}$$
$$= \binom{6}{2} (\frac{1}{2})^2 \left(\frac{1}{2}\right)^{6-2}$$
$$= \frac{15}{64}$$

Example 7.3

Verfy the following statement:

The mean of a Binomial distribution is 12 and its standard deviation is 4.

Solution:

Mean: $np = 12$

$$SD = \sqrt{npq} = 4$$

$npq = 4^2 = 16$, $\frac{np}{npq} = \frac{12}{16} = \frac{3}{4}$ which implies

$$q = 4/3 > 1$$

Since $p + q$ cannot be greater than unity, the Statement is wrong

Example 7.4

The probability that a student get the degree is 0.4 Determine the probability that out of 5 students (i) one will be graduate (ii) atleast one will be graduate

Solution:

Probability of getting a degree $p = 0.4$

$$\therefore \quad q = 1 - p$$
$$= 1 - 0.4$$
$$= 0.6$$

(i) P (one will be a graduate) $= P(X = 1) = 5C_1\ (0.4)(0.6)^4$

$$= 0.2592$$

(ii) P (atleast one will be a graduate) = $1 - P$ (none will be a graduate)

$$= 1 - 5C_0 (P^0)(q)^{5-0}$$
$$= 1 - 5C_0 (0.4)^0\ (0.6)^5$$
$$= 1 - 0.0777$$
$$= 0.9222$$

Example 7.5

In tossing of a five fair coin, find the chance of getting exactly 3 heads.

Solution :

Let X be a random variable follows binomial distribution with $p = q = 1/2$

$$P\text{ (3 heads)} = 5C_x \left(\tfrac{1}{2}\right)^x \left(\tfrac{1}{2}\right)^{5-x}$$

$$= 5C_3 \left(\tfrac{1}{2}\right)^3 \left(\tfrac{1}{2}\right)^{5-3}$$

$$= 5C_3 \left(\tfrac{1}{2}\right)^5$$

$$= \tfrac{5}{16}$$

Example 7.6

The mean of Binomials distribution is 20 and standard deviation is 4. Find the parameters of the distribution.

Solution

The parameters of Binomial distribution are n and p

For Binomial distribution Mean = $np = 20$

Standard deviation = $\sqrt{npq} = 4$

$$\therefore \quad npq = 16$$

$$\Rightarrow \quad npq/np = 16/20 = 4/5$$

$$q = \tfrac{4}{5}$$

$$\Rightarrow \quad p = 1-q = 1-(4/5) = 1/5$$

Since $\quad np = 20$

$$n = \frac{20}{p}$$

$$n = 100$$

Example 7.7

If x is a binomially distributed random variable with $E(x) = 2$ and van $(x) = \tfrac{4}{3}$. Find $P(x=5)$

Solution:

The p.m.f. Binomial distribution is

$p(x) = {}^nC_x p^x \quad q^{n-x}$

Given that $E(x) = 2$

For the Binomial distribution mean is given by $np = 2$ ... (1)

Given that var $(x) = 4/3$

For Binomial distribution variance is given by $npq = \frac{4}{3}$ ... (2)

$$\frac{(2)}{(1)} = \frac{\frac{4}{3}}{2} = \frac{4}{6} = \frac{2}{3}$$

$q = 2/3$ and $p = 1-2/3 = 1/3$

Substitute is (1) we get

$n = 6$

Hence, $P(X=5) = 6C_5 \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^{6-5} = 0.0108$

Example 7.8

If on the average rain falls on 9 days in every thirty days, find the probability that rain will fall on atleast two days of a given week.

Solution :

Probability of raining on a particular day is given by $p = 9/30 = 3/10$ and

$q = 1-p = 7/10$.

The BD is $P(X = x) = {}^nC_x p^x q^{n-x}$

There are 7 days in a week $P(X = x) = \binom{7}{x}\left(\frac{3}{10}\right)^x \left(\frac{7}{10}\right)^{7-x}$

The probability of raining for atleast 2 days is given by

$$P(X \geq 2) = 1 - P(X < 2)$$

$$= 1 - \left[P(X=0) + P(X=1)\right]$$

Here, $P(X=0) = \binom{7}{0}\left(\frac{3}{10}\right)^0\left(\frac{7}{10}\right)^{7-0}$

$= 0.0823$

and $P(X=1) = \binom{7}{1}\left(\frac{3}{10}\right)\left(\frac{7}{10}\right)^{7-1}$

$= 0.2471$

Therefore the required probability $= 1 - [P(x=0) + P(x=1)]$

$= 1 - \{0.082 + 0.247]$

$= 0.6706$

Example 7.9

What is the probability of guessing correctly atleast six of the ten answers in a TRUE/FALSE objective test?

Solution :

Probability p of guessing an answer correctly is $p = \frac{1}{2}$

$\Rightarrow \quad q = \frac{1}{2}$

Probability of guessing correctly x answers in 10 questions

$$P(X=x) = p(x) = {}^nC_x\, p^x q^{n-x} \quad = \quad 10Cx\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}$$

The required probability $P(X \geq 6) \;= P(6) + P(7) + P(8) + P(9) + P(10)$

$$= \left(\frac{1}{2}\right)^{10}\left[10C_6 + 10C_7 + 10C_8 + 10C_9 + 10C_{10}\right]$$

$$= \left[\frac{1}{1024}\right]\left[210 + 120 + 45 + 10 + 1\right]$$

$$= \frac{193}{512}$$

Example 7.10

If the chance of running a bus service according to schedule is 0.8, calculate the probability on a day schedule with 10 services : (i) exactly one is late (ii) atleast one is late

Solution :

Probability of bus running late is denoted as $p = 1\text{-}0.8 = 0.2$

Probability of bus running according to the schedule is $q = 0.8$

Also given that $n = 10$

The binomial distribution is $p(x) = 10C_x\left(0.2\right)^x\left(0.8\right)^{10-x}$ t

(i) probability that exactly one is late $P(y=1) = 10C_1\, p q^9$

$$= 10C_1\left(0.2\right)\left(0.8\right)^9$$

(ii) probability that at least one is late

$= 1 -$ probability that none is late

$= 1 - p(x=0)$

$= 1 - (0.8)^{10}$

Example 7.11

The sum and product of the mean and variance of a binomial distribution are 24 and 128. Find the distribution.

Solution:

For Binomial Distribution the mean is np and varaiance is npq

Given values are $np + npq = 24 \quad np(1 + q) = 24 \; - \; (1)$

Other term $np \times npq = 128 \quad n^2p^2q = 128 \; - \; (2)$

From (1) we get $np = 24/(1+q)$ which implies $n^2p^2 = (24/(1+q))2$

Substitute this value in equation (2) we get

$$\left(\frac{24}{1+q}\right)^2 q = 128 \qquad \text{which implies } 9q = 2(1+2q+q^2)$$

$(2q - 1)(q - 2) = 0$

Where $q = \frac{1}{2}$ and $p = \frac{1}{2}$

Substitute in (1) we get $n = 32$

Hence the binomial distribution $32C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{32-x}$

Example 7.12

Suppose A and B are two equally strong table tennis players. Which of the following two events is more probable:

(a) A beats B exactly in 3 games out of 4 or

(b) A beats B exactly in 5 games out of 8 ?

Solution :

Here $p = q = 1/2$

(a) probability of A beating B in exactly 3 games out of 4

$$\binom{4}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{4-3}$$

$= \frac{1}{4} = 25\%$

(b) probability of A beating B in exactly 5 games out of 8

$$\binom{8}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^{8-5}$$

$= \frac{7}{32} = 21.875\%$

Clearly, the first event is more probable.

Example 7.13

A pair of dice is thrown 4 times. If getting a doublet is considered a success, find the probability of 2 successes.

Solution :

In a throw of a pair of dice the doublets are (1,1) (2,2) (3,3) (4,4) (5,5) (6,6)

Probability of getting a doublet $p = 6/36 = 1/6$

$\Rightarrow q = 1 - p = 5/6$ and also $n = 4$ is given

The probabitliy of successes $= \binom{4}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{4-x}$

Therefore the probability of 2 successes are

$$P(X = 2) = \binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{4-2}$$
$$= 6 \times \frac{1}{36} \times \frac{25}{36}$$
$$= \frac{25}{216}$$

Exercise 7.1

1. Define Binomial distribution.
2. Define Bernoulli trials.
3. Dedrive the mean and variance of binomial distribution.
4. Write down the conditions for which the binomial distribution can be used.
5. Mention the properties of binomial distribution.
6. If 5% of the items produced turn out to be defective, then find out the probability that out of 10 items selected at random there are

   (i) exactly three defectives

   (ii) atleast two defectives

   (iii) exactly 4 defectives

   (iv) find the mean and variance
7. In a particular university 40% of the students are having news paper reading habit. Nine university students are selected to find their views on reading habit. Find the probability that

   (i) none of those selected have news paper reading habit

   (ii) all those selected have news paper reading habit

   (iii) atleast two third have news paper reading habit.
8. In a family of 3 children, what is the probability that there will be exactly 2 girls?
9. Defects in yarn manufactured by a local mill can be approximated by a distribution with a mean of 1.2 defects for every 6 metres of length. If lengths of 6 metres are to be inspected, find the probability of less than 2 defects.
10. If 18% of the bolts produced by a machine are defective, determine the probability that out of the 4 bolts chosen at random

(i) exactly one will be defective

(ii) none will be defective

(iii) atmost 2 will be defective

11. If the probability of success is 0.09, how many trials are needed to have a probability of atleast one success as 1/3 or more ?

12. Among 28 professors of a certain department, 18 drive foreign cars and 10 drive local made cars. If 5 of these professors are selected at random, what is the probability that atleast 3 of them drive foreign cars?

13. Out of 750 families with 4 children each, how many families would be expected to have (i) atleast one boy (ii) atmost 2 girls (iii) and children of both sexes? Assume equal probabilities for boys and girls.

14. Forty percent of business travellers carry a laptop. In a sample of 15 business travelers,

    (i) what is the probability that 3 will have a laptop?

    (ii) what is the probability that 12 of the travelers will not have a laptop?

    (iii) what is the probability that atleast three of the travelers have a laptop?

15. A pair of dice is thrown 4 times. If getting a doublet is considered a success, find the probability of 2 successes.

16. The mean of a binomial distribution is 5 and standard deviation is 2. Determine the distribution.

17. Determine the binomial distribution for which the mean is 4 and variance 3. Also find P(X=15)

18. Assume that a drug causes a serious side effect at a rate of three patients per one hundred. What is the probability that atleast one person will have side effects in a random sample of ten patients taking the drug?

19. Consider five mice from the same litter, all suffering from Vitamin A deficiency. They are fed a certain dose of carrots. The positive reaction means recovery from the disease. Assume that the probability of recovery is 0.73. What is the probability that atleast 3 of the 5 mice recover.

20. An experiment succeeds twice as often as it fails, what is the probability that in next five trials there will be (i) three successes and (ii) at least three successes

7.1.2 POISSON DISTRIBUTION

Poisson distribution was derived in 1837 by a French Mathematician Simeon D. Poisson. If n is large,, the evaluation of the binomial probabilities can involve complex computations, in such a case, a simple approximation to the binomial probabilities could be use. Such approximation of binomial when n is large and p is close to zero is called the Poisson distribution.

Poisson distribution occurs when there are events which do not occur as a definite number on trials but an events occurs rarely and the following examples may be analysed:

(i) Number of bacteria in one cubic centimeter.

(ii) Number of printing mistakes per page in a text book

(iii) the number of alpha particles emitted by a radioactive substance in a fraction of a second.

(iv) Number of road accidents occurring at a particular interval of time per day.

(v) Number of lightnings per second.

Poisson distribution is a limiting case of binomial distribution under the following conditions :

(i) n, the number of trials is indefinitely large i.e $n \to \infty$.

(ii) p, the constant probability of success in each trial is very small, i.e. $p \to 0$

(iii) $np = \lambda$ is finite. Thus $p = \frac{\lambda}{n}$ and $q = 1 - \left(\frac{\lambda}{n}\right)$ where λ is a positive real number.

Definition 7.2

A random variable X is said to follow a Poisson distribution with parameter λ if it assumes only non-negative values and its probability mass function is given by

$$P(x,\lambda) = P(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!}, & x = 0,1,2\ldots\ldots;\lambda > 0 \\ 0 & otherwise \end{cases}$$

Derivation of Mean and variance of Poisson distribution

$$\text{Mean } E(X) = \sum_{x=0}^{\infty} x\, p(x,\lambda)$$

$$= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda e^{-\lambda}\left\{\sum\left(\frac{\lambda^{x-1}}{(x-1)1!}\right)\right\}$$

$= \lambda e - \lambda(1 + \lambda + \lambda 2/2! + \ldots\ldots)$

$= \lambda e - \lambda\ e\lambda$

$= \lambda$

Variance $(X) = E(X^2) - E(X)^2$

Here $\quad E(X^2) = \sum_{x=0}^{\infty} x^2 p(x,\lambda)$

$$= \sum_{x=0}^{\infty} x^2 p(x,\lambda)$$

$$= \sum_{x=0}^{\infty}\{x(x-1)+x\}\,p(x,\lambda)$$

$$= \sum_{x=0}^{\infty}\{x(x-1)+x\}\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\Sigma x(x-1)\ \lambda x/x! \ + \Sigma\ x\ e^{-\lambda}\ \frac{\lambda^x}{x!}$$

$$= \lambda^2\ e^{-\lambda}\ \sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} + \lambda$$

$$= \lambda^2\ e^{-\lambda}\ e^{\lambda} + \lambda = \lambda^2 + \lambda$$

Variance $(X) = E(X^2) - E(X)^2$

$= \lambda^2 + \lambda - (\lambda)^2$

$= \lambda$

Properties of Poisson distribution :

1. Poisson distribution is the only distribution in which the mean and variance are equal .

Example 7.14

In a Poisson distribution the first probability term is 0.2725. Find the next Probability term

Solution :

Given that $p(0) = 0.2725$

$$\frac{e^{-\lambda}\lambda^0}{0!} = 0.2725$$

$$\Rightarrow e^{-\lambda} = 0.2725 \quad \text{(by using exponent table)}$$

$$\lambda = 1.3$$

$$\therefore p(X=1) = e^{-1.3}(1.3)/1!$$

$$= e^{-1.3}(1.3)$$

$$= 0.2725 \text{ x } 1.3$$

$$= 0.3543$$

Example 7.15

In a book of 520 pages, 390 typo-graphical errors occur. Assuming Poisson law for the number of errors per page, find the probability that a random sample of 5 pages will contain no error.

Solution :

The average number of typographical errors per page in the book is given by $\lambda =$ (390/520) = 0.75.

Hence using Poisson probability law, the probability of x errors per page is given by

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = e^{-0.75}\frac{(0.75)^x}{x!}, \; x = 0,1,2,3\ldots\ldots$$

The required probability that a random sample of 5 pages will contain no error is given by : $[P(X=0)]^5 = \left(e^{-0.75}\right)^5 = e^{-3.75}$

Example 7.16

An insurance company has discovered that only about 0.1 per cent of the population is involved in a certain type of accident each year. If its 10,000 policy holders were randomly selected from the population, what is the probability that not more than 5 of its clients are involved in such an accident next year? ($e^{-10} = .000045$)

Solution :

p = probability that a person will involve in an accident in a year

= 0.1/100 = 1/1000

given n = 10,000

so, $\lambda = np = 10000\left(\frac{1}{1000}\right) = 10$

Probability that not more than 5 will involve in such an accident in a year

$$P(X \leq 5) = P(X{=}0)+P(X{=}1)+P(X{=}2)+P(X{=}3)+P(X{=}4)+P(X{=}5)$$

$$= e^{-10}\left[\ 1+\frac{10}{1!}+\frac{10^2}{2!}+\frac{10^3}{3!}+\frac{10^4}{4!}+\frac{10^5}{5!}\right]$$

$$= 0.06651$$

Example 7.17

One fifth percent of the the blades produced by a blade manufacturing factory turn out to be defective. The blades are supplied in packets of 10. Use Poisson distribution to calculate the approximate number of packets containing no defective, one defective and two defective blades respectively in a consignment of 1,00,000 packets ($e^{-0.2}$ =.9802)

Solution :

P = 1/5/100 = 1/500 =0.002 $\quad n = 10\quad .\quad \lambda = np = 0.02$

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.02}(0.02)^x}{x!}$$

(i) Number of packets containing no defective $= N\,p(o) = 1{,}00{,}000 \times e^{-0.02}$

$= 98020$

(ii) Number of packets containing one defective $= N\,p(1) = 1{,}00{,}000 \times 0.9802 \times 0.02$

$= 1960$

(iii) Number of packets containing 2 defectives $= N\,p(2) = 20$

Example 7.18

If the probability that an individual suffers a bad reaction from injection of a given serum is 0.001, determines the probability that out of 2,000 individuals (a) exactly 3, and (b) more than 2 individuals will suffer a bad reaction.

Solution :

Consider a 2,000 individuals getting injection of a given serum , $n = 2000$

Let X be the number of individuals suffering a bad reaction.

Let p be the probability that an individual suffers a bad reaction = 0.001

and $q = 1 - p = 1 - 0.001 = 0.999$

Since n is large and p is small, Binomial Distribtuion approximated to poisson distribution

So, $\lambda = np = 2000 \times 0.001 = 2$

(i) Probability out of 2000, exactly 3 will suffer a bad reaction is

$$P(X=3) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-2}2^3}{3!} = 0.1804$$

(ii) Probability out of 2000, more than 2 individuals will suffer a bad reaction

$$= P(X > 2)$$

$$= 1 - [(P(X \leq 2)]$$

$$= 1 - [P(x=0) + P(x=1) + P(x=2)]$$

$$= 1 - \left[\frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!}\right]$$

$$= 1 - e^{-2}\left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!}\right)$$

$$= 0.323$$

Example 7.19

When counting red blood cells, a square grid is used, over which a drop of blood is evenly distributed. Under the microscope an average of 8 erythrocytes are observed per single square. What is the probability that exactly 5 erythrocytes are found in one square?

Solution :

Let X be a random variable follows poisson distribution with number of erythrocytes.

Hence, Mean $\lambda = 8$ erythrocytes/single square

$$P(\text{exactly 5 erythrocytes are in one square}) = P(X=5) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-8}8^5}{5!}$$

$$= \frac{0.000335 \times 32768}{120}$$

$$= 0.0916$$

The probability that exactly 5 erythrocytes are found in one square is 0.0916. i.e there are 9.16% chances that exactly 5 erythrocytes are found in one square.

Example 7.20

Assuming one in 80 births is a case of twins, calculate the probability of 2 or more sets of twins on a day when 30 births occur.

Solution :

Let x devotes the set of twins on a day

P(twin birth) $= p = 1/80 = 0.0125$ and $n = 30$

The value of mean $\lambda = np = 30 \times 0.0125 = 0.375$

Hence, X follows poisson distribution with $p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$

The probability is

$$P(2 \text{ or more}) = 1 - [p\ (x = 0) + p\ (x =)] = 1 - [\frac{e^{-0.375}(0.375)^{0}}{0!} + \frac{e^{-0.375}(0.375)^{1}}{1!}]$$

$$= 1 - e^{-0.375}\ [1+0.375]$$

$$= 1 - (0.6873 \times 1.375)$$

$$= 0.055$$

Exercise 7.2

1. Define Poisson distribution.
2. Write any 2 examples for Poisson distribution.
3. Write the conditions for which the poisson distribution is a limiting case of binomial distribution.
4. Derive the mean and variance of poisson distribution.
5. Mention the properties of poisson distribution.
6. The mortality rate for a certain disease is 7 in 1000. What is the probability for just 2 deaths on account of this disease in a group of 400? Given $e^{(-2.8)} = 0.06$
7. It is given that 5% of the electric bulbs manufactured by a company are defective. Using poisson distribution find the probability that a sample of 120 bulbs will contain no defective bulb.

8. A car hiring firm has two cars. The demand for cars on each day is distributed as a Poisson variate, wjth mean 1.5. Calculate the proportion of days on which

(i) Neither car is used (ii) Some demand is refused

9. The average number of phone calls per minute into the switch board of a company between 10.00 am and 2.30 pm is 2.5. Find the probability that during one particular minute there will be (i) no phone at all (ii) exactly 3 calls (iii) atleast 5 calls

10. The distribution of the number of road accidents per day in a city is poisson with mean 4. Find the number of days out of 100 days when there will be (i) no accident (ii) atleast 2 accidents and (iii) at most 3 accidents.

11. Assuming that a fatal accident in a factory during the year is 1/1200, calculate the probability that in a factory employing 300 workers there will be atleast two fatal accidents in a year. (given $e^{-0.25} = 0.7788$)

12. The average number of customers, who appear in a counter of a certain bank per minute is two. Find the probability that during a given minute (i) No customer appears (ii) three or more customers appear .

Continuous distribtuion

The binomial and Poisson distributions discussed in the previous chapters are the most useful theoretical distributions for discrete variables. In order to have mathematical distributions suitable for dealing with quantities whose magnitudes vary continuously like weight, heights of individual, a continuous distribution is needed. Normal distribution is one of the most widely used continuous distribution.

Normal distribution is the most important and powerful of all the distribution in statistics. It was first introduced by De Moivre in 1733 in the development of probability. Laplace (1749-1827) and Gauss (1827-1855) were also associated with the development of Normal distribution.

7.1.3 NORMAL DITRIBUTION

Definition 7.3

A random variable X is said to follow a normal distribution with parameters mean μ and variance σ^2 , if its probability density function is given by

$$f(x:\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad \begin{array}{c} -\infty < x < \infty, \\ -\infty < \mu < \infty, \\ \sigma > 0 \end{array}$$

Normal distribution is diagrammatically represented as follows :

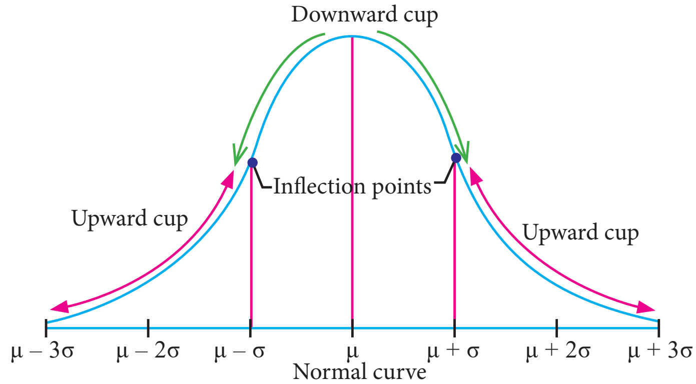


Fig 7.1

Normal distribution is a limiting case of Binomial distribution under the following conditions:

(i) n, the number of trials is infinitely large, i.e. $n \to \infty$

(ii) neither p (or q) is very small,

The normal distribution of a variable when represented graphically, takes the shape of a symmetrical curve, known as the Normal Curve. The curve is asymptotic to x-axis on its either side.

Chief Characterisitics or Properties of Normal Probabilty distribution and Normal probability Curve .

The normal probability curve with mean μ and standard deviation σ has the following properties :

(i) the curve is bell- shaped and symmetrical about the line x=u

(ii) Mean, median and mode of the distribution coincide.

(iii) x – axis is an asymptote to the curve. (tails of the cuve never touches the horizontal (x) axis)

(iv) No portion of the curve lies below the x-axis as $f(x)$ being the probability function can never be negative.

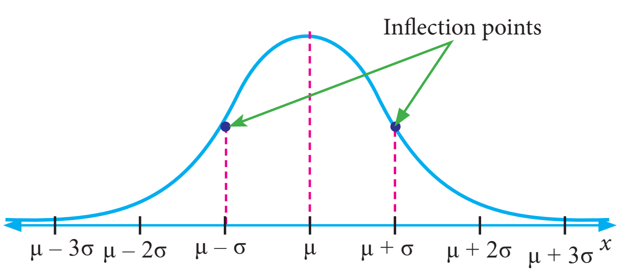


Fig 7.3

(v) The Points of inflexion of the curve are $x = \mu \pm \sigma$

(vi) The curve of a normal distribution has a single peak i.e it is a unimodal.

12th BM EM V-2.indb 63 30-03-2019 17:13:13

(vii) As x increases numerically, $f(x)$ decreases rapidly, the maximum probability occurring at the point $x = \mu$ and is given by $[p(x)]\max = 1/\sigma\sqrt{2\pi}$

(viii) The total area under the normal curve is equal to unity and the percentage distribution of area under the normal curve is given below

(a) About 68.27% of the area falls between $\mu - \sigma$ and $\mu + \sigma$

$P(\mu - \sigma < X < \mu + \sigma) = 0.6826$

(b) About 95.5% of the area falls between $\mu - 2\sigma$ and $\mu + 2\sigma$

$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9544$

(c) About 99.7% of the area falls between $\mu - 3\sigma$ and $\mu + 3\sigma$

$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9973$

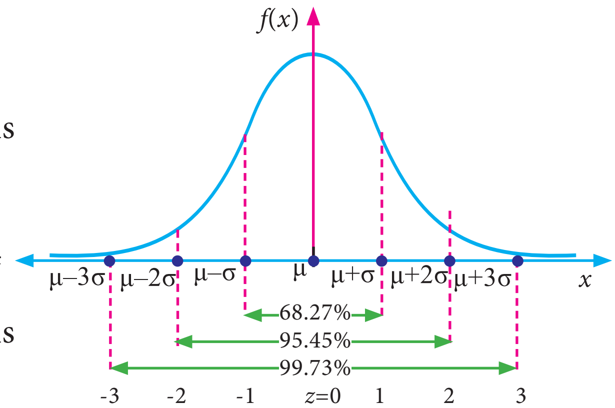


Fig 7.4

STANDARD NORMAL DISTRIBUTION

A random variable $Z = (X-\mu)/\sigma$ follows the standard normal distribution. Z is called the standard normal variate with mean 0 and standard deviation 1 i.e $Z \sim N(0,1)$. Its Probability density function is given by :

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty$$

1. The area under the standard normal curve is equal to 1.

2. 68.26% of the area under the standard normal curve lies between $z = -1$ and $Z = 1$

   95.44% of the area lies between $Z = -2$ and $Z = 2$

   99.74% of the area lies between $Z = -3$ and $Z = 3$

Example 7.21

What is the probability that a standard normal variate Z will be

(i) greater than 1.09

(ii) less than -1.65

(iii) lying between -1.00 and 1.96

(iv) lying between 1.25 and 2.75

Solution :

(i) greater than 1.09

The total area under the curve is equal to 1 , so that the total area to the right Z = 0 is 0.5 (since the curve is symmetrical). The area between Z = 0 and 1.09 (from tables) is 0.3621

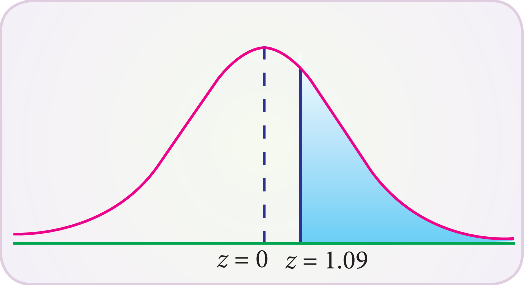


Fig 7.5

P(Z > 1.09) = 0.5000 - 0.3621 = 0.1379

The shaded area to the right of Z = 1.09 is the probability that Z will be greater than 1.09

(ii) less than −1.65

The area between -1.65 and 0 is the same as area between 0 and 1.65. In the table the area between zero and 1.65 is 0.4505 (from the table). Since the area to the left of zero is 0.5 , P(Z< 1.65) = 0.5000 – 0.4505 = 0.0495.

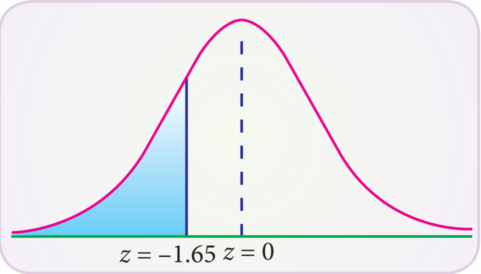


Fig 7.6

(iii) lying between -1.00 and 1.96

The probability that the random variable Z in between -1.00 and 1.96 is found by adding the corresponding areas :

Area between -1.00 and 1.96
= area between (-1.00 and 0) + area betwn (0 and 1.96)

$$P(-1.00 < Z < 1.96) = P(-1.00 < Z < 0) + P(0 < Z < 1.96)$$

$$= 0.3413 + 0.4750 \text{ (by tables)}$$

$$= 0.8163$$

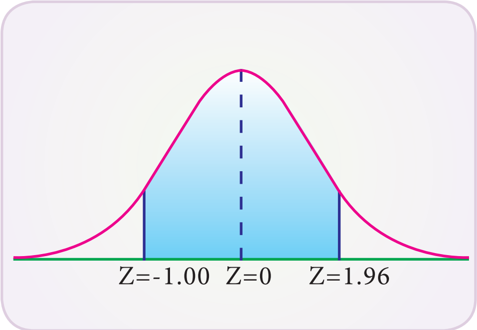


Fig 7.7

(iv) lying between 1.25 and 2.75

Area between Z = 1.25 and 2.75
= area betwn (z = 0 and z = 2.75)
– area betwn (z=0 and z = 1.25)

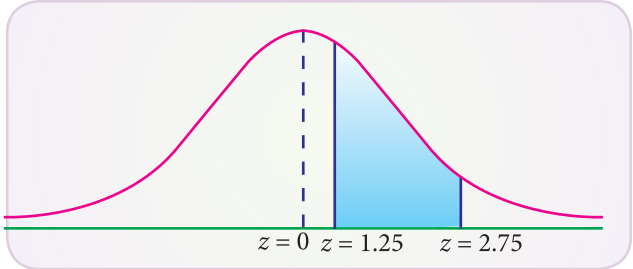


Fig 7.8

$P(1.25 < Z < 2.75) = P(0 < Z < 2.75) - P(0 < Z < 1.25)$

$= 0.4970 - 0.3944 = 0.1026$

Example 7.22

If X is a normal variate with mean 30 and SD 5. Find the probabilities that (i) $26 \le X \le 40$ (ii) $X > 45$

Solution :

Here mean μ= 30 and standard deviation σ = 5

(i) When $X = 26$ $\quad Z = (X - \mu)/\sigma = (26 - 30)/5 = -0.8$

And when $X = 40$, $\quad Z = \dfrac{40 - 30}{5} = 2$

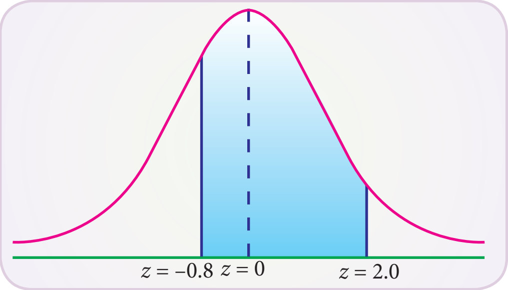


Fig 7.9

Therefore,

$P(26<X<40)) = P(-0.8 \le Z < 2)$

$= P(-0.8 \le Z \le 0) + p(0 \le Z \le 2)$

$= P(0 \le Z \le 0.8) + P(0 \le Z \le 2)$

$= 0.2881 + 0.4772$ (By tables)

$= 0.7653$

(ii) The probability that $X \ge 45$

When $X = 45$

$$Z = \frac{X - \mu}{\sigma} = \frac{45 - 30}{5} = 3$$

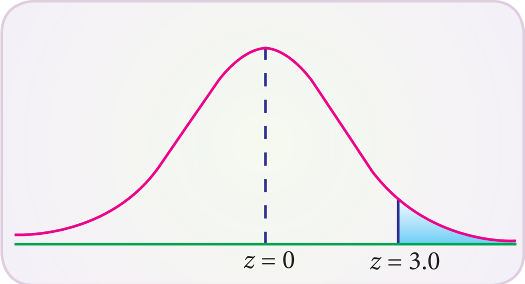


Fig 7.10

$P(X \ge 45) = P(Z \ge 3)$

$= 0.5 - 0.49865$

$= 0.00135$

Example 7.23

The average daily sale of 550 branch offices was Rs.150 thousand and standard deviation is Rs. 15 thousand. Assuming the distribution to be normal, indicate how many branches have sales between

(i) ₹ 1,25,000 and ₹ 1, 45, 000

(ii) ₹ 1,40,000 and ₹ 1,60,000

Solution :

Given that mean μ= 150 and standard deviation σ= 15

(i) when X = 125 thousand

$$Z = \frac{X-\mu}{\sigma} = \frac{125-150}{15} = -1.667$$

When X = 145 thousand

$$Z = \frac{X-\mu}{\sigma} = \frac{145-150}{15} = -0.33$$

Area between $Z = 0$ and $Z = -1.67$ is 0.4525

Area between $Z = 0$ and $Z = -0.33$ is 0.1293

$P(-1.667 \le Z \le -0.33) = 0.4525 - 0.1293$

$= 0.3232$

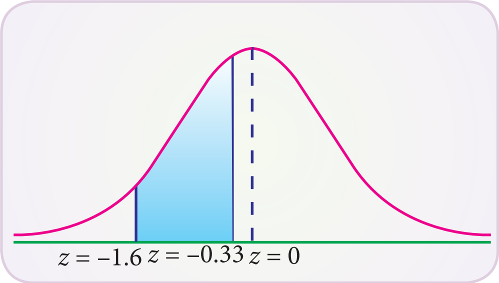


Fig 7.11

Therefore the number of branches having sales between 125 thousand and 145 thousand is 550 × 0.3232 = 178

(ii) When X = 140 thousand

$$Z = \frac{X-\mu}{\sigma} = \frac{140-150}{15} = -0.67$$

When X = 160 thousand

$$Z = \frac{X-\mu}{\sigma} = \frac{160-150}{15} = 0.67$$

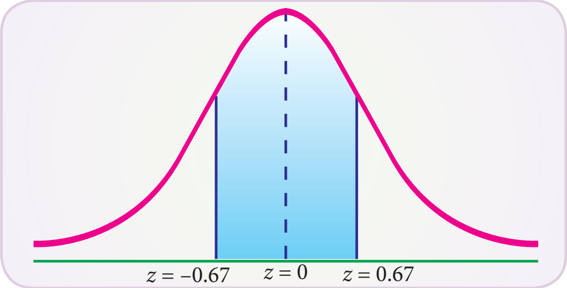


Fig 7.12

$P(-0.67 < Z < 0.67) = P(-0.67 < Z < 0) + P(0 < Z < 0.67)$

$= P(0 < Z < 0.67) + P(0 < Z < 0.67)$

$= 2 P(0 < Z < < 0.67)$

$= 2 \times 0.2486$

$= 0.4972$

Therefore, the number of branches having sales between Rs.140 thousand and Rs.160 thousand = 550 × 0.4972 = 273

Example 7.24

Assume the mean height of children to be 69.25 cm with a variance of 10.8 cm. How many children in a school of 1,200 would you expect to be over 74 cm tall?

Solution

Let the distribution of heights be normally distributed with mean mean 68.22 and standard deviation = 3.286

$$Z = \frac{X-\mu}{\sigma} = \frac{X-69.25}{3.286}$$

When $X = 74$

$$Z = \frac{X-\mu}{\sigma} = \frac{74-69.25}{3.286} = 1.4455$$

Now $P(Z > 74) = P(Z > 1.44)$

$= 0.5 - 0.4251$

$= 0.0749$

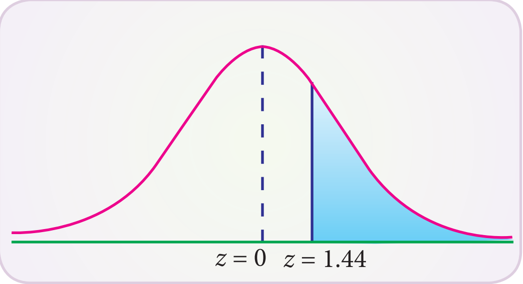


Fig 7.13

Expected number of children to be over 74 cm out of 1200 children

$= 1200 \times 0.0749 \approx 90$ children

Example 7.25

The marks obtained in a certain exam follow normal distribution with mean 45 and SD 10. If 1,300 students appeared at the examination, calculate the number of students scoring (i) less than 35 marks and (ii) more than 65 marks.

Solution :

Let X be the normal variate showing the score of the candidate with mean 45 and standard deviation 10.

(i) less than 35 marks

When $X = 35$

$$Z = \frac{X-\mu}{\sigma} = \frac{35-45}{10} = -1$$

$P(X < 35) = P(Z < -1)$

$P(Z > 1) = 0.5 - P(0 < Z < 1)$

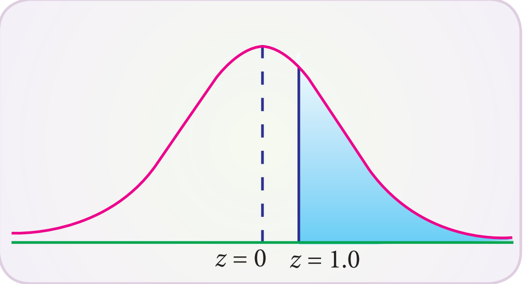


Fig 7.14

$= 0.5 - 0.3413$

$= 0.1587$

Expected number of students scoring less than 35 marks are 0.1587×1300

$= 206$

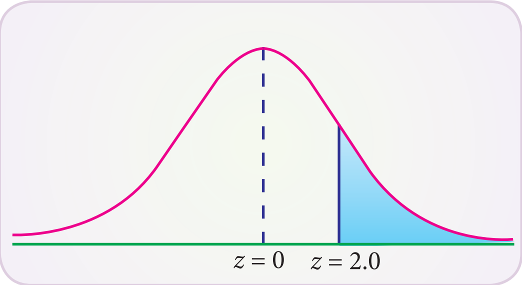


Fig 7.15

(ii) more than 65 marks

When $X = 65$

$$Z = \frac{X-\mu}{\sigma} = \frac{65-45}{10} = 2.0$$

$P(X > 65) = P(Z > 2.0)$

$0.5 - P(0 < Z < 2.0)$

$0.5 - 0.4772$

$= 0.0228$

Expected number of students scoring more than 65 marks are 0.0228 x 1300

$= 30$

Example 7.26

900 light bulbs with a mean life of 125 days are installed in a new factory. Their length of life is normally distributed with a standard deviation of 18 days. What is the expected number of bulbs expire in less than 95 days?

Solution :

Let X be the normal variate of life of light bulbs with mean 125 and standard deviation 18.

(i) less than 95 days

When $X = 95$

$$Z = \frac{X-\mu}{\sigma} = \frac{95-125}{18} = -1.667$$

$P(X < 95) = P(Z < -1.667)$

$= P(Z > 1.667)$

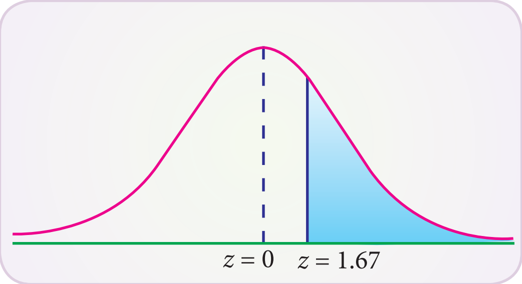


Fig 7.16

12th BM EM V-2.indb 69 30-03-2019 17:13:17

$$= 0.5 - P(0 < Z < 1.67)$$

$$= 0.5 - 0.4525$$

$$= 0.0475$$

No. of bulbs expected to expire in less than 95 days out of 900 bulbs

$$900 \times .0475 = 43 \text{ bulbs}$$

Example 7.27

Assume that the mean height of soldiers is 69.25 inches with a variance of 9.8 inches. How many soldiers in a regiment of 6,000 would you expect to be over 6 feet tall?

Solution :

Let X be the height of soldiers follows normal distribution with mean 69.25 inches and standard deviation 3.13. then the soldiers over 6 feet tall (6ft × 12= 72 inches)

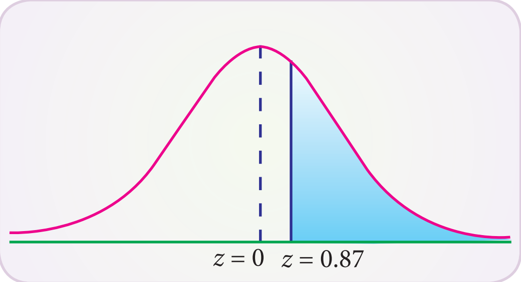


Fig 7.17

The standard normal variate

$$Z = \frac{X-\mu}{\sigma} = \frac{72-69.25}{3.13} = 0.8786$$

$$P(X > 72) = P(Z > 0.8786) = 0.5 - P(0 < Z < 0.88) = 0.5 - 0.3106 = 0.1894$$

Number of soldiers expected to be over 6 feet tall in 6000 are

$$6000 \times 0.1894 = 1136$$

Example 7.28

A bank manager has observed that the length of time the customers have to wait for being attended by the teller is normally distributed with mean time of 5 minutes and standard deviation of 0.6 minutes. Find the probability that a customer has to wait

(i) for less than 6 minutes

(ii) between 3.5 and 6.5 minutes

Solution :

Let X be the waiting time of a customer in the queue and it is normally distributed with mean 5 and SD 0.7.

(i) for less than 6 minutes

$$Z = \frac{X-\mu}{\sigma} = \frac{6-5}{0.7} = 1.4285$$

$$P(X < 6) = P(Z < 1.43)$$

$$= 0.5 + 0.4236$$

$$= 0.9236$$

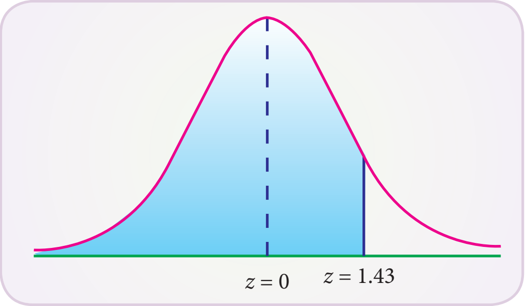


Fig 7.18

(ii) between 3.5 and 6.5 minutes

When $X = 3.5$

$$Z = \frac{X-\mu}{\sigma} = \frac{3.5-5}{0.7} = -2.1429$$

When $X = 6.5$

$$Z = \frac{X-\mu}{\sigma} = \frac{6.5-5}{0.7} = 2.1429$$

$P(3.5 < X < 6.5)$

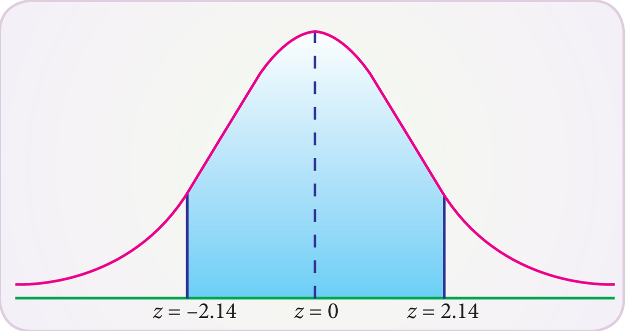


Fig 7.19

$$= P(-2.1429 < Z < 2.1429)$$

$$= P(0 < Z < 2.1429) + P(0 < Z < 2.1429)$$

$$= 2\,P(0 < Z < 2.1429)$$

$$= 2 \times .4838$$

$$= 0.9676$$

Example 7.29

A sample of 125 dry battery cells tested to find the length of life produced the following resultd with mean 12 and sd 3 hours. Assuming that the data to be normal distributed , what percentage of battery cells are expected to have life

(i) more than 13 hours

(ii) less than 5 hours

(iii) between 9 and 14 hours

Solution :

Let X denote the length of life of dry battery cells follows normal distribution with mean 12 and sd 3 hours

(i) more than 13 hours

$P(X > 13)$

When $X = 13$

$$Z = \frac{X-\mu}{\sigma} = \frac{13-12}{3} = 0.333$$

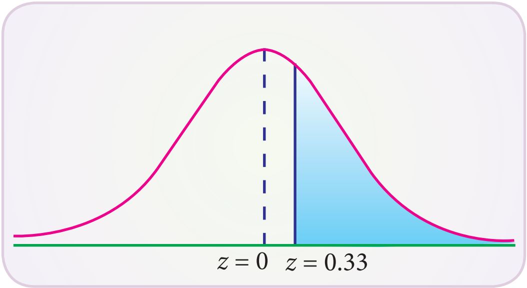


Fig 7.20

$P(X > 13) = P(Z>0.333) = 0.5 - 0.1293 = 0.3707$

The expected battery cells life to have more than 13 hours is

$125 \times 0.3707 = 46.34\%$

(ii) less than 5 hours

$P(X < 5)$

When $X = 5$ $\quad Z = \frac{X-\mu}{\sigma} = \frac{5-12}{3} = -2.333$

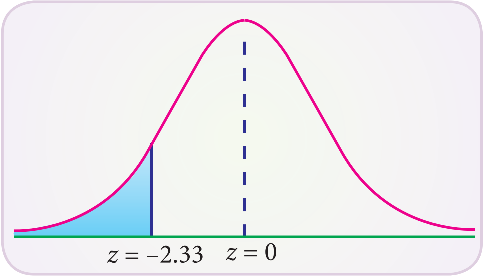


Fig 7.21

$P(X < 5) = P(Z < -2.333) = P(Z > 2.333)$

$= 0.5 - 0.4901 = 0.0099$

The expected battery cells life to have more than 13 hours is

$125 \times 0.0099 = 1.23\%$

(iii) between 9 and 14 hours

When $X = 9$

$$Z = \frac{X-\mu}{\sigma} = \frac{9-12}{3} = -1$$

When $X = 14$

$$Z = \frac{X-\mu}{\sigma} = \frac{14-12}{3} = 0.667$$

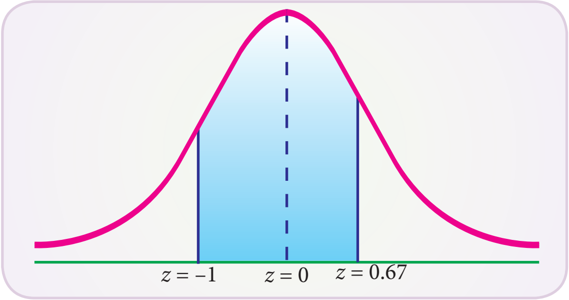


Fig 7.22

$P(9 < X < 14) = P(-1 < Z < 0.667)$

$= P(0 < Z < 1) + P(0 < Z < 0.667)$

$= 0.3413 + 0.2486$

$= 0.5899$

The expected battery cells life to have more than 13 hours is

125 x 0.5899 =73.73%

Example 7.30

Weights of fish caught by a traveler are approximately normally distributed with a mean weight of 2.25 kg and a standard deviation of 0.25 kg. What percentage of fish weigh less than 2 kg?

Solution :

We are given mean $\mu = 2.25$ and standard deviation $\sigma = 0.25$. Probability that weight of fish is less than 2 kg is $P(X < 2.0)$

When $x = 20$ $\quad Z = \dfrac{X-\mu}{\sigma} = \dfrac{2.0-2.25}{0.25} = P(Z< -1.0) = P(Z > 1.0)$

$$0.5 - 0.3413 = 0.1587$$

Therefore 15.87% of fishes weigh less than 2 kg.

Example 7.31

The average daily procurement of milk by village society in 800 litres with a standard deviation of 100 litres. Find out proportion of societies procuring milk between 800 litres to 1000 litres per day.

Solution :

We are given mean $\mu = 800$ and standard deviation $\sigma = 100$. Probability that the procurement of milk between 800 litres to 1000 litres per day is

$$P(800 < X < 1000)$$

$$P(\frac{800-800}{100} < z < \frac{1000-800}{100}$$

$$P(0 < Z < 2) = 0.4772 \text{ (table value)}$$

Therefore 47.75 percent of societies procure milk between 800 litres to 1000 litres per day.

Exercise 7.3

1. Define Normal distribution.
2. Define Standard normal variate.
3. Write down the conditions in which the Normal distribution is a limiting case of binomial distribution.
4. Write down any five chief characteristics of Normal probability curve.

5. In a test on 2,000 electric bulbs, it was found that bulbs of a particular make, was normally distributed with an average life of 2,040 hours and standard deviation of 60 hours. Estimate the number of bulbs likely to burn for (i) more than 2,150 hours (ii) less than 1,950 hours (iii) more 1,920 hours but less than 2,100 hours.

6. In a distribution 30% of the items are under 50 and 10% are over 86. Find the mean and standard deviation of the distribution.

7. X is normally distributed with mean 12 and sd 4. Find $P(X \leq 20)$ and $P(0 \leq X \leq 12)$

8. If the heights of 500 students are normally distributed with mean 68.0 inches and standard deviation 3.0 inches , how many students have height (a) greater than 72 inches (b) less than or equal to 64 inches (c) between 65 and 71 inches.

9. In a photographic process, the developing time of prints may be looked upon as a random variable having the normal distribution with a mean of 16.28 seconds and a standard deviation of 0.12 second. Find the probability that it will take less than 16.35 seconds to develop prints.

10. Time taken by a construction company to construct a flyover is a normal variate with mean 400 labour days and standard deviation of 100 labour days. If the company promises to construct the flyover in 450 days or less and agree to pay a penalty of ₹ 10,000 for each labour day spent in excess of 450. What is the probability that

    (i) the company pays a penalty of atleast ₹ 2,00,000?

    (ii) the company takes at most 500 days to complete the flyover?

RGECXV

Exercise 7.4

Choose the correct Answer

1. Normal distribution was invented by

   (a) Laplace (b) De-Moivre (c) Gauss (d) all the above

2. If X ~N(9,81) the standard normal variate Z will be

   (a) $Z = \dfrac{X-81}{9}$ (b) $Z = \dfrac{X-9}{81}$ (c) $Z = \dfrac{X-9}{9}$ (d) $Z = \dfrac{9-X}{9}$

3. If Z is a standard normal variate, the proportion of items lying between $Z = -0.5$ and $Z = -3.0$ is

   (a) 0.4987 (b) 0.1915 (c) 0.3072 (d) 0.3098

4. If X ~N(μ, σ^2), the maximum probability at the point of inflexion of normal distribution is

   (a) $\left(\dfrac{1}{\sqrt{2\pi}}\right)e^{\frac{1}{2}}$ (b) $\left(\dfrac{1}{\sqrt{2\pi}}\right)e^{\left(-\frac{1}{2}\right)}$ (c) $\left(\dfrac{1}{\sigma\sqrt{2\pi}}\right)e^{\left(-\frac{1}{2}\right)}$ (d) $\left(\dfrac{1}{\sqrt{2\pi}}\right)$

5. In a parametric distribution the mean is equal to variance is :

(a) binomial (b) normal (c) poisson (d) all the above

6. In turning out certain toys in a manufacturing company, the average number of defectives is 1%. The probability that the sample of 100 toys there will be 3 defectives is

(a) 0.0613 (b) 0.613 (c) 0.00613 (d) 0.3913

7. The parameters of the normal distribution $f(x)=\left(\frac{1}{\sqrt{72\pi}}\right)\frac{e^{-(x-10)^2}}{72}$ $-\infty < x < \infty$

(a) (10,6) (b) (10,36) (c) (6,10) (d) (36,10)

8. A manufacturer produces switches and experiences that 2 per cent switches are defective. The probability that in a box of 50 switches, there are atmost two defective is :

(a) $2.5\ e^{-1}$ (b) e^{-1} (c) $2\ e^{-1}$ (d) none of the above

9. An experiment succeeds twice as often as it fails. The chance that in the next six trials, there shall be at least four successes is

(a) 240/729 (b) 489/729 (c) 496/729 (d) 251/729

10. If for a binomial distribution b(n,p) mean = 4 and variance = 4/3, the probability, $P(X \geq 5)$ is equal to :

(a) $(2/3)^6$ (b) $(2/3)^5(1/3)$ (c) $(1/3)^6$ (d) $4(2/3)^6$

11. The average percentage of failure in a certain examination is 40. The probability that out of a group of 6 candidates atleast 4 passed in the examination are :

(a) 0.5443 (b) 0.4543 (c) 0.5543 (d) 0.4573

12. Forty percent of the passengers who fly on a certain route do not check in any luggage. The planes on this route seat 15 passengers. For a full flight, what is the mean of the number of passengers who do not check in any luggage?

(a). 6.00 (b.) 6.45 (c). 7.20 (d.) 7.50

13. Which of the following statements is/are true regarding the normal distribution curve?

(a) it is symmetrical and bell shaped curve

(b) it is asymptotic in that each end approaches the horizontal axis but never reaches it

(c) its mean, median and mode are located at the same point

(d) all of the above statements are true.

14. Which of the following cannot generate a Poisson distribution?

(a) The number of telephone calls received in a ten-minute interval

(b) The number of customers arriving at a petrol station

(c) The number of bacteria found in a cubic feet of soil

(d) The number of misprints per page

15. The random variable X is normally distributed with a mean of 70 and a standard deviation of 10. What is the probability that X is between 72 and 84?

(a) 0.683 (b) 0.954 (c) 0.271 (d) 0.340

16. The starting annual salaries of newly qualified chartered accountants (CA's) in South Africa follow a normal distribution with a mean of ₹ 180,000 and a standard deviation of ₹ 10,000. What is the probability that a randomly selected newly qualified CA will earn between ₹ 165,000 and ₹ 175,000 per annum?

(a) 0.819 (b) 0.242 (c) 0.286 (d) 0.533

17. In a large statistics class the heights of the students are normally distributed with a mean of 172cm and a variance of 25cm. What proportion of students are between 165cm and 181cm in height?

(a) 0.954 (b) 0.601 (c) 0.718 (d) 0.883

18. A statistical analysis of long-distance telephone calls indicates that the length of these calls is normally distributed with a mean of 240 seconds and a standard deviation of 40 seconds. What proportion of calls lasts less than 180 seconds?

(a) 0.214 (b) 0.094 (c) 0933 (d) 0.067

19. Cape town is estimated to have 21% of homes whose owners subscribe to the satelite service, DSTV. If a random sample of your home in taken, what is the probability that all four home subscribe to DSTV?

(a) 0.2100 (b) 0.5000 (c) 0.8791 (d) 0.0019

20. Using the standard normal table, the sum of the probabilities to the right of $z = 2.18$ and to the left of $z = -1.75$ is:

(a) 0.4854 (b) 0.4599 (c) 0.0146 (d) 0.0547

21. The time until first failure of a brand of inkjet printers is normally distributed with a mean of 1,500 hours and a standard deviation of 200 hours. What proportion of printers fails before 1000 hours?

(a) 0.0062 (b) 0.0668 (c) 0.8413 (d) 0.0228

12th BM EM V-2.indb 76 30-03-2019 17:13:26

22. The weights of newborn human babies are normally distributed with a mean of 3.2kg and a standard deviation of 1.1kg. What is the probability that a randomly selected newborn baby weighs less than 2.0kg?

(a) 0.138 (b) 0.428 (c) 0.766 (d) 0.262

23. Monthly expenditure on their credit cards, by credit card holders from a certain bank, follows a normal distribution with a mean of ₹ 1,295.00 and a standard deviation of ₹ 750.00. What proportion of credit card holders spend more than ₹ 1,500.00 on their credit cards per month?

(a) 0.487 (b) 0.392 (c) 0.500 (d) 0.791

24. Let z be a standard normal variable. If the area to the right of z is 0.8413, then the value of z must be:

(a) 1.00 (b) –1.00 (c) 0.00 (d) –0.41

25. If the area to the left of a value of z (z has a standard normal distribution) is 0.0793, what is the value of z?

(a) -1.41 (b) 1.41 (c) –2.25 (d) 2.25

26. If $P(Z > z) = 0.8508$ what is the value of z (z has a standard normal distribution)?

(a) –0.48 (b) 0.48 (c) –1.04 (d) 1.04

27. If $P(Z > z) = 0.5832$ what is the value of z (z has a standard normal distribution)?

(a) –0.48 (b) 0.48 (c) 1.04 (d) –0.21

28. In a binomial distribution, the probability of success is twice as that of failure. Then out of 4 trials, the probability of no success is

(a) 16/81 (b) 1/16 (c) 2/27 (d) 1/81

Miscellaneous Problems

1. A manufacturer of metal pistons finds that on the average, 12% of his pistons are rejected because they are either oversize or undersize. What is the probability that a batch of 10 pistons will contain

   (a) no more than 2 rejects? (b) at least 2 rejects?

2. Hospital records show that of patients suffering from a certain disease 75% die of it. What is the probability that of 6 randomly selected patients, 4 will recover?

3. If electricity power failures occur according to a Poisson distribution with an average of 3 failures every twenty weeks, calculate the probability that there will not be more than one failure during a particular week.

4. Vehicles pass through a junction on a busy road at an average rate of 300 per hour.
   1. Find the probability that none passes in a given minute.
   2. What is the expected number passing in two minutes?
5. Entry to a certain University is determined by a national test. The scores on this test are normally distributed with a mean of 500 and a standard deviation of 100. Raghul wants to be admitted to this university and he knows that he must score better than at least 70% of the students who took the test. Raghul takes the test and scores 585. Will he be admitted to this university?
6. The time taken to assemble a car in a certain plant is a random variable having a normal distribution of 20 hours and a standard deviation of 2 hours. What is the probability that a car can be assembled at this plant in a period of time .
   a) less than 19.5 hours?
   b) between 20 and 22 hours?
7. The annual salaries of employees in a large company are approximately normally distributed with a mean of $50,000 and a standard deviation of $20,000.
   (a) What percent of people earn less than $40,000?
   (b) What percent of people earn between $45,000 and $65,000?
   (c) What percent of people earn more than $70,00
8. X is a normally normally distributed variable with mean $\mu = 30$ and standard deviation $\sigma = 4$. Find
   (a) $P(x < 40)$ (b) $P(x > 21)$ (c) $P(30 < x < 35)$
9. The birth weight of babies is Normally distributed with mean 3,500g and standard deviation 500g. What is the probability that a baby is born that weighs less than 3,100g?

People's monthly electric bills in chennai are normally distributed with a mean of ₹ 225 and a standard deviation of ₹ 55. Those people spend a lot of time online. In a group of 500 customers, how many would we expect to have a bill that is ₹ 100 or less?

Summary

- Conditions for the binomial probability distribution are
  (i) the trials are independent
  (ii) the number of trials is finite
  (iii)each trial has only two possible outcomes called success and failure.
  (iv) the probability of success in each trial is a constant.

- The probability for exactly x success in n independent trials is given by

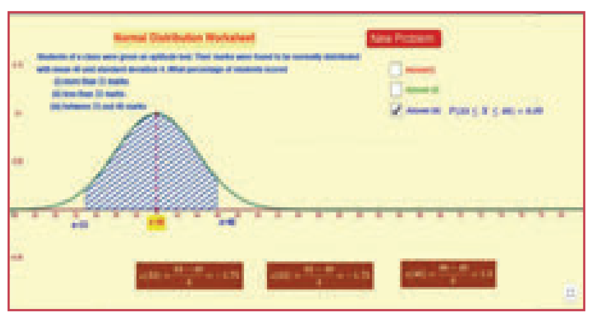
  $p(x) = \binom{n}{x} p^x q^{n-x}$ where $x = 0,1,2,3\ldots..n$ and $q = 1 - p$
- The parameters of the binomial distributions are n and p
- The mean of the binomial distribution is np and variance are npq
- Poisson distribution as limiting form of binomial distribution when n is large, p is small and np is finite.
- The Poisson probability distribution is $p(x) = e^{-\lambda}\dfrac{\lambda^x}{x!}$ X = 0,1,2,3…

  Where $\lambda = np$
- The mean and variance of the poisson distribtution is λ.
- The λ is the only parameter of poisson distribution.
- Poisson distribution can never be symmetrical.
- It is a distribution for rare events.
- Normal distribution is the limiting form of binomial distribution when n is large and neither p nor q is small
- The normal probability distribution is given by $f(x) = \left(\dfrac{1}{\sigma\sqrt{2\pi}}\right)(e^{-1/2(x-\mu/\sigma)^2})$
- The mean of the distribution is μ
- The sd of the distribution is σ.
- It is a symmetrical distribution
- The graph of the distribution is bell shaped
- In normal distribution the mean, median and mode are equal
- The points of inflexion are $\mu - \sigma$ and $\mu + \sigma$
- The normal curve approaches the horizontal axis asymptotically
- Area Property : In a normal distribution about 68% of the item will lie between $\mu - \sigma$ and $\mu + \sigma$. About 95% will lie between are $\mu -2\sigma$ and $\mu + 2\sigma$. About 99% will lie between $\mu -3\sigma$ and $\mu +3\sigma$.
- Standard normal random variate is denoted as $Z = (X - \mu)/\sigma$
- The standard normal probability distribution is $1/\sqrt{2\pi}(e^{\frac{-z^2}{2}})$
- The mean of the distribution is zero and SD is unity
- The points of inflexion are at $z = -1$ and $z = +1$

GLOSSARY	
bell shaped curve	மணி வடிவ வளைவரை
Binomial	ஈருறுப்பு
continuous distribution	தொடர்ச்சியான பரவல்
discrete distribution	தனிநிலைப் பரவல் / தனித்த பரவல்
Distribution	பரவல்
Independent	சார்பற்ற
Normal	இயல்நிலை
parameter	பண்பளவை
point of inflexion	வளைவு மாற்றுப்புள்ளி
random experiment	சமவாய்ப்பு சோதனை
Random variable	இயைபிலா மாறி
Sample space	கூறுவெளி
Skewness	கோட்ட அளவை
Standard deviation	திட்ட விலக்கம்
Symmetry	சமச்சீர்

ICT Corner

Expected Result is shown in this picture

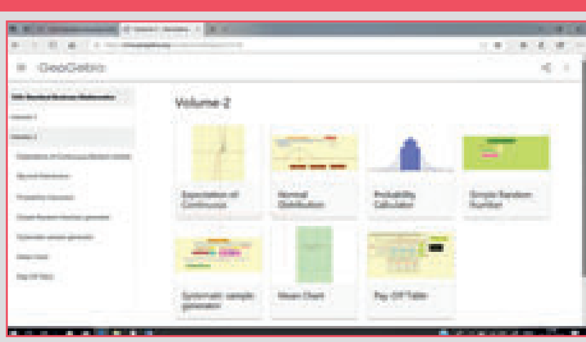


Step – 1 : Open the Browser, type the URL Link given below (or) Scan the QR Code. GeoGebra work Book named "12th Standard Business Mathematics " will open. In the work book there are two Volumes. Select "Volume-2".

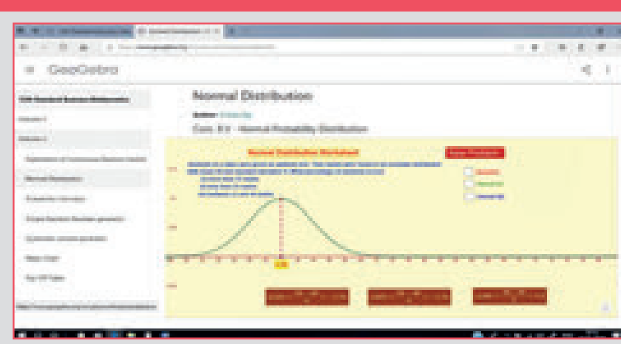
Step - 2 : Select the worksheet named "Normal Distribution"

Click on "New Problem" and click on Answer-I, Answer-ii and Answer -iii to see the answer for respective questions. Though the Normal Curve seems touching the bottom line, it don't touch when we zoom. Work out the answer and check yourself and analyse the shaded region.

Step 1



Step 2



Scan the QR Code.

B284_12_BUS_MAT_EM

8 Sampling Techniques and Statistical Inference

Pandurang Vasudeo Sukhatme
(July 27, 1911 - January 28, 1997)

Introduction

Pandurang Vasudeo Sukhatme (1911–1997) was an award-winning Indian statistician. He is known for his pioneering work of applying random sampling methods in agricultural statistics and in biometry, in the 1940s. He was also influential in the establishment of the Indian Agricultural Statistics Research Institute. As a part of his work at the Food and Agriculture Organization in Rome, he developed statistical models for assessing the dimensions of hunger and future food supplies for the world. He also developed methods for measuring the size and nature of the protein gap.

In any statistical investigation, the interest lies in the assessment of one or more characteristics relating to the individuals belonging to a group. When all the individuals present in the study are investigated, it is called complete enumeration, but in practice, it is very difficult to investigate all the individuals present in the study. So the technique of sampling is done which states that a part of the individuals are selected for the study and the assessment is made from the selected group of individuals. For example

(i) A housewife tastes a spoonful whatever she cooks to check whether it tastes good or not.

(ii) A few drops of our blood are tested to check about the presence or absence of a disease.

(iii) A grain merchant takes out a handful of grains to get an idea about the quality of the whole consignment.

These are typical examples where decision making is done on the basis of sample information. So sampling is the process of choosing a representative sample from a given population.

Learning Objectives

After studying this chapter students are able to understand

- sampling techniques
- random sampling

- simple random sampling
- stratified random sampling
- systematic sampling
- sampling and non-sampling errors
- sampling distribution
- statistical inference
- estimation
- test of statistical hypothesis

XEPSAT

8.1 Sampling

Sampling is the procedure or process of selecting a sample from a population. Sampling is quite often used in our day-to-day practical life.

8.1.1 Basic concepts of sampling

Population

The group of individuals considered under study is called as population. The word population here refers not only to people but to all items that have been chosen for the study. Thus in statistics, population can be number of bikes manufactured in a day or week or month, number of cars manufactured in a day or week or month, number of fans, TVs, chalk pieces, people, students, girls, boys, any manufacturing products, etc…

Finite and infinite population:

When the number of observations/individuals/products is countable in a group, then it is a finite population. Example: weights of students of class XII in a school.

When the number of observations/individuals/products is uncountable in a group, then it is an infinite population. Example: number of grains in a sack, number of germs in the body of a sick patient.

Sample and sample size

A selection of a group of individuals from a population in such a way that it represents the population is called as sample and the number of individuals included in a sample is called the sample size.

Parameter and statistic

Parameter: The statistical constants of the population like mean (μ), variance(σ^2) are referred as population parameters.

Statistic : Any statistical measure computed from sample is known as statistic.

Note

In practice, the parameter values are not known and their estimates based on the sample values are generally used.

Types of sampling

The technique or method of selecting a sample is of fundamental importance in the theory of sampling and usually depends upon the nature of the data and the type of enquiry. The procedures of selecting a sample may be broadly classified as

1. Non-Random sampling or Non-probability sampling.
2. Random Sampling or Probability sampling.

Note

Here we confine ourselves to Random sampling or Probability sampling

Random sampling or Probability sampling

Random sampling refers to selection of samples from the population in a random manner. A random sample is one where each and every item in the population has an equal chance of being selected.

" Every member of a parent population has had equal chances of being included".- Dr. Yates

"A random sample is a sample selected in such a way that every item in the population has an equal chance of being included".-Harper

The following are different types of probability sampling:

(i) Simple random sampling

(ii) Stratified random sampling

(iii) Systematic sampling

(i) Simple random sampling

In this technique the samples are selected in such a way that each and every unit in the population has an equal and independent chance of being selected as a sample. Simple random sampling may be done, with or without replacement of the samples selected. In a simple random sampling with replacement there is a possibility of selecting the same sample any number of times. So, simple random sampling without replacement is followed. Thus in simple random sampling from a population of N units, the probability of drawing any unit at the first draw is $\frac{1}{N}$, the probability of drawing any unit in the second draw from among the available $(N-1)$ units is $\frac{1}{(N-1)}$, and so on. . Several methods have been adopted for random selection of the samples from the population. Of those, the following two methods are generally used and which are described below.

(A) Lottery method

This is the most popular and simplest method when the population is finite. In this method, all the items of the population are numbered on separate slips of paper of same size, shape and colour. They are folded and placed in a container and shuffled thoroughly. Then the required numbers of slips are selected for the desired sample size. The selection of items thus depends on chance.

For example, if we want to select 10 students, out of 100 students, then we must write the names/roll number of all the 100 students on slips of the same size and mix them, then we make a blindfold selection of 10 students. This method is called unrestricted random sampling, because units are selected from the population without any restriction. This method is mostly used in lottery draws. If the population or universe is infinite, this method is inapplicable.

(B) Table of Random number

When the population size is large, it is difficult to number all the items on separate slips of paper of same size, shape and colour. The alternative method is that of using the table of random numbers. The most practical, easy and inexpensive method of selecting a random sample can be done through "Random Number Table". The random number table has been so constructed that each of the digits 0,1,2,…,9 will appear approximately with the same frequency and independently of each other.

The various random number tables available are

a. L.H.C. Tippett random number series

b. Fisher and Yates random number series

c. Kendall and Smith random number series

d. Rand Corporation random number series.

Tippett's table of random numbers is most popularly used in practice. Given below the first forty sets from Tippett's table as an illustration of the general appearance of random numbers:

2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

Suppose, if we want to select the required number of samples from a population of size N(<99) then any two digit random number can be selected from (00 to 99) from the above random number table. Similarly if N(< 999) or (< 9999), then any three digit random number or four digit random number can be selected from (000 to 999) or (0000 to 9999).

The procedure of selecting the random samples consists of following steps:

1. Identify the N units in the population with the numbers from 1 to N.
2. Select at random, any page from the 'Random Number Table'.
3. Select the required number of samples from any row or column or diagonal.

One may question, as to how it is ensured that these digits are random. It may be pointed out that the digits in the table were chosen horizontally but the real guarantee of their randomness lies in practical tests. Tippett's numbers have been subjected to numerous tests and used in many investigations and their randomness has been very well established for all practical purposes.

An example to illustrate how Tippett's table of random numbers may be used is given below. Suppose we have to select 20 items out of 6,000. The procedure is to number all the 6,000 items from 1 to 6,000. A page from Tippett's table may be selected and the first twenty numbers ranging from 1 to 6,000 are noted down. If the numbers are above 6000, choose the next number ranging from 1 to 6000. Items bearing those numbers will be selected as samples from the population. Making use of the portion of the random number table given,the requiredrandom samples are shaded. Here, we consider row wise selection of random numbers.

2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

If the population size is 1,000 and suppose we want to select 15 items out of 1,000. All itemsfrom 1 to 1000 should be numbered as 0001 to 1000.Now, we may now select 15 numbers from the random number table.The procedure will be different, as Tippett's random numbers are available only in four digits. Thus, we can select the first three digits from the four digit random sample number.Making use of the portion of the random number table given, the required random samples are shaded in RED colour. Here, we consider row wise selection of random numbers.

2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

If the population size is 100 and suppose we want to select 10 items out of 100. All itemsfrom 1 to 100 should be numbered as 001 to 100.Now, we may now select 10 numbers from the random number table. The procedure will be different, as Tippett's random numbers are available only in four digits. Thus, we can select the first two digits from the four digit random sample number. Making use of the portion of the random number table given, the required random samples are shaded in RED colour.Here, we consider row wise selection of random numbers.

2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

Merits and demerits of SimpleRandomSampling:

Merits

1. Personal bias is completely eliminated.
2. This method is economical as it saves time, money and labour.
3. The method requires minimum knowledge about the population in advance.

Demerits

1. This requires a complete list of the population but such up-to-date lists are not available in many enquiries.
2. If the size of the sample is small, then it will not be a representative of the population.

Example 8.1

Using the Kendall-Babington Smith - Random number table,Draw5 random samples.

23	15	75	48	59	01	83	72	59	93	76	24	97	08	86	95	23	03	67	44
05	54	55	50	43	10	53	74	35	08	90	61	18	37	44	10	96	22	13	43
14	87	16	03	50	32	40	43	62	23	50	05	10	03	22	11	54	36	08	34
38	97	67	49	51	94	05	17	58	53	78	80	59	01	94	32	42	87	16	95
97	31	26	17	18	99	75	53	08	70	94	25	12	58	41	54	88	21	05	13

Solution:

23	15	75	48	59	01	83	72	59	93	76	24	97	08	86	95	23	03	67	44
05	54	55	50	43	10	53	74	35	08	90	61	18	37	44	10	96	22	13	43
14	87	16	03	50	32	40	43	62	23	50	05	10	03	22	11	54	36	08	34
38	97	67	49	51	94	05	17	58	53	78	80	59	01	94	32	42	87	16	95
97	31	26	17	18	99	75	53	08	70	94	25	12	58	41	54	88	21	05	13

There many ways to select 5random samples from the given Kendall-Babington Smith - Random number table. Assume that at random we select 3rd column 1stvalue. This location gives the digit to be 75. So the first sample will be 75, and then the next choices can be in the same 3rd column which follows as 55, 16, 67 and 26. Therefore 75, 55,16,67 and 26 will be used as random samples. The various shaded numbers can be taken as 5 random sample numbers. Apart from this, one can select any 5 random sample numbers as they like.

Example 8.2

Using the following Tippett's random number table,

2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

Draw a sample of 15 houses from Cauvery Street which has 83 houses in total.

Solution:

There many ways to select 15 random samples from the given Tippet's random number table. Since the population size is 83(two-digit number). Here the door numbers are assigned from 1 to 83. Assume that at random we first choose 2nd column. So the first sample is 66 and other 14 samples are 74, 52, 39, 15, 34, 11, 14, 13, 27, 61, 79, 72, 35, and 60. If the numbers are above 83, choose the next number ranging from 1 to 83.

2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

Example 8.3

Using the following random number table,

Tippet's random number table							
2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

Draw a sample of 10 children with theirheight from the population of 8,585 children as classified hereunder.

Height (cm)	105	107	109	111	113	115	117	119	121	123	125
Number of children	2	4	14	41	83	169	394	669	990	1223	1329
Height(cm)	127	129	131	133	135	137	139	141	143	145	
No. of children	1230	1063	646	392	202	79	32	16	5	2	

Solution:

The first thing is to number the population(8585 children). The numbering has already been provided by the frequency table. There are 2 children with height of 105 cm, therefore we assign number 1 and 2 to the children those in the group 105 cm, number 3 to 6 is assigned to those in the group 107 cm and similarly all other children are assigned the numbers. In the last group 145 cms there are two children with assigned number 8584 and 8585.

Height (cm.)	**Number of children**	**Cumulative Frequency**
105	2	2
107	4	6
109	14	20
111	41	61
113	83	144
115	169	313
117	394	707
119	669	1376
121	990	2366
123	1223	3589
125	1329	4918
127	1230	6148
129	1063	7211
131	646	7857
133	392	8249
135	202	8451
137	79	8530
139	32	8562
141	16	8578
143	5	8583
145	2	8585
Total	8585	

Now we take 10 samples from the tables, since the population size is in 4 digits we can use the given random number table. Select the10 random numbers from 1 to 8585 in the table, Here, we consider column wise selection of random numbers, starting from first column.

Tippet's random number table							
2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

The children with assigned number 2952 is selected and then see the cumulative frequency table where 2952 is present, now select the corresponding row height which is 123 cm, similarly all the selected random numbers are considered for the selection of the child with their corresponding height. The following table shows all the selected 10 children with their heights.

Child with assigned Number	2952	4167	2670	0560	2754
Corresponding Height (cms)	123	125	123	117	123
Child with assigned Number	6641	7483	5246	3992	1545
Corresponding Height (cms)	129	131	127	125	121

Example 8.4

Using the following random number table (Kendall-Babington Smith)

23	15	75	48	59	01	83	72	59	93	76	24	97	08	86	95	23	03	67	44
05	54	55	50	43	10	53	74	35	08	90	61	18	37	44	10	96	22	13	43
14	87	16	03	50	32	40	43	62	23	50	05	10	03	22	11	54	36	08	34
38	97	67	49	51	94	05	17	58	53	78	80	59	01	94	32	42	87	16	95
97	31	26	17	18	99	75	53	08	70	94	25	12	58	41	54	88	21	05	13

Draw a random sample of 10 four- figure numbers startingfrom 1550 to 8000.

Solution:

Here, we have to select 10 random numbers ranging from 1550 to 8000 but the given random number table has only 2 digit numbers. To solve this, two - 2 digit numbers can be combined together to make a four- figure number. Let us select the 5thand 6th column and combine them to form a random number, then select the random number with given range. This gives 5random numbers, similarly 8th and 9th is selected and combined to form a random numbers, then select the random number with given range. This gives

5random numbers, totally 10 four- figure numbers have been selected. The following table shows the 10 random numbers which are combined and selected.

23	15	75	48	59	01	83	72	59	93	76	24	97	08	86	95	23	03	67	44
05	54	55	50	43	10	53	74	35	08	90	61	18	37	44	10	96	22	13	43
14	87	16	03	50	32	40	43	62	23	50	05	10	03	22	11	54	36	08	34
38	97	67	49	51	94	05	17	58	53	78	80	59	01	94	32	42	87	16	95
97	31	26	17	18	99	75	53	08	70	94	25	12	58	41	54	88	21	05	13

Therefore the selected 10 random numbers are

5901	4310	5032	5194	1899
7259	7435	4362	1758	5308

(ii) Stratified Random Sampling

Definition 8.1

In stratified random sampling, first divide the population into sub-populations, which are called strata. Then, the samples are selected from each of the strata through random techniques.The collection of all the samples from all strata gives the stratified random samples.

When the population is heterogeneous or different segments or groups with respect to the variable or characteristic under study, then Stratified Random Sampling method is studied. First, the population is divided into homogeneous number of sub-groups or strata before the sample is drawn. A sample is drawn from each stratum at random. Following steps are involved for selecting a random sample in a stratified random sampling method.

(a) The population is divided into different classes so that each stratum will consist of more or less homogeneous elements. The strata are so designed that they do not overlap each other.

(b) After the population is stratified, a sample of a specified size is drawn at random from each stratum using Lottery Method or Table of Random Number Method.

Stratified random sampling is applied in the field of the different legislative areas as strata in election polling, division of districts (strata) in a state etc…

Example 8.5

From the following data, select 68 random samples from the populationof heterogeneous group with size of 500 through stratified random sampling, considering the following categories as strata.

Category1: Lower income class -39%

Category2: Middle income class - 38%

Category3: Upper income class- 23%

Solution

Stratum	Homogenous group	Percentage from population	Number of people in each strata	Random Samples
Category1	Lower income class	39	$\frac{39}{100} \times 500 = 195$	$195 \times \frac{68}{500} = 26.5 \sim 26$
Category2	Middle income class	38	$\frac{38}{100} \times 500 = 190$	$190 \times \frac{68}{500} = 26.5 \sim 26$ $190 \times \frac{68}{500} = 26.5 \sim 26$
Category3	Upper income class	23	$\frac{23}{100} \times 500 = 115$	$115 \times \frac{68}{500} = 15.6 \sim 16$
Total		**100**	**500**	**68**

Merits

(a) A random stratified sample is superior to a simple random sample because it ensures representation of all groups and thus it is more representative of the population which is being sampled.

(b) A stratified random sample can be kept small in size without losing its accuracy.

(c) It is easy to administer,if the population under study is sub-divided.

(d) It reduces the time and expenses in dividing the strata into geographical divisions, since the government itself had divided the geographical areas.

Demerits

(a) To divide the population into homogeneous strata (if not divided), it requires more money, time and statistical experience which is a difficult one.

(b) If proper stratification is not done, the sample will have an effect of bias.

(c) There is always a possibility of faulty classification of strata and hence increases variability.

(iii) Systematic Sampling

Definition 8.2

In a systematic sampling, randomly select the first sample from the first k units. Then every k^{th} member, starting with the first selected sample, is included in the sample.

Systematic sampling is a commonly used technique, if the complete and up-to-date list of the sampling units is available. We can arrange the items in numerical, alphabetical, geographical or in any other order. The procedure of selecting the samples starts with selecting the first sample at random, the rest being automatically selected according to some pre-determined pattern. A systematic sample is formed by selecting every item from the population, where k refers to the sample interval. The sampling interval can be determined by dividing the size of the population by the size of the sample to be chosen.

That is $k = \frac{N}{n}$,where k is an integer.

k = Sampling interval , Size of the population, Sample size.

Procedure for selection of samples by systematic sampling method

(i) If we want to select a sample of 10 students from a class of 100 students, the sampling interval is calculated as $k = \frac{N}{n} = \frac{100}{10} = 10$.

Thus sampling interval = 10 denotes that for every 10 samples one sample has to be selected.

(ii) The firstsample is selected from the first 10 (sampling interval) samples through random selection procedures.

(iii) If the selected first random sample is 5, then the rest of the samples are automatically selected by incrementing the value of the sampling interval (k = 10) i.e., 5, 15 , 25, 35, 45, 55, 65, 75, 85, 95.

Example:

Suppose we have to select 20 items out of 6,000. The procedure is to number all the 6,000 items from 1 to 6,000. The sampling interval is calculated as $k = \frac{N}{n} = \frac{6000}{20} = 300$. Thus sampling interval = 300 denotes that for every 300 samples one sample has to be selected. The first sample is selected from the first 300 (sampling interval) samples through random selection procedures. If the selected first random sample is 50, then the rest of the samples are automatically selected by incrementing the value of the sampling interval (k=300) ie,50, 350, 650, 950, 1250, 1550, 1850, 2150, 2450, 2750, 3050, 3350, 3650, 3950, 4250, 4550, 4850, 5150, 5450, 5750. Items bearing those numbers will be selected as samples from the population.

Merits

1. This is simple and convenient method.
2. This method distributes the sample more evenly over the entire listed population.
3. The time and work is reduced much.

Demerits

1. Systematic samples are not random samples.
2. If N is not a multiple of n, then the sampling interval (k) cannot be an integer, thus sample selection becomes difficult.

8.1.2 Sampling and Non-Sampling Errors:

A sample is a part of the whole population. A sample drawn from the population depends upon chance and as such all the characteristics of the population may not be present in the sample drawn from the same population. The errors involved in the collection, processing and analysis of the data may be broadly classified into two categories namely,

(i) Sampling Errors

(ii) Non-Sampling Errors

(i) Sampling Errors

Errors, which arise in the normal course of investigation or enumeration on account of chance, are called sampling errors. Sampling errors are inherent in the method of sampling. They may arise accidentally without any bias or prejudice. Sampling Errors arise primarily due to the following reasons:

(a) Faulty selection of the sample instead of correct sample by defective sampling technique.

(b) The investigator substitutes a convenient sample if the original sample is not available while investigation.

(c) In area surveys, while dealing with border lines it depends upon the investigator whether to include them in the sample or not. This is known as Faulty demarcation of sampling units.

(ii)Non-Sampling Errors

The errors that arise due to human factors which always vary from one investigator to another in selecting, estimating or using measuring instruments(tape, scale)are called Non-Sampling errors.It may arise in the following ways:

(a) Due to negligence and carelessness of the part of either investigator or respondents.

(b) Due to lack of trained and qualified investigators.

(c) Due to framing of a wrong questionnaire.

(d) Due to apply wrong statistical measure

(e) Due to incomplete investigation and sample survey.

8.1.3 Sampling distribution

Definition 8.3

Sampling distribution of a statistic is the frequency distribution which is formed with various values of a statistic computed from different samples of the same size drawn from the same population.

For instance if we draw a sample of size n from a given finite population of size N, then the total number of possible samples is ${}^{N}C_{n} = \frac{N!}{n!(N-n)!} = k$ (say). For each of these k samples we can compute some statistic, $t = t(x_1, x_2, x_3, ... x_n)$, in particular the mean $\overline{x}$, the variance S2, etc., is given below

Sample Number	Statistic		
		$\overline{x}$	S^2
1	t_1	$\overline{x}_1$	S_1^2
2	t_2	$\overline{x}_2$	S_2^2
3	t_3	$\overline{x}_3$	S_3^2
. . .	. . .	. . .	. . .
	t_k	$\overline{x}_k$	S_k^2

The set of the values of the statistic so obtained, one for each sample, constitutes the sampling distribution of the statistic.

Standard Error

The standard deviation of the sampling distribution of a statistic is known as its Standard Error abbreviated as S.E. The Standard Errors (S.E.) of some of the well-known statistics, for large samples, are given below, where n is the sample size, σ^2 is the population variance.

S.No	Statistic	Standard Error
1.	Sample mean	$\sigma/\sqrt{n}$
2.	Observed sample proportion	$\sqrt{PQ/n}$

3.	Sample standard deviation	$\sqrt{\sigma^2/2n}$
4.	Sample variance	$\sigma^2\sqrt{2/n}$
5.	Sample quartiles	$1.36263\,\sigma/\sqrt{n}$
6.	Sample median	$1.25331\,\sigma/\sqrt{n}$
7.	Sample correlation coefficient	$(1-\rho^2)/\sqrt{n}$

8.1.4 Computing standard error in simple cases

Example 8.6

A server channel monitored for an hour was found to have an estimated mean of 20 transactions transmitted per minute. The variance is known to be 4. Find the standard error.

Solution:

Given $\sigma^2 = 4$ which implies $\sigma = 2$, $n = 1$ hour = 60 min, $\bar{X} = 20$ /min

Standard Error = $\dfrac{\sigma}{\sqrt{n}} = \dfrac{2}{\sqrt{60}} = 0.2582$

Example 8.7

Find the sample size for the given standard deviation 10 and the standard error with respect of sample mean is 3.

Solution:

Given $\sigma = 10$, S.E. $\bar{X}$ =3 We know that S.E = $\sigma/\sqrt{n}$

Therefore, $3 = \dfrac{10}{\sqrt{n}} \Rightarrow \sqrt{n} = \dfrac{10}{3}$

Taking Squaring on both sides we get

$$n = \left(\frac{10}{3}\right)^2 = \frac{100}{9} = 11.11 \cong 11\text{, The required sample size is 11.}$$

Example 8.8

`A die is thrown 9000 times and a throw of 3 or 4 is observed 3240 times. Find the standard error of the proportion for an unbiased die .

Solution :

If the occurrence of 3 or 4 on the die is called a success, then

12th BM EM V-2.indb 95 30-03-2019 17:13:36

Sample size $= 9000$; Number of Success $= 3240$

$$\text{Sample proportion } = p = \frac{3240}{9000} = 0.36$$

Population proportion (P) = Prob(getting 3 or 4 when a die is thrown)

$$= \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3} = 0.3333$$

Thus $P = 0.3333$ and $Q = 1-P=1-0.3333 = 0.6667$

The S.E for sample proportion is given by

$$S.E. = \sqrt{\frac{PQ}{N}} = \sqrt{\frac{(0.3333)(0.6667)}{9000}} = 0.00496$$

Hence the standard error for sample proportion is S.E=0.00496.

Example 8.9

The standard deviation of a sample of size 50 is 6.3. Determine the standard error whose population standard deviation is 6?

Solution:

Sample size $n = 50$

Sample S.D $s = 6.3$

Population S.D $\sigma = 6$

The standard error for sample S.D is given by

$$S.E. = \sqrt{\frac{\sigma^2}{2n}} = \frac{6}{\sqrt{2(50)}} = \frac{6}{\sqrt{100}} = 1.8974$$

Thus standard error for sample S.D = 1.8974

Example 8.10

A sample of 100 students is chosen from a large group of students. The average height of these students is 162 cm and standard deviation (S.D) is 8 cm. Obtain the standard error for the average height of large group of students of 160 cm?

Solution:

Given = 100, $\bar{x}$ =162 cm, s = 8 cm is known in this problem

since σ is unknown , so we consider $\breve{\sigma} = s$ and μ= 160 cm

$$S.E. = \frac{\breve{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n}} = \frac{8}{\sqrt{100}} = 0.8$$

Therefore the standard error for the average height of large group of students of 160 cm is 0.8.

Exercise 8.1

1. What is population?
2. What is sample?
3. What is statistic?
4. Define parameter.
5. What is sampling distribution of a statistic?
6. What is standard error?
7. Explain in detail about simple random sampling with a suitable example.
8. Explain the stratified random sampling with a suitable example.
9. Explain in detail about systematic random sampling with example.
10. Explain in detail about sampling error.
11. Explain in detail about non-sampling error.
12. State any two merits of simple random sampling.
13. State any three merits of stratified random sampling.
14. State any two demerits of systematic random sampling.
15. State any two merits for systematic random sampling.
16. Using the following Tippet's random number table

2952	6641	3992	9792	7969	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2670	7483	3408	2762	3563	1089	6913	7991
0560	5246	1112	6107	6008	8125	4233	8776
2754	9143	1405	9025	7002	6111	8816	6446

Draw a sample of 10 three digit numbers which are even numbers.

17. A wholesaler in apples claims that only 4% of the apples supplied by him are defective. A random sample of 600 apples contained 36 defective apples. Calculate the standard error concerning of good apples
18. A sample of 1000 students whose mean weight is 119 lbs(pounds) from a school in Tamil Nadu State was taken and their average weight was found to be 120 lbs with a standard deviation of 30 lbs. Calculate standard error of mean.

19. A random sample of 60 observations was drawn from a large population and its standard deviation was found to be 2.5. Calculate the suitable standard error that this sample is taken from a population with standard deviation 3?

20. In a sample of 400 population from a village 230 are found to be eaters of vegetarian items and the rest non-vegetarian items. Compute the standard error assuming that both vegetarian and non-vegetarian foods are equally popular in that village?

Statistical Inference

One of the main objectives of any statistical investigation is to draw inferences about a population from the analysis of samples drawn from that population. Statistical Inference provides us how to estimate a value from the sample and test that value for the population. This is done by the two important classifications in statistical inference,

(i) Estimation (ii) Testing of Hypothesis

8.2 Estimation:

It is possible to draw valid conclusion about the population parameters from sampling distribution. Estimation helps in estimating an unknown population parameter such as population mean, standard deviation, etc., on the basis of suitable statistic computed from the samples drawn from population.

Estimation:

Definition 8.4

The method of obtaining the most likely value of the population parameter using statistic is called estimation.

Estimator:

Definition 8.5

Any sample statistic which is used to estimate an unknown population parameter is called an estimator ie., an estimator is a sample statistic used to estimate a population parameter.

Estimate:

Definition 8.6

When we observe a specific numerical value of our estimator, we call that value is an estimate. In other words, an estimate is a specific observed value of a statistic.

Characteristic of a good estimator

A good estimator must possess the following characteristic:

(i) Unbiasedness (ii) Consistency(iii) Efficiency (iv) Sufficiency.

(i) Unbiasedness: An estimator $T_n = T(x_1, x_2, ..., x_n)$ is said to be an unbiased estimator of $\gamma(\theta)$ *if* $E(T_n) = \gamma(\theta)$, for all $\theta \ \varepsilon \ \theta$ (parameter space), (i.e)An estimator is said to be unbiased if its expected value is equal to the population parameter. Example: $E(\bar{x}) = \mu$

(ii) Consistency: An estimator $T_n = T(x_1, x_2, ..., x_n)$ is said to be consistent estimator of $\gamma(\theta)$, if T_n converges to $\gamma(\theta)$ in Probability, i.e., $T_n \xrightarrow{P} \gamma(\theta) \, as \, n \rightarrow \infty$, for all $\theta \ \varepsilon \Theta$.

(iii) Efficiency:If T_1 is the most efficient estimator with variance V_1 and T_2 is any other estimator with variance V_2, then the efficiency E of T_2 is defined as $E = \frac{V_1}{V_2}$ Obviously, E cannot exceed unity.

(iv) Sufficiency: If $T = t(x_1, x_2, ..., x_n)$ is an estimator of a parameter θ, based on a sample $x_1, x_2, ..., x_n$ of size n from the population with density $f(x, \theta)$ such that the conditional distribution of $x_1, x_2, ..., x_n$ given T, is independent of θ, then T is sufficient estimator for θ.

8.2.1 Point and Interval Estimation:

To estimate an unknown parameter of the population, concept of theory of estimation is used.There are two types of estimation namely,

1. Point estimation
2. Interval estimation

1. Point Estimation

When a single value is used as an estimate, the estimate is called a point estimate of the population parameter. In other words, an estimate of a population parameter given by a single number is called as point estimation.

For example

(i) 55 is the mean mark obtained by a sample of 5 students randomly drawn from a class of 100 students is considered to be the mean marks of the entire class. This single value 55is a point estimate.

(ii) 50 kg is the average weight of a sample of 10 students randomly drawn from a class of 100 students is considered to be the average weight of the entire class. This single value 50 is a point estimate.

Note

The sample mean $(\bar{x})$ is the sample statistic used as an estimate of population mean (μ)

Instead of considering, the estimated value of the population parameter to be a single value, we might consider an interval for estimating the value of the population parameter. This concept is known as interval estimation and is explained below.

2. Interval Estimation

Generally, there are situations where point estimation is not desirable and we are interested in finding limits within which the parameter would be expected to lie is called an interval estimation.

For example,

If T is a good estimator of θ with standard error s then, making use of general property of the standard deviations, the uncertainty in T, as an estimator of θ, can be expressed by statements like " We are about 95% certain that the unknown θ, will lie somewhere between T-2s and T+2s", "we are almost sure that θ will in the interval (T-3s and T+3s)" such intervals are called confidence intervals and is explained below.

Confidence interval

After obtaining the value of the statistic 't' (sample) from a given sample, Can we make some reasonable probability statements about the unknown population parameter 'θ'?. This question is very well answered by the technique of Confidence Interval. Let us choose a small value of α which is known as level of significance(1% or 5%) and determine two constants say, c_1 and c_2 such that $P(c_1 < \theta < c_2 | t) = 1 - \alpha$.

The quantities c_1 and c_2, so determined are known as the Confidence Limits and the interval $\left[c_1, c_2\right]$ within which the unknown value of the population parameter is expected to lie is known as Confidence Interval. $(1-\alpha)$ is called as confidence coefficient.

Confidence Interval for the population mean for Large Samples (when is known)

If we take repeated independent random samples of size n from a population with an unknown mean but known standard deviation, then the probability that the true population mean μ will fall in the following interval is $(1-\alpha)$ i.e

$$P = \left(\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = (1-\alpha)$$

So, the confidence interval for population mean (μ), when standard deviation (σ) is known and is given by $\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$.

For the computation of confidence intervals and for testing of significance, the critical values

Z_α at the different level of significance is given in the following table:

Normal Probability Table

Critical Values Z_α	Level of significance (α)			
	1%	2%	5%	10%
Two-tailed test	$\lvert Z_\alpha \rvert$=2.58	$\lvert Z_\alpha \rvert$=2.33	$\lvert Z_\alpha \rvert$=1.96	$\lvert Z_\alpha \rvert$=1.645
Right tailed test	Z_α =2.33	Z_α =2.055	Z_α =1.645	Z_α =1.28
Left tailed test	Z_α = -2.33	$Z_\alpha = -2.055$	$Z_\alpha = -1.645$	$Z_\alpha = -1.28$

The calculation of confidence interval is illustrated below.

Example 8.11

A machine produces a component of a product with a standard deviation of 1.6 cm in length. A random sample of 64 componentsvwas selected from the output and this sample has a mean length of 90 cm. The customer will reject the part if it is either less than 88 cm or more than 92 cm. Does the 95% confidence interval for the true mean length of all the components produced ensure acceptance by the customer?

Solution:

Here μ is the mean length of the components in the population.

The formula for the confidence interval is

$$\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Here σ = 1.6, $Z_{\alpha/2}$ = 1.96, $\bar{x}$ = 90 and n = 64

Then $S.E. = \frac{\sigma}{\sqrt{n}} = \frac{1.6}{\sqrt{64}} = 0.2$

Therefore, $90 - (1.96 \times 0.2) \leq \mu \leq 90 + (1.96 \times 0.2)$

$(89.61 \leq \mu \leq 90.39)$

This implies that the probability that the true value of the population mean length of the components will fall in this interval (89.61,90.39) at 95% . Hence we concluded that 95% confidence interval ensures acceptance of the component by the consumer.

Example 8.12

A sample of 100 measurements at breaking strength of cotton thread gave a mean of 7.4 and a standard deviation of 1.2 gms. Find 95% confidence limits for the mean breaking strength of cotton thread.

Solution:

Given, sample size = 100, $\bar{x}$ = 7.4, since σ is unknown but s = 1.2 is known.

In this problem, we consider $\breve{\sigma} = s$, $Z_{\alpha/2} = 1.96$

$$S.E. = \frac{\breve{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n}} = \frac{1.2}{\sqrt{100}} = 0.12$$

Hence 95% confidence limits for the population mean are

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$7.4 - (1.96 \times 0.12) \le \mu \le 7.4 + (1.96 \times 0.12)$$

$$7.4 - 0.2352 \le \mu \le 7.4 + 0.2352$$

$$7.165 \le \mu \le 7.635$$

This implies that the probability that the true value of the population mean breaking strength of the cotton threads will fall in this interval (7.165,7.635) at 95% .

Example 8.13

The mean life time of a sample of 169 light bulbs manufactured by a company is found to be 1350 hours with a standard deviation of 100 hours. Establish 90% confidence limits within which the mean life time of light bulbs is expected to lie.

Solution:

Given: n = 169, $\bar{x}$ = 1350 hours, σ = 100 hours, since the level of significance is (100-90)% =10% thus α is 0.1, hence the significant value at 10% is $Z_{\alpha/2}$ = 1.645

$$S.E. = \frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{169}} = 7.69$$

Hence 90% confidence limits for the population mean are

$$\bar{x} - Z_{\alpha/2} SE < \mu < \bar{x} + Z_{\alpha/2} SE$$

$$1350 - (1.645 \times 7.69) \le \mu$$

$$1337.35 \le \mu \le 1362.65$$

Hence the mean life time of light bulbs is expected to lie between the interval (1337.35, 1362.65)

8.3 Hypothesis Testing

One of the important areas of statistical analysis is testing of hypothesis. Often, in real life situations we require to take decisions about the population on the basis of sample information. Hypothesis testing is also referred to as "Statistical Decision Making". It employs statistical techniques to arrive at decisions in certain situations where there is an element of uncertainty on the basis of sample, whose size is fixed in advance. So statistics helps us in arriving at the criterion for such decision is known as Testing of hypothesis which was initiated by J. Neyman and E.S. Pearson.

For Example: We may like to decide on the basis of sample data whether a new vaccine is effective in curing cold, whether a new training methodology is better than the existing one, whether the new fertilizer is more productive than the earlier one and so on.

8.3.1 Meaning : Null Hypothesis and Alternative Hypothesis - Level of Significants and Type of Errors

Statistical Hypothesis

Statistical hypothesis is some assumption or statement, which may or may not be true, about a population.

There are two types of statistical hypothesis

(i) Null hypothesis (ii) Alternative hypothesis

Null Hypothesis

Definition 8.7

According to Prof. R.A.Fisher, "Null hypothesis is the hypothesis which is tested for possible rejection under the assumption that it is true", and it is denoted by H_0 .

For example: If we want to find the population mean has a specified value μ_0 , then the null hypothesis H_0 is set as follows $H_0 : \mu = \mu_0$

Alternative Hypothesis

Any hypothesis which is complementary to the null hypothesis is called as the alternative hypothesis and is usually denoted by H_1.

For example: If we want to test the null hypothesis that the population has specified mean μ i.e., $H_0 : \mu = \mu_0$ then the alternative hypothesis could be any one among the following:

(i) $H_1 : \mu \neq \mu_0$ ($\mu >$ or $\mu < \mu_0$)

(ii) $H_1 : \mu > \mu_0$

(iii) $H_1 : \mu < \mu_0$

The alternative hypothesis in $H_1 : \mu \neq \mu_0$ is known as two tailed alternative test. Two tailed test is one where the hypothesis about the population parameter is rejected for the value of sample statistic falling into either tails of the sampling distribution. When the hypothesis about the population parameter is rejected only for the value of sample statistic falling into one of the tails of the sampling distribution, then it is known as one-tailed test. Here $H_1 : \mu > \mu_0$ and $H_1 : \mu < \mu_0$ are known as one tailed alternative.

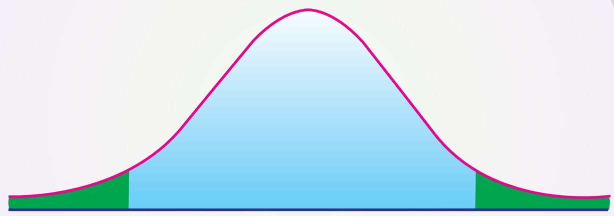

Two tailed diagram

Fig.8.1

Right tailed test: $H_1 : \mu > \mu_0$ is said to be right tailed test where the rejection region or critical region lies entirely on the right tail of the normal curve.

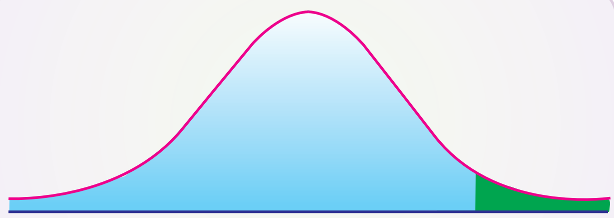

Right tailed diagram

Fig.8.2

Left tailed test: $H_1 : \mu < \mu_0$ is said to be left tailed test where the critical region lies entirely on the left tail of the normal curve.

(diagram)

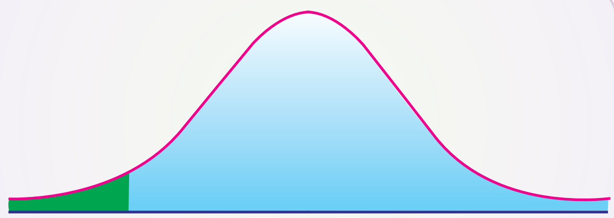

Left tailed test

Fig.8.3

Types of Errors in Hypothesis testing

There is every chance that a decision regarding a null hypothesis may be correct or may not be correct. There are two types of errors. They are

Type I error: The error of rejecting H_0 when it is true.

Type II error: The error of accepting when H_0 it is false.

Critical region or Rejection region

A region corresponding to a test statistic in the sample space which tends to rejection of H_0 is called critical region or region of rejection.

Note

The region complementary to the critical region is called the region of acceptance.

Level of significance

The probability of type I error is known as level of significance and it is denoted by . The level of significance is usually employed in testing of hypothesis are 5% and 1%. The level of significance is always fixed in advance before collecting the sample information.

Critical values or significant values

The value of test statistic which separates the critical (or rejection) region and the acceptance region is called the critical value or significant value. It depend upon

(i) The level of significance

(ii) The alternative hypothesis whether it is two-tailed or single tailed.

For large samples, the standardized variable corresponding to the statistic viz.,

$$Z = \frac{t - E(\mathrm{t})}{\sqrt{Var(t)}} = \frac{t - E(t)}{S.E.(t)} \sim N(0,1) \text{ as } n \to \infty \qquad \ldots (1)$$

The value of Z given by (1) under the null hypothesis is known as test statistic. The critical values of Z at commonly used level of significance for both two tailed and single tailed tests are given in the normal probability table (Refer the normal probably Table).

Since for large n, almost all the distributions namely, Binomial, Poisson, etc., can be approximated very closely by a normal probability curve, we use the normal test of significance for large samples.

Note

For practical purposes, the sample may be regarded as large if $n > 30$

8.3.2 Testing Procedure : Large sample theory and test of significants for single mean

The following are the steps involved in hypothesis testing problems

1. Null hypothesis: Set up the null hypothesis H_0
2. Alternative hypothesis: Set up the alternative hypothesis . This will enable us to decide whether we have to use two tailed test or single tailed test (right or left tailed)
3. Level of significance: Choose the appropriate level of significant (α) depending on the reliability of the estimates and permissible risk. This is to be fixed before sample is drawn. i.e., α is fixed in advance.
4. Test statistic : Compute the test statistic

$$Z = \frac{t - E(\mathrm{t})}{\sqrt{Var(t)}} = \frac{t - E(t)}{S.E.(t)} \sim N(0,1) \text{ as } n \to \infty$$

5. Conclusion: We compare the computed value of Z in step 4 with the significant value or critical value or table value Z_α at the given level of significance .

(i) If $|Z| < Z_\alpha$ i.e., if the calculated value of is less than critical value we say it is not significant. This may due to fluctuations of sampling and sample data do not provide us sufficient evidence against the null hypothesis which may therefore be accepted.

(ii) If $|Z| > Z_\alpha$ i.e., if the calculated value of is greater than critical value Z_α then we say it is significant and the null hypothesis is rejected at level of significance α .

Test of significance for single mean

Let $x_i, (i = 1, 2, 3, ..., n)$ is a random sample of size from a normal population with mean μ and variance σ^2 then the sample mean is distributed normally with mean and variance $\frac{\sigma^2}{n}$, i.e., $\overline{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. Thus for large samples, the standard normal variate corresponding to $\overline{x}$ is :

$$Z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

Under the null hypothesis that the sample has been drawn from a population with mean and variance σ^2, i.e., there is no significant difference between the sample mean $(\overline{x})$ and the population mean (α), the test statistic (for large samples) is:

$$Z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Remark:

If the population standard deviation σ is unknown then we use its estimate provided by the sample variance given by $\hat{\sigma}^2 = s_2 \Rightarrow \hat{\sigma} = s$

Example 8.14

An auto company decided to introduce a new six cylinder car whose mean petrol consumption is claimed to be lower than that of the existing auto engine. It was found that the mean petrol consumption for the 50 cars was 10 km per litre with a standard deviation of 3.5 km per litre. Test at 5% level of significance, whether the claim of the new car petrol consumption is 9.5 km per litre on the average is acceptable.

Solution:

Sample size n =50 Sample mean $\overline{x}$ = 10 km Sample standard deviation s = 3.5 km

Population mean μ =9.5 km

Since population SD is unknown we consider $\sigma = s$

The sample is a large sample and so we apply Z-test

Null Hypothesis: There is no significant difference between the sample average and the company's claim, i.e., $H_0 : \mu = 9.5$

Alternative Hypothesis: There is significant difference between the sample average and the company's claim, i.e., $H_1 : \mu \neq 9.5$ (two tailed test)

The level of significance $\alpha = 5\% = 0.05$

Applying the test statistic

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1); \quad Z = \frac{10 - 9.5}{\frac{3.5}{\sqrt{50}}} \sim N(0,1) = \frac{0.5}{0.495} = 1.01$$

Thus the calculated value 1.01 and the significant value or table value $Z_{\alpha/2} = 1.96$

Comparing the calculated and table value ,Here $Z < Z_{\alpha/2}$ i.e., 1.01<1.96.

Inference:Since the calculated value is less than table value i.e., $Z < Z_{\alpha/2}$ at 5% level of sinificance, the null hypothesis H_0 is accepted. Hence we conclude that the company's claim that the new car petrol consumption is 9.5 km per litre is acceptable.

Example 8.15

A manufacturer of ball pens claims that a certain pen he manufactures has a mean writing life of 400 pages with a standard deviation of 20 pages. A purchasing agent selects a sample of 100 pens and puts them for test. The mean writing life for the sample was 390 pages. Should the purchasing agent reject the manufactures claim at 1% level?

Solution:

Sample size n =100, Sample mean $\bar{x}$ = 390 pages, Population mean μ = 400 pages

Population SD σ = 20 pages

The sample is a large sample and so we apply Z -test

Null Hypothesis: There is no significant difference between the sample mean and the population mean of writing life of pen he manufactures, i.e., $H_0 : \mu = 400$

Alternative Hypothesis: There is significant difference between the sample mean and the population mean of writing life of pen he manufactures, i.e., $H_1 : \mu \neq 400$ (two tailed test)

The level of significance $\alpha = 1\% = 0.01$

Applying the test statistic

$$Z=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1);\ Z=\frac{390-400}{\frac{20}{\sqrt{100}}}=\frac{-10}{2}=-5,\ \therefore |Z|=5$$

Thus the calculated value $|Z|$ = 5 and the significant value or table value $Z_{\alpha/2}$ = 2.58

Comparing the calculated and table values, we found $Z > Z_{\alpha/2}$ i.e., 5 > 2.58

Inference: Since the calculated value is greater than table value i.e., $Z > Z_{\alpha/2}$ at 1% level of significance, the null hypothesis is rejected and Therefore we concluded that $\mu \neq 400$ and the manufacturer's claim is rejected at 1% level of significance.

Example 8.16

(i) A sample of 900 members has a mean 3.4 cm and SD 2.61 cm. Is the sample taken from a large population with mean 3.25 cm. and SD 2.62 cm?

(ii) If the population is normal and its mean is unknown, find the 95% and 98% confidence limits of true mean.

Solution:

(i) Given:

Sample size n =900, Sample mean $\bar{x}$ = 3.4 cm, Sample SD s = 2.61 cm

Population mean μ= 3.25 cm, Population SD σ = 2.61 cm

Null Hypothesis H_0 : μ = 3.25 cm (the sample has been drawn from the population mean μ = 3.25 cm and SD σ = 2.61 cm)

Alternative Hypothesis H_1 : $\mu \neq$ 3.25 cm (two tail) i.e., the sample has not been drawn from the population mean μ = 3.25 cm and SD σ = 2.61 cm.

The level of significance $\alpha = 5\% = 0.05$

Teststatistic:

$$Z=\frac{3.4-3.25}{\frac{2.61}{\sqrt{900}}}=\frac{0.15}{0.087}=1.724$$

$$\therefore Z=1.724$$

Thus the calculated and the significant value or table value $Z_{\alpha/2}$ = 1.96

Comparing the calculated and table values, $Z < Z_{\alpha/2}$ i.e., 1.724 < 1.96

Inference:Since the calculated value is less than table value i.e., $Z > Z_{\alpha/2}$ at 5% level of significance, the null hypothesis is accepted. Hence we conclude that the data doesn't provide us any evidence against the null hypothesis. Therefore, the sample has been drawn from the population mean μ = 3.25 cm and SD, σ = 2.61 cm.

(ii) Confidence limits

95% confidential limits for the population mean μ are :

$$\bar{x} - Z_{\alpha/2} SE \leq \mu \leq \bar{x} + Z_{\alpha/2} SE$$

$$3.4 - (1.96 \times 0.087) \leq \mu \leq 3.4 + (1.96 \times 0.087)$$

$$3.229 \leq \mu \leq 3.571$$

98% confidential limits for the population mean are :

$$\bar{x} - Z_{\alpha/2} SE \leq \mu \leq \bar{x} + Z_{\alpha/2} SE$$

$$3.4 - (2.33 \times 0.087) \leq \mu \leq 3.4 + (2.33 \times 0.087)$$

$$3.197 \leq \mu \leq 3.603$$

Therefore,95% confidential limits is (3.229,3.571) and 98% confidential limits is (3.197,3.603).

Example 8.17

The mean weekly sales of soap bars in departmental stores were 146.3 bars per store. After an advertising campaign the mean weekly sales in 400 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the advertising campaign successful?

Solution:

Sample size n = 400 stores

Sample mean $\bar{x}$ = 153.7 bars

Sample SD s = 17.2 bars

Population mean m = 146.3 bars

Since population *SD* is unknown we can consider the sample *SD* $s = \sigma$

Null Hypothesis. The advertising campaign is not successful i.e, $H_0 : \mu = 146.3$ (There is no significant difference between the mean weekly sales of soap bars in department stores before and after advertising campaign)

Alternative Hypothesis $H_1 : \mu > 143.3$ (Right tail test). The advertising campaign was successful

Level of significance α = 0.05

Test statistic

$$Z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$Z = \frac{153.7 - 146.3}{\frac{17.2}{\sqrt{400}}}$$

$$= \frac{7.4}{0.86} = 8.605$$

$$\therefore Z = 8.605$$

Comparing the calculated value $Z = 8.605$ and the significant value or table value $Z_{\alpha} = 1.645$. we get 8.605 > 1.645

Inference: Since, the calculated value is much greater than table value i.e., $Z > Z_{\alpha}$, it is highly significant at 5% level of significance. Hence we reject the null hypothesis H_0 and conclude that the advertising campaign was definitely successful in promoting sales.

Example 8.18

The wages of the factory workers are assumed to be normally distributed with mean and variance 25. A random sample of 50 workers gives the total wages equal to ₹ 2,550. Test the hypothesis μ = 52, against the alternative hypothesis μ = 49 at 1% level of significance.

Solution:

Sample size n = 50 workers

Total wages Σx = 2550

Sample mean $\overline{x} = \frac{total\,wages}{n} = \frac{\Sigma x}{n} = \frac{2550}{50} = 51$ units

Population mean μ = 52

Population variance σ^2 = 25

Population SD σ = 5

Under the null hypothesis $H_0 : \mu$ = 52

Against the alternative hypothesis $H_1 : \mu \neq 52$ (Two tail)

Level of significance μ = 0.01

Test statistic $Z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

$$Z = \frac{51 - 52}{\frac{5}{\sqrt{50}}} = \frac{-1}{0.7071} = -1.4142$$

Since alternative hypothesis is of two tailed test we can take $|Z| = 1.4142$

Critical value at 1% level of significance is $Z_{\alpha/2} = 2.58$

Inference: Since the calculated value is less than table value i.e., $Z < Z_{\alpha/2}$ at 1% level of significance, the null hypothesis H_0 is accepted. Therefore, we conclude that there is no significant difference between the sample mean and population mean $\mu = 52$ and SD $\sigma = 5$.

Example 8.19

An ambulance service claims that it takes on the average 8.9 minutes to reach its destination in emergency calls. To check on this claim, the agency which licenses ambulance services has them timed on 50 emergency calls, getting a mean of 9.3 minutes with a standard deviation of 1.6 minutes. What can they conclude at the level of significance

Solution:

Sample size	$n = 50$
Sample mean	$\bar{x} = 9.3$ minutes
Sample S.D	$s = 1.6$ minutes
Population mean	$\mu = 8.9$ minutes
Null hypothesis	$H_0 : \mu = 8.9$
Alternative hypothesis	$H_1 : \mu = 8.9$(Two tail)
Level of significance	$\mu = 0.05$
Test statistic	$Z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$
	$Z = \dfrac{9.3 - 8.9}{\frac{1.6}{\sqrt{50}}} = \dfrac{0.4}{0.2263} = 1.7676$
Calculated value	$Z = 1.7676$

Critical value at 5% level of significance is $Z_{\alpha/2} = 1.96$

Inference: Since the calculated value is less than table value i.e., $Z < Z_{\alpha/2}$ at 5% level of significance, the null hypothesis is accepted. Therefore we conclude that an ambulance service claims on the average 8.9 minutes to reach its destination in emergency calls.

Exercise 8.2

1. Mention two branches of statistical inference?
2. What is an estimator?

3. What is an estimate?
4. What is point estimation?
5. What is interval estimation?
6. What is confidence interval?
7. What is null hypothesis? Give an example.
8. Define alternative hypothesis.
9. Define critical region.
10. Define critical value.
11. Define level of significance
12. What is type I error
13. What is single tailed test.
14. A sample of 100 items, draw from a universe with mean value 4 and S.D 3, has a mean value 63.5. Is the difference in the mean significant?
15. A sample of 400 individuals is found to have a mean height of 67.47 inches. Can it be reasonably regarded as a sample from a large population with mean height of 67.39 inches and standard deviation 1.30 inches?
16. The average score on a nationally administered aptitude test was 76 and the corresponding standard deviation was 8. In order to evaluate a state's education system, the scores of 100 of the state's students were randomly selected. These students had an average score of 72. Test at a significance level of 0.05 if there is a significant difference between the state scores and the national scores.
17. The mean breaking strength of cables supplied by a manufacturer is 1,800 with a standard deviation 100. By a new technique in the manufacturing process it is claimed that the breaking strength of the cables has increased. In order to test this claim a sample of 50 cables is tested. It is found that the mean breaking strength is 1,850. Can you support the claim at 0.01 level of significance.

Exercise 8.3

Choose the correct Answer

1. A ………… may be finite or infinite according as the number of observations or items in it is finite or infinite.

   (a) Population (b) census (c) parameter (d) none of these

2. A ……… of statistical individuals in a population is called a sample.

(a) Infinite set (b) finite subset (c) finite set (d) entire set

3. A finite subset of statistical individuals in a population is called ………...

(a) a sample (b) a population (c) universe (d) census

4. Any statistical measure computed from sample data is known as …………...

(a) parameter (b) statistic (c) infinite measure (d) uncountable measure

5. A…………..is one where each item in the universe has an equal chance of known opportunity of being selected.

(a) Parameter (b) random sample (c) statistic (d) entire data

6. A random sample is a sample selected in such a way that every item in the population has an equal chance of being included

(a) Harper (b) Fisher (c) Karl Pearson (d) Dr. Yates

4D17MR

7. Which one of the following is probability sampling

(a) purposive sampling (b) judgment sampling

(c) simple random sampling (d) Convenience sampling

8. In simple random sampling from a population of units, the probability of drawing any unit at the first draw is

(a) $\frac{n}{N}$ (b) $\frac{1}{N}$ (c) $\frac{N}{n}$ (d) 1

9. In …………….. the heterogeneous groups are divided into homogeneous groups.

(a) Non-probability sample (b) a simple random sample

(c) a stratified random sample (d) systematic random sample

10. Errors in sampling are of

(a) Two types (b) three types (c) four types (d) five types

11. The method of obtaining the most likely value of the population parameter using statistic is called ……………………

(a) estimation (b) estimator (c) biased estimate (d) standard error.

12th BM EM V-2.indb 113 30-03-2019 17:14:07

12. An estimator is a sample statistic used to estimate a

(a) population parameter (b) biased estimate

(c) sample size (d) census

13. ....is a relative property, which states that one estimator is efficient relative to another.

(a) efficiency (b) sufficiency (c) unbiased (d) consistency

14. If probability $P[|\breve{\theta}-\theta|<\varepsilon]\rightarrow 1\mu$ as $n\rightarrow\infty'$, for any positive ε then $\breve{\theta}$ is said to ....estimator of θ.

(a) efficient (b) sufficient (c) unbiased (d) consistent

15. An estimator is said to be .... if it contains all the information in the data about the parameter it estimates.

(a) efficient (b) sufficient (c) unbiased (d) consistent

16. An estimate of a population parameter given by two numbers between which the parameter would be expected to lie is called an................interval estimate of the parameter.

(a)point estimate (b) interval estimation (c) standard error (d) confidence

17. A ________ is a statement or an assertion about the population parameter.

(a) hypothesis (b) statistic (c) sample (d) census

18. Type I error is

(a) Accept H_0 when it is true (b) Accept H_0 when it is false

(c) Reject H_0 when it is true (d) Reject H_0 when it is false.

19. Type II error is

(a) Accept H_0 when it is wrong (b) Accept H_0 when it is true

(c) Reject H_0 when it is true (d) Reject H_0 when it is false

20. The standard error of sample mean is

(a) $\frac{\sigma}{\sqrt{2n}}$ (b) $\frac{\sigma}{n}$ (c) $\frac{\sigma}{\sqrt{n}}$ (d) $\frac{\sigma^2}{\sqrt{n}}$

Miscellaneous Problems

1. Explain the types of sampling.

2. Write short note on sampling distribution and standard error.

3. Explain the procedures of testing of hypothesis
4. Explain in detail about the test of significance for single mean
5. Determine the standard error of proportion for a random sample of 500 pineapples was taken from a large consignment and 65 were found to be bad.
6. A sample of 100 students are drawn from a school. The mean weight and variance of the sample are 67.45 kg and 9 kg. respectively. Find (a) 95% and (b) 99% confidence intervals for estimating the mean weight of the students.
7. The mean I.Q of a sample of 1600 children was 99. Is it likely that this was a random sample from a population with mean I.Q 100 and standard deviation 15 ? (Test at 5% level of significance)

Summary

- **Sampling:** It is the procedure or process of selecting a sample from a population.
- **Population:** The group of individuals considered under study is called as population.
- **Sample :**Aselection of a group of individuals from a population.
- **Sample size :**The number of individuals included in a sample.
- **Simple Random Sampling :**The samples are selected in such a way that each and every unit in the population has an equal and independent chance of being selected as a sample.
- **Stratified Random Sampling:** When the population is heterogeneous, the population is divided into homogeneous number of sub-groups or strata. A sample is drawn from each stratum at random.
- **Systematic Sampling:** Select the first sample at random, the rest being automatically selected according to some predetermined pattern.
- **Sampling Distribution:** Sampling distribution of a statistic is the frequency distribution which is formed with various values of a statistic computed from different samples of the same size drawn from the same population.
- **Standard Error:** The standard deviation of the sampling distribution of a statistic is known as its Standard Error.
- **Statistical Inference:** To draw inference about a population of any statistical investigation from the analysis of samples drawn from that population.
- **Estimation :**The method of obtaining the most likely value of the population parameter using statistic is called estimation.

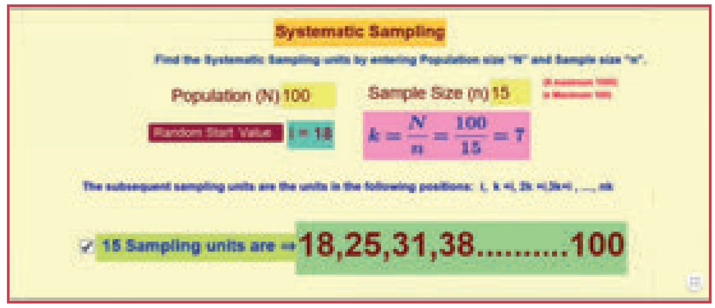
- **Point Estimation:** When a single value is used as an estimate, it is called as point estimation.
- **Interval Estimation:** An interval within which the parameter would be expected to lie is called interval estimation.
- **Test of Statistical Hypothesis:** Statistical technique to arrive at a decision in certain situations where there is an element of uncertainty on the basis of sampl
- **Null Hypothesis:** The hypothesis which is tested for possible rejection under the assumption that it is true", denoted by H_0.
- **Alternative Hypothesis:**The hypothesis which is complementary to the null hypothesis is called as the alternative hypothesis, denoted by H_1.
- **Type I error:**The error of rejecting H_0 when it is true.
- Type II **error:**The error of accepting H_0 when it is false.
- **Test of significance for single mean:**

$$Z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

GLOSSARY	
Alternative hypothesis	மாற்று கருதுகோள்
Confidence interval	நம்பிக்கைஇடைவெளி
Estimation	மதிப்பிடுதல்
Interval Estimation	இடைவெளிமதிப்பீடு
Non-Sampling Errors	கூறற்றபிழை
Null hypothesis	இன்மை கருதுகோள்
Parameter	தொகுதிப் பண்பளவை
Point Estimation	புள்ளிமதிப்பீடு
Population	முழுமைத்தொகுதி
Sample	கூறு
Sample size	கூறின்அளவு
Sampling	கூறெடுப்பு
Sampling distribution	கூறுபரவல்
Sampling Errors	கூறுபிழை
Simple Random Sampling	எளியசமவாய்ப்புகூறெடுப்பு
Standard Error	திட்டபிழை
Statistic	கூறு பண்பளவை / மாதிரிப் பண்பளவை
Statistical Inference	புள்ளியியல் அனுமானம்
Stratified Random Sampling	படுகைசமவாய்ப்புகூறெடுப்பு
Systematic Sampling	முறைசார் கூறெடுத்தல்

ICT Corner

Expected Result is shown in this picture

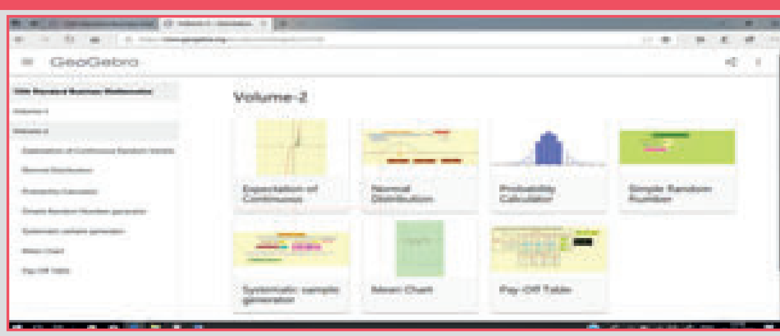


Step – 1 : Open the Browser, type the URL Link given below (or) Scan the QR Code. GeoGebra work Book named "12th Standard Business Mathematics " will open. In the work book there are two Volumes. Select "Volume-2".

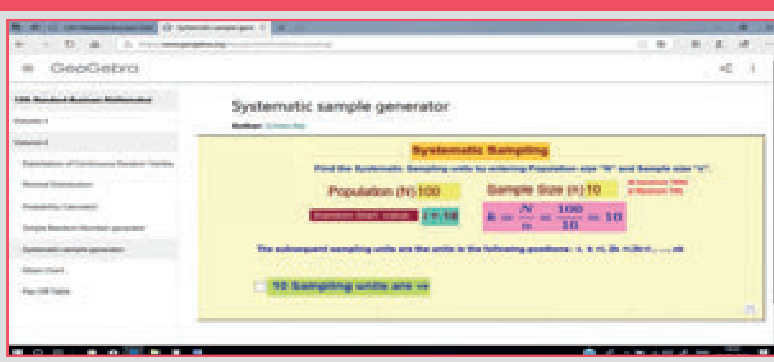
Step - 2 : Select the worksheet named "Systematic Sampling- Generator."

Enter the Population size, Sample size and random start value in the respective boxes. Then click on the Sampling units to view the Systematic sample units.

Step 1



Step 2



Scan the QR Code.

B284_12_BUS_MAT_EM

9 Applied Statistics

Walter Shewhart (March 18, 1891-March 11, 1967)

Introduction

The term Applied Statistics refers to the use of statistical theory to conduct operational activities in a variety of fields in real life situations. Today, the applications of statistics are an indispensable part of any and every activity. There are many fields in which statistical concepts can be applied, some of them are business decision making, finance, marketing, economics, social sciences, industry, agriculture etc… An important aspect of applied statistics is to study about the present and future behaviour of the activities performed in an industry.

Walter Andrew Shewhart (pronounced like "shoe-heart",) was an American physicist, engineer and statistician, sometimes known as the father of statistical quality control and also related to the Shewhart cycle.

In this chapter we would be studying about the theoretical and application of the statistical methods of Time Series, Index Number and Statistical Quality Control. Each one of them has its importance in its field of application. Statistical analysis has been widely used for scientific research, surveys, experiments etc. The reliability of the interpretation of the statistical analysis depends upon the informations collected and represented.

Learning Objectives

After studying this chapter the students are able to understand

- Time series data
- Components of Time Series
- Moving Averages
- Seasonal Variation
- Index Numbers

YP1GD5

- Weighted Index Number
- Tests for an Ideal Index Number
- Statistical Quality Control
- Causes for Variation
- Process Control and Product Control

9.1 Time Series Analysis

Time Series analysis is one of the statistical methods used to determine the patterns in data collected for a period of time. Generally, each of us should know about the past data to observe and understand the changes that have taken place in the past and current time. One can also identify the regular or irregular occurrence of any specific feature over a time period in a time series data. Most of the time series data relates to fields like Economics, Business, Commerce, etc… For example Production of a product, Cost of a product, Sales of a product, National income, Salary of an individual, etc.. By close observation of time series data, one can predict and plan for future operations in industries and other fields.

Definition 9.1

A Time Series consists of data arranged chronologically – *Croxton & Cowden.*

When quantitative data are arranged in the order of their occurrence, the resulting series is called the Time Series – *Wessel & Wallet.*

9.1.1 Meaning, Uses and Basic Components

Meaning:

A time series consists of a set of observations arranged in chronological order (either ascending or descending). Time Series has an important objective to identify the variations and try to eliminate the variations and also helps us to estimate or predict the future values.

Why should we learn Time Series?

It helps in the analysis of the past behavior.

It helps in forecasting and for future plans.

It helps in the evaluation of current achievements.

It helps in making comparative studies between one time period and others.

Therefore time series helps us to study and analyze the time related data which involves in business fields, economics, industries, etc…

Components of Time Series

There are four types of components in a time series. They are as follows;

(i) Secular Trend (ii) Seasonal variations

(iii) Cyclic variations (iv) Irregular variations

(i) Secular Trend

It is a general tendency of time series to increase or decrease or stagnates during a long period of time. An upward tendency is usually observed in population of a country, production, sales, prices in industries, income of individuals etc., A downward tendency is observed in deaths, epidemics, prices of electronic gadgets, water sources, mortality rate etc…. It is not necessarily that the increase or decrease should be in the same direction throughout the given period of time.

(ii) Seasonal Variations

As the name suggests, tendency movements are due to nature which repeat themselves periodically in every seasons. These variations repeat themselves in less than one year time. It is measured in an interval of time. Seasonal variations may be influenced by natural force, social customs and traditions. These variations are the results of such factors which uniformly and regularly rise and fall in the magnitude. For example, selling of umbrellas' and raincoat in the rainy season, sales of cool drinks in summer season, crackers in Deepawali season, purchase of dresses in a festival season, sugarcane in Pongal season.

(iii) Cyclic Variations

These variations are not necessarily uniformly periodic in nature. That is, they may or may not follow exactly similar patterns after equal intervals of time. Generally one cyclic period ranges from 7 to 9 years and there is no hard and fast rule in the fixation of years for a cyclic period. For example, every business cycle has a Start- Boom- Depression- Recover, maintenance during booms and depressions, changes in government monetary policies, changes in interest rates.

(iv) Irregular Variations

These variations do not have particular pattern and there is no regular period of time of their occurrences. These are accidently changes which are purely random or unpredictable. Normally they are short-term variations, but its occurrence sometimes has its effect so intense that they may give rise to new cyclic or other movements of variations. For example floods, wars, earthquakes, Tsunami, strikes, lockouts etc…

Mathematical Model for a Time Series

There are two common models used for decomposition of a time series into its components, namely additive and multiplicative model.

(i) Additive Model:

This model assumes that the observed value is the sum of all the four components of time series. (i.e) Y= T+S+C+I

where Y = Original value , T = Trend Value , S = Seasonal component

C = Cyclic component , I = Irregular component

The additive model assumes that all the four components operate independently. It also assumes that the behavior of components is of an additive character.

(ii) Multiplicative Model:

This model assumes that the observed value is obtained by multiplying the trend(T) by the rates of other three components. $Y = T \times S \times C \times I$

where Y = Original value , T = Trend Value , S = Seasonal component

C = Cyclic component , I = Irregular component

This model assumes that the components due to different causes are not necessarily independent and they can affect one another. It also assumes that the behavior of components is of a multiplicative character.

9.1.2 Measurements of Trends

Following are the methods by which we can measure the trend.

(i) Freehand or Graphic Method.

(ii) Method of Semi-Averages.

(iii) Method of Moving Averages.

(iv) Method of Least Squares.

(i) Freehand or Graphic Method.

It is the simplest and most flexible method for estimating a trend. We will see the working procedure of this method.

Procedure:

(a) Plot the time series data on a graph.
(b) Draw a freehand smooth curve joining the plotted points.
(c) Examine the direction of the trend based on the plotted points.
(d) Draw a straight line which will pass through the maximum number of plotted points.

Example 9.1

Fit a trend line by the method of freehand method for the given data.

Year	2000	2001	2002	2003	2004	2005	2006	2007
Sales	30	46	25	59	40	60	38	65

Solution:

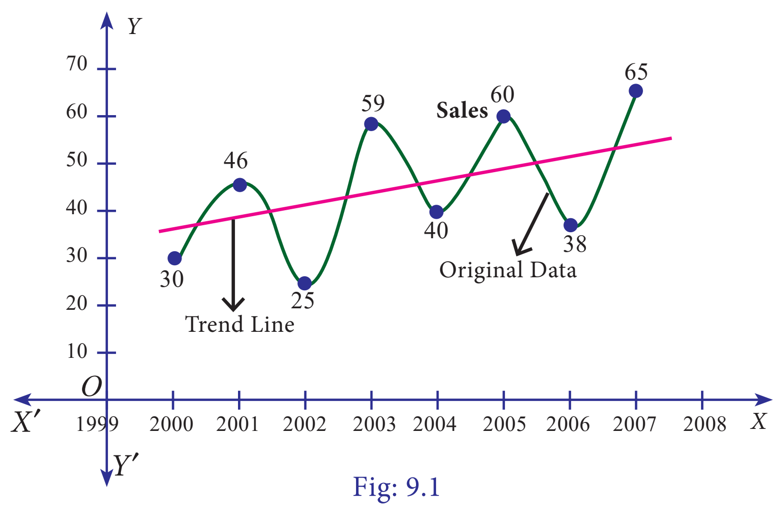


Fig: 9.1

Note

The trend drawn by the freehand method can be extended to predict the future values of the given data. However, this method is subjective in nature, predictions obtained by this method depends on the personal bias and judgement of the investigator handling the data.

(ii) Method of Semi-Averages

In this method, the semi-averages are calculated to find out the trend values. Now, we will see the working procedure of this method.

Procedure:

(i) The data is divided into two equal parts. In case of odd number of data, two equal parts can be made simply by omitting the middle year.

(ii) The average of each part is calculated, thus we get two points.

(iii) Each point is plotted at the mid-point (year) of each half.

(iv) Join the two points by a straight line.

(v) The straight line can be extended on either side.

(vi) This line is the trend line by the methods of semi-averages.

Example 9.2

Fit a trend line by the method of semi-averages for the given data.

Year	2000	2001	2002	2003	2004	2005	2006
Production	105	115	120	100	110	125	135

Solution:

Since the number of years is odd(seven), we will leave the middle year's production value and obtain the averages of first three years and last three years.

Year	Production	Average
2000	105	
2001	115	$\dfrac{105+115+120}{3}=113.33$
2002	120	
2003	100 (left out)	
2004	110	
2005	125	$\dfrac{110+125+135}{3}=123.33$
2006	135	

Table 9.1

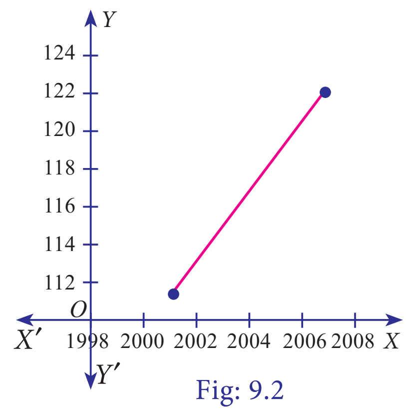


Fig: 9.2

Example 9.3

Fit a trend line by the method of semi-averages for the given data.

Year	1990	1991	1992	1993	1994	1995	1996	1997
Sales	15	11	20	10	15	25	35	30

Solution:

Since the number of years is even(eight), we can equally divide the given data it two equal parts and obtain the averages of first four years and last four years.

Year	Production	Average
1990	15	$\dfrac{15+11+20+10}{4}=14$
1991	11	
1992	20	
1993	10	
1994	15	
1995	25	
1996	35	$\dfrac{15+25+35+30}{4}=26.25$
1997	30	

Table 9.2

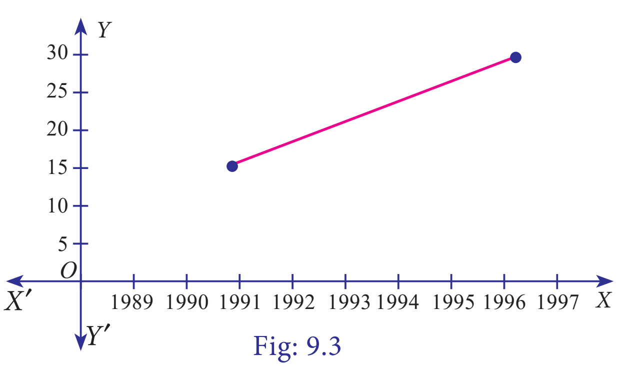


Fig: 9.3

Note

(i) The future values can be predicted.

(ii) The trend values obtained by this method and the predicted values are not precise.

9.1.3 Method of Moving Averages

Moving Averages Method gives a trend with a fair degree of accuracy. In this method, we take arithmetic mean of the values for a certain time span. The time span can be three-years, four -years, five- years and so on depending on the data set and our interest. We will see the working procedure of this method.

Procedure:

(i) Decide the period of moving averages (three- years, four -years).

(ii) In case of odd years, averages can be obtained by calculating,

$$\frac{a+b+c}{3}, \frac{b+c+d}{3}, \frac{c+d+e}{3}, \frac{d+e+f}{3}, \ldots\ldots\ldots$$

(iii) If the moving average is an odd number, there is no problem of centering it, the average value will be centered besides the second year for every three years.

(iv) In case of even years, averages can be obtained by calculating,

$$\frac{a+b+c+d}{4}, \frac{b+c+d+e}{4}, \frac{c+d+e+f}{4}, \frac{d+e+f+g}{4}, \ldots$$

(v) If the moving average is an even number, the average of first four values will be placed between 2nd and 3rd year, similarly the average of the second four values will be placed between 3rd and 4th year. These two averages will be again averaged and placed in the 3rd year. This continues for rest of the values in the problem. This process is called as centering of the averages.

Example 9.4

Calculate three-yearly moving averages of number of students studying in a higher secondary school in a particular village from the following data.

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Number of students	332	317	357	392	402	405	410	427	435	438

Solution:

Computation of three- yearly moving averages.

Year	Number of students	3- yearly moving Total	3- yearly moving Averages
1995	332	---	---
1996	317	1006	335.33
1997	357	1066	355.33
1998	392	1151	383.67
1999	402	1199	399.67
2000	405	1217	405.67
2001	410	1242	414.00
2002	427	1272	424.00
2003	435	1300	433.33
2004	438	---	---

Table 9.3

Example 9.5

Calculate four-yearly moving averages of number of students studying in a higher secondary school in a particular city from the following data.

Year	2001	2002	2003	2004	2005	2006	2007	2008
Sales	124	120	135	140	145	158	162	170

Solution:

Computation of four- yearly moving averages.

Year	Sales	4-yearly centered moving total	4-yearly moving Average	4-yearly centered moving Average
2001	124	--	--	--
2002	120	--	--	--
		519	129.75	
2003	135	--		132.37
		540	135.00	
2004	140	--		139.75
		578	144.50	
2005	145	--		147.87
		605	151.25	
2006	158	--		155.00
		635	158.75	
2008	162	--		162.50
		665	166.25	
2007	170	--	--	-
2008	175	--	--	-

Table 9.4

Note

The calculated 4-yearly centered moving average belongs to the particular year present in that row eg; 132.37 belongs to the year 2003.

9.1.4 Method of Least Squares

The line of best fit is a line from which the sum of the deviations of various points is zero. This is the best method for obtaining the trend values. It gives a convenient basis for calculating the line of best fit for the time series. It is a mathematical method for measuring trend. Further the sum of the squares of these deviations would be least when compared with other fitting methods. So, this method is known as the Method of Least Squares and satisfies the following conditions:

(i) The sum of the deviations of the actual values of Y and $\hat{Y}$ (estimated value of Y) is Zero. that is $\Sigma(Y-\hat{Y}) = 0$.

(ii) The sum of squares of the deviations of the actual values of Y and $\hat{Y}$ (estimated value of Y) is least. that is $\Sigma(Y-\hat{Y})^2$ is least ;

Procedure:

(i) The straight line trend is represented by the equation $Y = a + bX$ ...(1)

where Y is the actual value, X is time, a, b are constants

(ii) The constants 'a' and 'b' are estimated by solving the following two normal

Equations $\Sigma Y = n\,a + b\,\Sigma X$ ...(2)

$\Sigma XY = a\,\Sigma X + b\,\Sigma X^2$ ...(3)

Where 'n' = number of years given in the data.

(iii) By taking the mid-point of the time as the origin, we get $\Sigma X = 0$

(iv) When $\Sigma X = 0$, the two normal equations reduces to

$$\Sigma Y = n\,a + b\,(0) \quad ;\ a = \frac{\sum Y}{n} = \overline{Y}$$

$$\Sigma XY = a(0) + b\,\Sigma X^2 \quad ;\ b = \frac{\sum XY}{\sum X^2}$$

The constant 'a' gives the mean of Y and 'b' gives the rate of change (slope).

DO YOU KNOW?

The constant 'a' takes only positive values, but constant 'b' takes both positive and negative values

(v) By substituting the values of 'a' and 'b' in the trend equation (1), we get the Line of Best Fit.

Example 9.6

Given below are the data relating to the production of sugarcane in a district.

Fit a straight line trend by the method of least squares and tabulate the trend values.

Year	2000	2001	2002	2003	2004	2005	2006
Prod. of Sugarcane	40	45	46	42	47	50	46

Solution:

Computation of trend values by the method of least squares (ODD Years).

Year(x)	Production of Sugarcane(Y)	X=(x–2003)	X^2	XY	Trend values(Yt)
2000	40	–3	9	–120	42.04
2001	45	–2	4	–90	43.07
2002	46	–1	1	–46	44.11
2003	42	0	0	0	45.14
2004	47	1	1	47	46.18
2005	50	2	4	100	47.22
2006	46	3	9	138	48.25
N= 7	ΣY = 316	ΣX = 0	ΣX^2= 28	ΣXY=29	$\Sigma\ Yt$ = 316

Table 9.5

$$a=\frac{\sum Y}{n}=\frac{316}{7}=45.143\ ;\ b=\frac{\sum XY}{\sum X^2}=\frac{29}{28}=1.036$$

Therefore, the required equation of the straight line trend is given by

$$Y = a + bX \quad ;$$

$$Y = 45.143 + 1.036\ (x - 2003)$$

The trend values can be obtained by

When $X = 2000$, $Yt = 45.143 + 1.036(2000–2003) = 42.035$

When $X = 2001$, $Yt = 45.143 + 1.036(2001–2003) = 43.071$,

similarly other values can be obtained.

Example 9.7

Given below are the data relating to the sales of a product in a district.

Fit a straight line trend by the method of least squares and tabulate the trend values.

Year	1995	1996	1997	1998	1999	2000	2001	2002
Sales	6.7	5.3	4.3	6.1	5.6	7.9	5.8	6.1

Solution:

Computation of trend values by the method of least squares.

In case of EVEN number of years, let us consider

$$X=\frac{(x-Arithimetic\ mean\ of\ two\ middle\ years)}{0.5}$$

Year(x)	Sales(Y)	$X = \dfrac{(x\ -1998.5)}{0.5}$	XY	X^2	Trend values(Y_t)
1995	6.7	–7	46.9	49	5.6166
1996	5.3	–5	26.5	25	5.7190
1997	4.3	–3	12.9	9	5.8214

1998	6.1	–1	6.1	1	5.9238
1999	5.6	7	39.2	49	6.0261
2000	7.9	5	39.5	25	6.1285
2001	5.8	3	17.4	9	6.2309
2002	6.1	1	6.1	1	6.3333
N=8	47.8	$\Sigma X = 0$	194.6	168	

Table 9.6

$$a = \frac{\sum Y}{n} = \frac{47.8}{8} = 5.975 \quad ; \; b = \frac{\sum XY}{\sum X^2} = \frac{4.3}{42} = 0.10238$$

Therefore, the required equation of the straight line trend is given by

$$Y = a + bX \quad ; \; Y = 5.975 + 0.10238\, X.$$

When $X = 1995$, $Y_t = 5.975 + 0.10238 \left(\frac{1995 - 1998.5}{0.5}\right) = 5.6166$

When $X = 1996$, $Y_t = 5.975 + 0.10238 \left(\frac{1996 - 1998.5}{0.5}\right) = 5.7190$

similarly other values can be obtained.

Note

(i) Future forecasts made by this method are based only on trend values.

(ii) The predicted values are more reliable in this method than the other methods.

9.1.5 Methods of measuring Seasonal Variations By Simple Averages :

Seasonal Variations can be measured by the method of simple average. The data should be available in season wise likely weeks, months, quarters.

Method of Simple Averages:

This is the simplest and easiest method for studying Seasonal Variations. The procedure of simple average method is outlined below.

Procedure:

(i) Arrange the data by months, quarters or years according to the data given.

(ii) Find the sum of the each months, quarters or year.

(iii) Find the average of each months, quarters or year.

(iv) Find the average of averages, and it is called Grand Average (G)

(v) Compute Seasonal Index for every season (i.e) months, quarters or year is given by

$$\text{Seasonal Index (S.I)} = \frac{Seasonal\ Average}{Grand\ average} \times 100$$

(vi) If the data is given in months

$$\text{S.I for Jan} = \frac{Monthly\ Average\,(for\ Jan)}{Grand\ average} \times 100$$

$$\text{S.I for Feb} = \frac{Monthly\ Average\,(for\ Feb)}{Grand\ average} \times 100$$

Similarly we can calculate SI for all other months.

(vii) If the data is given in quarter

$$\text{S.I for I Quarter} = \frac{Average\ of\ I\ quarter}{Grand\ average} \times 100$$

$$\text{S.I for II Quarter} = \frac{Average\ of\ II\ quarter}{Grand\ average} \times 100$$

$$\text{S.I for III Quarter} = \frac{Average\ of\ III\ quarter}{Grand\ average} \times 100$$

$$\text{S.I for IV Quarter} = \frac{Average\ of\ IV\ quarter}{Grand\ average} \times 100$$

Example 9.8

Calculate the seasonal index for the monthly sales of a product using the method of simple averages.

Months / Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
2001	15	41	25	31	29	47	41	19	35	38	40	30
2002	20	21	27	19	17	25	29	31	35	39	30	44
2003	18	16	20	28	24	25	30	34	30	38	37	39

Solution: Computation of seasonal Indices by method of simple averages.

months / Years	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
2001	15	41	25	31	29	47	41	19	35	38	40	30
2002	20	21	27	19	17	25	29	31	35	39	30	44
2003	18	16	20	28	24	25	30	34	30	38	37	39
Monthly Total	53	78	72	78	70	97	100	84	100	115	107	113
Monthly Averages	17.67	26	24	26	23.33	32.33	33.33	28	33.33	38.33	35.67	37.67

Table 9.7

$$\text{S.I for Jan} = \frac{Monthly\ Average\,(for\ Jan)}{Grand\,average} \times 100$$

$$\text{Grand Average} = \frac{355.582}{12} = 29.63$$

$$\text{S.I for Jan} = \frac{17.666}{29.361} X100 = 59.62\,; \quad \text{S.I for Feb} = \frac{26}{29.361} X100 = 87.77\,;$$

Similarly other seasonal index values can be obtained.

Months	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
Seasonal Index	59.62	87.77	80.99	87.77	78.74	109.12	112.49	94.49	112.49	129.36	120.36	126.89

Example 9.9

Calculate the seasonal index for the quarterly production of a product using the method of simple averages.

Year	I Quarter	II Quarter	III Quarter	IV Quarter
2005	255	351	425	400
2006	269	310	396	410
2007	291	332	358	395
2008	198	289	310	357
2009	200	290	331	359
2010	250	300	350	400

Solution :

Computation of Seasonal Index by the method of simple averages.

Year	I Quarter	II Quarter	III Quarter	IV Quarter
2005	255	351	425	400
2006	269	310	396	410
2007	291	332	358	395
2008	198	289	310	357
2009	200	290	331	359
2010	250	300	350	400
Quarterly Total	1463	1872	2170	2321
Quarterly Averages	243.83	312	361.67	386.83

Table 9.8

$$\text{S.I for I Quarter} = \frac{\textit{Average of I quarter}}{\textit{Grand average}} \times 100$$

$$\text{Grand Average} = \frac{1304.333}{4} = 326.0833$$

$$\text{S.I for I Q} = \frac{243.8333}{326.0833} \times 100 = 74.77\,; \quad \text{S.I for II Q} = \frac{312}{326.0833} \times 100 = 95.68\,;$$

$$\text{S.I for III Q} = \frac{361.6667}{326.0833} \times 100 = 110.91\,; \quad \text{S.I for IV Q} = \frac{386.833}{326.0833} \times 100 = 118.63$$

Exercise 9.1

1. Define Time series.
2. What is the need for studying time series?
3. State the uses of time series.
4. Mention the components of the time series.
5. Define secular trend.
6. Write a brief note on seasonal variations
7. Explain cyclic variations
8. Discuss about irregular variation
9. Define seasonal index.
10. Explain the method of fitting a straight line.
11. State the two normal equations used in fitting a straight line.
12. State the different methods of measuring trend.
13. Compute the average seasonal movement for the following series

year	Quarterly Production			
	I	II	III	IV
2002	3.5	3.8	3.7	3.5
2003	3.6	4.2	3.4	4.1
2004	3.4	3.9	3.7	4.2
2005	4.2	4.5	3.8	4.4
2006	3.9	4.4	4.2	4.6

14. The following figures relates to the profits of a commercial concern for 8 years

Year	1986	1987	1988	1989	1990	1991	1992	1993
Profit (₹)	15,420	15,470	15,520	21,020	26,500	31,950	35,600	34,900

Find the trend of profits by the method of three yearly moving averages.

15. Find the trend of production by the method of a five-yearly period of moving average for the following data:

Year	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Production('000)	126	123	117	128	125	124	130	114	122	129	118	123

16. The following table gives the number of small-scale units registered with the Directorate of Industries between 1985 and 1991. Show the growth on a trend line by the free hand method.

Year	1985	1986	1987	1988	1989	1990	1991	1992
No. of units (in'000)	10	22	36	62	55	40	34	50

17. The annual production of a commodity is given as follows :

Year	1995	1996	1997	1998	1999	2000	2001
Production (in tones)	155	162	171	182	158	180	178

Fit a straight line trend by the method of least squares.

18. Determine the equation of a straight line which best fits the following data

Year	2000	2001	2002	2003	2004
Sales (₹ '000)	35	36	79	80	40

Compute the trend values for all years from 2000 to 2004

19. The sales of a commodity in tones varied from January 2010 to December 2010 as follows:

in year 2010	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sales (in tones)	280	240	270	300	280	290	210	200	230	200	230	210

Fit a trend line by the method of semi-average.

20. Use the method of monthly averages to find the monthly indices for the following data of production of a commodity for the years 2002, 2003 and 2004.

2002	15	18	17	19	16	20	21	18	17	15	14	18
2003	20	18	16	13	12	15	22	16	18	20	17	15
2004	18	25	21	11	14	16	19	20	17	16	18	20

21. Calculate the seasonal indices from the following data using the average from the following data using the average method:

	I Quarterly	II Quarterly	III Quarterly	IV Quarterly
2008	72	68	62	76
2009	78	74	78	72
2010	74	70	72	76
2011	76	74	74	72
2012	72	72	76	68

22. The following table shows the number of salesmen working for a certain concern:

Year	1992	1993	1994	1995	1996
No. of salesmen	46	48	42	56	52

Use the method of least squares to fit a straight line and estimate the number of salesmen in 1997.

9.2 INDEX NUMBER

Introduction:

Index Numbers are the indicators which reflect the changes over a specified period of time in price of different commodities, production, sales, cost of living etc... Index Numbers are statistical methods used to measure the relative change in the level of a variable or group of variables with respect to time, geographical location or other characteristics such as income, profession etc. The variables may be

(i) The price of a particular commodity. For example gold, silver, iron (or) a group of commodities. For example consumer goods, household food items etc..

(ii) The volume of export and import, agricultural and industrial production.

(iii) National income of a country, cost of living of persons belonging to a particular income group.

9.2.1 Meaning, Classifications and Uses

Suppose we want to measure the general changes in the price level of consumer goods, the price changes are not directly measurable, as the prices of various commodities are in different units For example rice, wheat and sugar are in kilograms, milk, petrol, oil are in litres, clothes are in metres etc… Further, the price and quantity of some commodities may increase or decrease during the two time periods. Therefore, index number gives a single representative value which gives the general level of the prices of the commodities in a given group over a specified time period.

Definition 9.2

"An Index Number is a device which shows by its variations the Changes in a magnitude which is not capable of accurate measurements in itself or of direct valuation in practice". - *Wheldon*

"An Index number is a statistical measure of fluctuations in a variable arranged in the form of a series and using a base period for making comparisons"
– *Lawrence J Kalpan*

Classification of Index Numbers:

Index number can be classified as follows,

(i) Price Index Number

It measures the general changes in the retail or wholesale price level of a particular or group of commodities.

(ii) Quantity Index Number

These are indices to measure the changes in the quantity of goods manufactured in a factory.

(iii) Cost of living Index Number

These are intended to study the effect of change in the price level on the cost of living of different classes of people.

Uses of Index number

(i) It is an important tool for the formulating decision and management policies.

(ii) It helps in studying the trends and tendencies.

(iii) It determines the inflation and deflation in an economy.

Construction of Index Number

There are two types in construction of index number.

(i) Unweighted Index Number (ii) Weighted Index Number

We confine ourselves to Weighted Index Number.

9.2.2 Weighted Index Number

In general, all the commodities cannot be given equal importance, so we can assign weights to each commodity according to their importance and the index number computed from these weights are called as weighted index number. The weights can be production, consumption values. If 'w' is the weight attached to a commodity, then the price index is given by,

Price Index (P_{01}) = $\dfrac{\sum p_1 w}{\sum p_0 w}\times 100$, Let us consider the following notations,

p_1 - current year price $\qquad$ p_0 - base year price

q_1 - current year quantity $\qquad$ q_0 - base year quantity

where suffix '0' represents base year and '1' represents current year.

Laspeyre's price index number

$$P_{01}^{L}=\frac{\sum p_1 q_0}{\sum p_0 q_0}\times 100$$

Paasche's price index number

$$P_{01}^{P}=\frac{\sum p_1 q_1}{\sum p_0 q_1}\times 100$$

Fisher's price index number

$$P_{01}^{F}=\sqrt{P_{01}^{L}\times P_{01}^{P}}$$

$$P_{01}^{F}=\sqrt{\frac{\sum p_1q_0\times\sum p_1q_1}{\sum p_0q_0\times\sum p_0q_1}}\times 100$$

Note

To get exact Fisher's price index number, one should use formula method rather than using $P_{01}^{F}=\sqrt{P_{01}^{L}\times P_{01}^{P}}$.

In Laspeyre's price index number, the quantity of the base year is used as weight.

In Paasche's price index number, the quantity of the current year is used as weight.

Example 9.10

Calculate the Laspeyre's, Paasche's and Fisher's price index number for the following data. Interpret on the data.

Commodities	Price		Quantity	
	2000	2010	2000	2010
Rice	38	35	6	7
Wheat	12	18	7	10
Rent	10	15	10	15
Fuel	25	30	12	16
Miscellaneous	30	33	8	10

DO YOU KNOW?

Fisher's price index number is the geometric mean between Laspeyre's and Paasche's price index number

Solution

Commodities	Price		Quantity		$p_0 q_0$	$p_0 q_1$	$p_1 q_0$	$p_1 q_1$
	2000 (p_0)	2010 (p_1)	2000 (q_0)	2010 (q_1)				
Rice	38	35	6	7	228	266	210	245
Wheat	12	18	7	10	84	120	126	180
Rent	10	15	10	15	100	150	150	225
Fuel	25	30	12	16	300	400	360	480
Miscellaneous	30	33	8	10	240	300	264	330
				Total	**952**	**1236**	**1110**	**1460**

Table 9.9

Laspeyre's price index number

$$P_{01}^{L} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 = \frac{1110}{952} \times 100 = 116.60$$

On an average, there is an increase of 16.60 % in the price of the commodities when the year 2000 compared with the year 2010.

Paasche's price index number

$$P_{01}^{P} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100 = \frac{1460}{1236} \times 100 = 118.12$$

On an average, there is an increase of 18.12 % in the price of the commodities when the year 2000 compared with the year 2010.

Fisher's price index number

$$P_{01}^{F} = \sqrt{\frac{\sum p_1 q_0 \times \sum p_1 q_1}{\sum p_0 q_0 \times \sum p_0 q_1}} \times 100 = \sqrt{\frac{1110 \times 1460}{952 \times 1236}} \times 100 = 117.36$$

On an average, there is an increase of 17.36 % in the price of the commodities when the year 2000 compared with the year 2010.

Example 9.11

Construct the Laspeyre's , Paasche's and Fisher's price index number for the following data. Comment on the result.

Commodities	Base Year		Current Year	
	Price	Quantity	Price	Quantity
Rice	15	5	16	8
Wheat	10	6	18	9
Rent	8	7	15	8
Fuel	9	5	12	6
Transport	11	4	11	7
Miscellaneous	16	6	15	10

Solution:

Commodities	Base Year		Current Year		$p_0 q_0$	$p_0 q_1$	$p_1 q_0$	$p_1 q_1$
	Price (p_0)	Quantity (p_0)	Price	Quantity				
Rice	15	5	16	8	75	120	80	128
Wheat	10	6	18	9	60	90	108	162
Rent	8	7	15	8	56	64	105	120
Fuel	9	5	12	6	45	54	60	72
Transport	11	4	11	7	44	77	44	77
Miscellaneous	16	6	15	10	96	160	90	150
				Total	**376**	**565**	**487**	**709**

Table 9.10

Laspeyre's price index number

$$P_{01}^{L} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 = \frac{487}{376} \times 100 = 129.5212$$

Paasche's price index number

$$P_{01}^{P} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100 = \frac{709}{565} \times 100 = 125.4867$$

Fisher's price index number

$$P_{01}^{F} = \left(\sqrt{\frac{\sum p_1 q_0 \times \sum p_1 q_1}{\sum p_0 q_0 \times \sum p_0 q_1}}\right) \times 100 = \left(\sqrt{\frac{487 \times 709}{376 \times 565}}\right) \times 100 = 127.4879$$

On an average, there is an increase of 29.52 % , 25.48% and 27.48% in the price of the commodities by Laspeyre's, Paasche's, Fisher's price index number respectively, when the base year compared with the current year.

9.2.3 Test of adequacy for an Index Number

Index numbers are studied to know the relative changes in price and quantity for any two years compared. There are two tests which are used to test the adequacy for an index number. The two tests are as follows,

(i) Time Reversal Test (ii) Factor Reversal Test

The criterion for a good index number is to satisfy the above two tests.

Time Reversal Test

It is an important test for testing the consistency of a good index number. This test maintains time consistency by working both forward and backward with respect to time (here time refers to base year and current year). Symbolically the following relationship should be satisfied, $P_{01} \times P_{10} = 1$

Fisher's index number formula satisfies the above relationship

$$P_{01}^{F} = \sqrt{\frac{\sum p_1q_0 \times \sum p_1q_1}{\sum p_0q_0 \times \sum p_0q_1}}$$

when the base year and current year are interchanged, we get

$$P_{10}^{F} = \sqrt{\frac{\sum p_0q_1 \times \sum p_0q_0}{\sum p_1q_1 \times \sum p_1q_0}}$$

$$P_{01}^{F} \times P_{10}^{F} = 1$$

Note

Ignore the factor 100 in each Index number

Factor Reversal Test

This is another test for testing the consistency of a good index number. The product of price index number and quantity index number from the base year to the current year should be equal to the true value ratio. That is, the ratio between the total value of current period and total

value of the base period is known as true value ratio. Factor Reversal Test is given by,

$$P_{01} \times Q_{01} = \frac{\sum p_1q_1}{\sum p_0q_0}$$

where $$P_{01} = \sqrt{\frac{\sum p_1q_0 \times \sum p_1q_1}{\sum p_0q_0 \times \sum p_0q_1}}$$

Now interchanging P by Q, we get

$$Q_{01} = \sqrt{\frac{\sum q_1p_0 \times \sum q_1p_1}{\sum q_0p_0 \times \sum q_0p_1}}$$

where P_{01} is the relative change in price.

Q_{01} is the relative change in quantity.

Example 9.12

Calculate Fisher's price index number and show that it satisfies both Time Reversal Test and Factor Reversal Test for data given below.

Commodities	Price		Quantity	
	2003	2009	2003	2009
Rice	10	13	4	6
Wheat	15	18	7	8
Rent	25	29	5	9
Fuel	11	14	8	10
Miscellaneous	14	17	6	7

DO YOU KNOW?

Since Fisher's Price Index Number satisfies both TRT & FRT, it is termed as an Ideal Index Number

Solution:

Commodities	Price		Quantity		p_0q_0	p_0q_1	p_1q_0	p_1q_1
	2003 (p_0)	2009 (p_1)	2003 (q_0)	2009 (q_1)				
Rice	10	13	4	6	40	60	52	78
Wheat	15	18	7	8	105	120	126	144
Rent	25	29	5	9	125	225	145	261
Fuel	11	14	8	10	88	110	140	140
Miscellaneous	14	17	6	7	84	98	102	119
				Total	442	613	565	742

Table 9.11

Fisher's price index number

$$P_{01}^{F}=\left(\sqrt{\frac{\sum p_1q_0 \times \sum p_1q_1}{\sum p_0q_0 \times \sum p_0q_1}}\right)\times 100=\left(\sqrt{\frac{565\times 742}{442\times 613}}\right)\times\ 100 = 124.3898$$

Time Reversal Test: $P_{01}\times P_{10}=1$

$$P_{01}\times P_{10}=\sqrt{\left(\frac{\sum p_1q_0 \times \sum p_1q_1 \times \sum p_0q_1 \times \sum p_0q_0}{\sum p_0q_0 \times \sum p_0q_1 \times \sum p_1q_1 \times \sum p_1q_0}\right)}$$

$$P_{01}\times P_{10}\ =\sqrt{\left(\frac{565\times 742\times 613\times 442}{442\times 613\times 742\times 565}\right)}$$

$$P_{01}\times P_{10}\ = 1$$

Factor Reversal Test

$$P_{01}\ \times\ Q_{01}=\frac{\sum p_1q_1}{\sum p_0q_0}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{\sum p_1q_0 \times \sum p_1q_1 \times \sum q_1p_0 \times \sum q_1p_1}{\sum p_0q_0 \times \sum p_0q_1 \times \sum q_0p_0 \times \sum q_0p_1}\right)}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{565 \times 742 \times 613 \times 742}{442 \times 613 \times 442 \times 565}\right)}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{742 \times 742}{442 \times 442}\right)} = \frac{742}{442} \qquad \Rightarrow P_{01} \times Q_{01} = \frac{\sum p_1q_1}{\sum p_0q_0}$$

Example 9.13

Calculate Fisher's price index number and show that it satisfies both Time Reversal Test and Factor Reversal Test for data given below.

Commodities	Base Year		Current Year	
	Price	Quantity	Price	Quantity
Rice	10	5	11	6
Wheat	12	6	13	4
Rent	14	8	15	7
Fuel	16	9	17	8
Transport	18	7	19	5
Miscellaneous	20	4	21	3

Solution:

Commodities	Base Year		Current Year		p_0q_0	p_0q_1	p_1q_0	p_1q_1
	Price (p_0)	Quantity (q_0)	Price (p_1)	Quantity (q_1)				
Rice	10	5	11	6	50	60	55	66
Wheat	12	6	13	4	72	48	78	52
Rent	14	8	15	7	112	98	120	105
Fuel	16	9	17	8	144	128	153	136
Transport	18	7	19	5	126	90	133	95
Miscellaneous	20	4	21	3	80	60	84	63
				Total	584	484	623	517

Table 9.12

Fisher's price index number

$$P_{01}^{F} = \left(\sqrt{\frac{\sum p_1q_0 \times \sum p_1q_1}{\sum p_0q_0 \times \sum p_0q_1}}\right) \times 100 = \left(\sqrt{\frac{623 \times 517}{584 \times 484}}\right) \times 100 = 106.74$$

Time Reversal Test: $P_{01} \times P_{10} = 1$

$$P_{01} \times P_{10} = \sqrt{\left(\frac{\sum p_1 q_0 \times \sum p_1 q_1 \times \sum p_0 q_1 \times \sum p_0 q_0}{\sum p_0 q_0 \times \sum p_0 q_1 \times \sum p_1 q_1 \times \sum p_1 q_0} \right)}$$

$$P_{01} \times P_{10} = \sqrt{\left(\frac{623 \times 517 \times 484 \times 584}{584 \times 484 \times 517 \times 623} \right)}$$

$$P_{01} \times P_{10} = 1$$

Factor Reversal Test

$$P_{01} \times Q_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{\sum p_1 q_0 \times \sum p_1 q_1 \times \sum q_1 p_0 \times \sum q_1 p_1}{\sum p_0 q_0 \times \sum p_0 q_1 \times \sum q_0 p_0 \times \sum q_0 p_1} \right)}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{623 \times 517 \times 484 \times 517}{584 \times 484 \times 584 \times 623} \right)}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{517 \times 517}{585 \times 584} \right)} = \frac{517}{584}$$

$$P_{01} \times Q_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

Example 9.14

Construct Fisher's price index number and prove that it satisfies both Time Reversal Test and Factor Reversal Test for data following data.

Commodities	Base Year		Current Year	
	Price	Quantity	Price	Quantity
Rice	40	5	48	4
Wheat	45	2	42	3
Rent	90	4	95	6
Fuel	85	3	80	2
Transport	50	5	65	8
Miscellaneous	65	1	72	3

Solution:

Commodities	Base Year		Current Year		p_0q_0	p_0q_1	p_1q_0	p_1q_1
	Price (p_0)	Quantity (q_0)	Price (p_1)	Quantity (q_1)				
Rice	40	5	48	4	200	160	240	192
Wheat	45	2	42	3	90	135	84	126
Rent	90	4	95	6	360	540	380	570
Fuel	85	3	80	2	255	170	240	160
Transport	50	5	65	8	250	400	325	520
Miscellaneous	65	1	72	3	65	195	72	216
				Total	1220	1600	1341	1784

Table 9.13

Fisher's price index number

$$P_{01}^{F} = \left(\sqrt{\frac{\sum p_1q_0 \times \sum p_1q_1}{\sum p_0q_0 \times \sum p_0q_1}}\right) \times 100 = \left(\sqrt{\frac{1341 \times 1784}{1220 \times 1600}}\right) \times 100 = 110.706$$

Time Reversal Test: $P_{01} \times P_{10} = 1$

$$P_{01} \times P_{10} = \sqrt{\left(\frac{\sum p_1q_0 \times \sum p_1q_1 \times \sum p_0q_1 \times \sum p_0q_0}{\sum p_0q_0 \times \sum p_0q_1 \times \sum p_1q_1 \times \sum p_1q_0}\right)}$$

$$P_{01} \times P_{10} = \sqrt{\left(\frac{1341 \times 1784 \times 1600 \times 1220}{1220 \times 1600 \times 1784 \times 1341}\right)}$$

$$P_{01} \times P_{10} = 1$$

Factor Reversal Test

$$P_{01} \times Q_{01} = \frac{\sum p_1q_1}{\sum p_0q_0}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{\sum p_1q_0 \times \sum p_1q_1 \times \sum q_1p_0 \times \sum q_1p_1}{\sum p_0q_0 \times \sum p_0q_1 \times \sum q_0p_0 \times \sum q_0p_1}\right)}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{1341 \times 1784 \times 1600 \times 1784}{1220 \times 1600 \times 1220 \times 1341}\right)}$$

$$P_{01} \times Q_{01} = \sqrt{\left(\frac{1784 \times 1784}{1220 \times 1220}\right)} = \frac{1784}{1220} \Rightarrow P_{01} \times Q_{01} = \frac{\sum p_1q_1}{\sum p_0q_0}$$

9.2.4 Construction of Cost of Living Index Number

Cost of Living Index Number is constructed to study the effect of changes in the price of goods and services of consumers for a current period as compared with base period. The change in the cost of living index number between any two periods means the change in income which will be necessary to maintain the same standard of living in both the periods. Therefore the cost of living index number measures the average increase in the cost to maintain the same standard of life. Further, the consumption habits of people differ widely from class to class (rich, poor, middle class) and even with the region. The changes in the price level affect the different classes of people, consequently the general price index numbers fail to reflect the effect of changes in their cost of living of different classes of people. Therefore, cost of living index number measures the general price movement of the commodities consumed by different classes of people.

Uses of Cost of Living Index Number

(i) It indicates whether the real wages of workers are rising or falling for a given time.

(ii) It is used by the administrators for regulating dearness allowance or grant of bonus to the workers.

Methods of constructing Cost of Living Index Number

The cost of living index number can be constructed by the following methods,

(i) Aggregate Expenditure Method (or) Weighted Aggregative Method.

(ii) Family Budget Method.

Aggregate Expenditure Method

This is the most common method used to calculate cost of living index number. In this method, weights are assigned to various commodities consumed by a group in the base year. In this method the quantity of the base year is used as weight.

The formula is given by,

$$\text{Cost of Living Index Number} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

Note

The formula for Aggregate Expenditure Method to calculate Cost of Living Index Number is same as formula of Laspeyre's Method.

Family Budget Method

In this method, the weights are calculated by multiplying prices and quantity of the base year. (i.e.) $V = \sum p_0 q_0$. The formula is given by,

$$\text{Cost of Living Index Number} = \frac{\sum PV}{\sum V}$$

where $P = \frac{p_1}{p_0} \times 100$ is the price relative.

$V = \sum p_0 q_0$ is the value relative.

Note

This method is same as the weighted average of price relative method.

When to Use ?

When the Price and Quantity are given, Aggregate Expenditure Method is used

When the Price and Weight are given, Family Budget Method is used.

Example 9.15

Calculate the cost of living index number for the following data.

Commodities	Quantity	Price	
	2005	2005	2010
A	10	7	9
B	12	6	8
C	17	10	15
D	19	14	16
E	15	12	17

Solution:

Commodities	Quantity	Price		P_1Q_0	P_0Q_0
	2005 (Q_0)	2005 (P_0)	2010 (P_1)		
A	10	7	9	90	70
B	12	6	8	96	72
C	17	10	15	255	170
D	19	14	16	304	266
E	15	12	17	255	180
			Total	1000	758

Table 9.14

$$\text{Cost of Living Index Number} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 = \frac{1000}{758} \times 100 = 131.926$$

Example 9.16

Calculate the cost of living index number for the year 2015 with respect to base year 2010 of the following data.

Commodities	Number of Units (2010)	Price (2010)	Price (2015)
Rice	5	1500	1750
Sugar	3.5	1100	1200
Pulses	3	800	950
Cloth	2	1200	1550
Ghee	0.75	550	700
Rent	12	2500	3000
Fuel	8	750	600
Misc	10	3200	3500

Solution:

Here the base year quantities are given, therefore we can apply Aggregate Expenditure Method.

Commodities	Number of Units q_0 (2010)	Price (2010) p_0	Price (2015) p_1	p_0q_0	p_1q_0
Rice	5	1500	1750	7500	8750
Sugar	3.5	1100	1200	3850	4200
Pulses	3	800	950	2400	2850
Cloth	2	1200	1550	2400	3100
Ghee	0.75	550	700	412.5	525
Rent	12	2500	3000	30000	36000
Fuel	8	750	600	6000	4800
Misc	10	3200	3500	32000	35000
			Total	84562.5	95225

Table 9.15

$$\text{Cost of Living Index Number} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 = \frac{95225}{84562.5} \times 100 = 112.609$$

Hence, the Cost of Living Index Number for a particular class of people for the year 2015 is increased by 12.61 % as compared to the year 2010.

Example 9.17

Calculate the cost of living index number by consumer price index number for the year 2016 with respect to base year 2011 of the following data.

Commodities	Price		Quantity
	Base year	Current year	
Rice	32	48	25
Sugar	25	42	10
Oil	54	85	6
Coffee	250	460	1
Tea	175	275	2

Solution:

Here the base year quantities are given, therefore we can apply Aggregate Expenditure Method.

Commodities	Price		Quantity (q_0)	p_0q_0	p_1q_0
	Base year (p_0)	Current year (p_1)			
Rice	32	48	25	800	1200
Sugar	25	42	10	250	420
Oil	54	85	6	324	510
Coffee	250	460	1	250	460
Tea	175	275	2	350	550
			Total	1974	3140

Table 9.16

$$\text{Cost of Living Index Number} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 = \frac{3140}{1974} \times 100 = 159.0679$$

Hence, the Cost of Living Index Number for a particular class of people for the year 2016 is increased by 59.0679 % as compared to the year 2011.

Example 9.18

Construct the cost of living index number for 2011 on the basis of 2007 from the given data using family budget method.

Commodities	Price		Weights
	2007	2011	
A	350	400	40
B	175	250	35
C	100	115	15
D	75	105	20
E	60	80	25

Solution:

Commodities	Price		Weights	$P=\frac{P_1}{P_0}\times 100$	PV
	2007 (P_0)	2011 (P_1)	(V)		
A	350	400	40	114.285	4571.4
B	175	250	35	142.857	4999.99
C	100	115	15	115	1725
D	75	105	20	114	2280
E	60	80	25	133.333	3333.33
		Total	135		16909.7315

Table 9.17

$$\text{Cost of Living Index Number} = \frac{\sum PV}{\sum V} = \frac{16909.7315}{135} = 125.2572$$

Hence, the Cost of Living Index Number for a particular class of people for the year 2011 is increased by 25.25 % as compared to the year 2007.

Exercise 9.2

1. Define Index Number.
2. State the uses of Index Number.
3. Mention the classification of Index Number.
4. Define Laspeyre's price index number.
5. Explain Paasche's price index number.
6. Write note on Fisher's price index number.
7. State the test of adequacy of index number.
8. Define Time Reversal Test.
9. Explain Factor Reversal Test.
10. Define true value ratio.
11. Discuss about Cost of Living Index Number.
12. Define Family Budget Method.
13. State the uses of Cost of Living Index Number.
14. Calculate by a suitable method, the index number of price from the following data:

Commodity	2002		2012	
	Price	Quantity	Price	Quantity
A	10	20	16	10
B	12	34	18	42
C	15	30	20	26

15. Calculate price index number for 2005 by (a) Laspeyre's (b) Paasche's method

Commodity	1995		2005	
	Price	Quantity	Price	Quantity
A	5	60	15	70
B	4	20	8	35
C	3	15	6	20

16. Compute (i) Laspeyre's (ii) Paasche's (iii) Fisher's Index numbers for the 2010 from the following data.

Commodity	Price		Quantity	
	2000	2010	2000	2010
A	12	14	18	16
B	15	16	20	15
C	14	15	24	20
D	12	12	29	23

17. Using the following data, construct Fisher's Ideal index and show how it satisfies Factor Reversal Test and Time Reversal Test?

Commodity	Price in Rupees per unit		Number of units	
	Base year	Current year	Base year	Current year
A	6	10	50	56
B	2	2	100	120
C	4	6	60	60
D	10	12	50	24
E	8	12	40	36

18. Using Fisher's Ideal Formula, compute price index number for 1999 with 1996 as base year, given the following:

Year	Commodity: A		Commodity: B		Commodity: C	
	Price (Rs.)	Quantity (Kg)	Price (Rs.)	Quantity (Kg)	Price (Rs.)	Quantity (Kg)
1996	5	10	8	6	6	3
1999	4	12	7	7	5	4

19. Calculate Fisher's index number to the following data. Also show that it satisfies Time Reversal Test.

Commodity	2016		2017	
	Price (Rs.)	Quantity (Kg)	Price (Rs.)	Quantity (Kg)
Food	40	12	65	14
Fuel	72	14	78	20
Clothing	36	10	36	15
Wheat	20	6	42	4
Others	46	8	52	6

20. The following are the group index numbers and the group weights of an average working class family's budget. Construct the cost of living index number:

Groups	Food	Fuel and Lighting	Clothing	Rent	Miscellaneous
Index Number	2450	1240	3250	3750	4190
Weight	48	20	12	15	10

21. Construct the cost of living Index number for 2015 on the basis of 2012 from the following data using family budget method.

Commodity	Price		Weights
	2012	2015	
Rice	250	280	10
Wheat	70	85	5
Corn	150	170	6
Oil	25	35	4
Dhal	85	90	3

22. Calculate the cost of living index by aggregate expenditure method:

Commodity	Weights 2010	Price (Rs.)	
		2010	2015
P	80	22	25
Q	30	30	45
R	25	42	50
S	40	25	35
T	50	36	52

9.3 Statistical Quality Control (SQC)

Introduction

It is one of the most important applications of statistical techniques in industry. The term Quality means a level or standard of a product which depends on Material, Manpower, Machines, and Management (4M's). Quality Control ensures the quality specifications all along them from the arrival of raw materials through each of their processing to the final delivery of goods. This technique is used in almost all production industries such as automobile, textile, electrical equipment, biscuits, bath soaps, chemicals, petroleum products etc.

9.3.1 Meaning

Quality Control is a powerful technique used to diagnose the lack of quality in any of the raw materials, processes, machines etc… It is essential that the end products should possess the qualities that the consumer expects from the manufacturer.

9.3.2 Causes of Variation

There are two causes of variation which affects the quality of a product, namely

1. Chance Causes (or) Random causes

2. Assignable Causes

Chance Causes

These are small variations which are natural and inherent in the manufacturing process. The variation occurring due to these causes is beyond the human control and cannot be prevented or eliminated under any circumstances. The minor causes which do not affect the quality of the products to an extent are called as Chance Causes (or) Random causes. For example Rain, floods, power cuts, etc…

Assignable Causes

The second type of variation which is present in any production process is due to non-random causes. The assignable causes may occur in at any stage of the process, right from the arrival of the raw materials to the final delivery of the product. Some of the important factors of assignable causes are defective raw materials, fault in machines, unskilled manpower, worn out tools, new operation, etc.

The main purpose of SQC is to device statistical techniques which would help in elimination of assignable causes and bring the production process under control.

9.3.3 Process Control and Product Control

The main objective in any production process is to control and maintain a satisfactory quality level of the manufactured product. This is done by 'Process Control'. In Process Control the proportion of defective items in the production process is to be minimized and it is achieved through the technique of control charts. Product Control means that controlling the quality of the product by critical examination through sampling inspection plans. Product Control aims at a certain quality level to be guaranteed to the customers. It attempts to ensure that the product sold does not contain a large number of defective items. Thus it is concerned with classification of raw materials, semi-finished goods or finished goods into acceptable or rejectable products.

Control Charts

In an industry, there are two kinds of problems to be faced, namely

(i) To check whether the process is conforming to its standard level.

(ii) To improve the standard level and reduce the variability.

Shewhart's control charts provide an answer to both. It is a simple technique used for detecting patterns of variations in the data. Control charts are simple to construct and easy to interpret. A typical control charts consists of the following three lines.

(i) Centre Line (CL) indicates the desired standard level of the process.

(ii) Upper Control Limit (UCL) indicates the upper limit of tolerance.

(iii) Lower Control Limit (LCL) indicates the lower limit of tolerance.

If the data points fall within the control limits, then we can say that the process is in control, instead if one or more data points fall outside the control limits, then we can say that the process is out of control.

For example, the following diagram shows all the three control lines with the data points plotted, since all the points falls within the control limits, we can say that the process is in control.

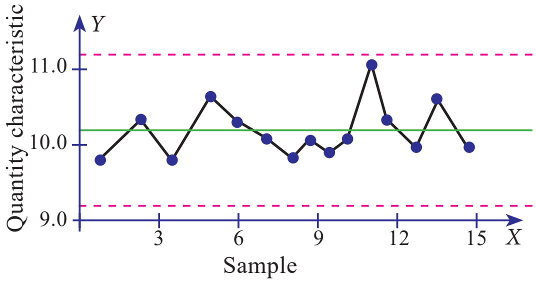


Fig 9.4

Control Charts for Variables

These charts may be applied to any quality characteristic that can be measured quantitatively. A quality characteristic which can be expressed in terms of a numerical value is called as a variable. Many quality characteristics such as dimensions like length, width, temperature, tensile strength etc… of a product are measurable and are expressed in a specific unit of measurements. The variables are of continuous type and are regarded to follow normal probability law. For quality control of such data, there are two types of control charts used. They are as follows :

(i) Charts for Mean ($\overline{X}$) (ii) Charts for Range (R)

9.3.4 Construction of $\overline{X}$ and R charts

Any production process is not perfect enough to produce all the products exactly the same. Some amount of variation is inherent in any production process. This variation is a total of number of characteristics of the production process such as raw materials, machine setting, operators, handling new operations and new machines, etc. The $\overline{X}$ chart is used to show the quality averages of the samples taken from the given process. The *R* chart is used to show the variability or dispersion of the samples taken from the given process. The control limits of the $\overline{X}$ and *R* charts shows the presence or absence of assignable causes in the production process. Both $\overline{X}$ and *R* charts are usually required for decision making to accept or reject the process.

The procedure for constructing $\overline{X}$ and *R* charts are outlined below.

Procedure for $\overline{X}$

(i) Let X_1,X_2, X_3, etc. be the samples selected, each containing '*n*' observations (usually n = 4, 5 or 6)

(ii) Calculate mean for each samples $\overline{X_1}, \overline{X_2}, \overline{X_3}$.... by using

$\overline{X_i} = \dfrac{\sum X_i}{n}, i = 1,2,3,4,....$ where $\sum X_i$ = total of '*n*' values included in the sample X_i.

(iii) Find the mean $\left(\overline{\overline{X}}\right)$ of the sample means.

$\overline{\overline{X}} = \dfrac{\sum \overline{X}}{number\ of\ sample\ means}$ where $\sum \overline{X}$ = total of all the sample means.

Procedure for *R* -Charts.

Calculate $R = x_{\max} - x_{\min}$

Let R_1, R_2, R_3… be the ranges of the 'n' samples. The average range is given by

$$\bar{R} = \frac{\sum R}{n}$$

The calculation of control limits for $\bar{X}$ chart in two different cases is

Case (i) when $\bar{X}$ and SD are given	Case (ii) when $\bar{X}$ and SD are not given
$UCL = \bar{\bar{X}} + 3\frac{\sigma}{\sqrt{n}}$	$UCL = \bar{\bar{X}} + A_2\bar{R}$
$CL = \bar{\bar{X}}$	$CL = \bar{\bar{X}}$
$LCL = \bar{\bar{X}} - 3\frac{\sigma}{\sqrt{n}}$	$LCL = \bar{\bar{X}} - A_2\bar{R}$

Table 9.18

The calculation of control limits for R chart in two different cases are

Case (i) when SD is given	Case (ii) when SD is not given
$UCL = \bar{R} + 3\sigma_R$	$UCL = D_4\bar{R}$
$CL = \bar{R}$	$CL = \bar{R}$
$LCL = \bar{R} - 3\sigma_R$	$LCL = D_3\bar{R}$

Table 9.19

The values of A_2, D_3 and D_4 are given in the table.

Example 9.19

A machine drills hole in a pipe with a mean diameter of 0.532 cm and a standard deviation of 0.002 cm. Calculate the control limits for mean of samples 5.

Solution:

Given $\bar{\bar{X}} = 0.532$, $\sigma = 0.002$, $n = 5$

The control limits for $\bar{X}$ chart is

$$UCL = \bar{\bar{X}} + 3\frac{\sigma}{\sqrt{n}} = 0.532 + 3\frac{0.002}{\sqrt{5}} = 0.5346$$

$$CL = \bar{\bar{X}} = 0.532$$

$$UCL = \overline{\overline{X}} - 3\frac{\sigma}{\sqrt{n}} = 0.532 - 3\frac{0.002}{\sqrt{5}} = 0.5293$$

Example 9.20

The following data gives the readings for 8 samples of size 6 each in the production of a certain product. Find the control limits using mean chart.

Sample	1	2	3	4	5	6
Mean	300	342	351	319	326	333
Range	25	37	20	28	30	22

Given for $n = 6$, $A_2 = 0.483$,

Solution:

Sample	1	2	3	4	5	6	**Total**
Mean	300	342	351	319	326	333	**1971**
Range	25	37	20	28	30	22	**162**

Table 9.20

The control limits for $\overline{X}$ chart is

$$\overline{\overline{X}} = \frac{\sum \overline{X}}{number\ of\ samples} = \frac{1971}{6} = 328.5 \qquad \overline{R} = \frac{\sum R}{n} = \frac{162}{6} = 27$$

$$UCL = \overline{\overline{X}} + A_2\overline{R} = 328.5 + 0.483(27) = 341.54$$

$$CL = \overline{\overline{X}} = 328.5$$

$$LCL = \overline{\overline{X}} - A_2\overline{R} = 328.5 - 0.483(27) = 315.45$$

Example 9.21

The data shows the sample mean and range for 10 samples for size 5 each. Find the control limits for mean chart and range chart.

Sample	1	2	3	4	5	6	7	8	9	10
Mean	21	26	23	18	19	15	14	20	16	10
Range	5	6	9	7	4	6	8	9	4	7

Solution:

Sample	1	2	3	4	5	6	7	8	9	10	**Total**
Mean	21	26	23	18	19	15	14	20	16	10	**182**
Range	5	6	9	7	4	6	8	9	4	7	**65**

Table 9.21

The control limits for $\overline{X}$ chart is

$$\overline{\overline{X}} = \frac{\sum \overline{X}}{number\ of\ samples} = \frac{182}{10} = 18.2 \qquad \overline{R} = \frac{\sum R}{n} = \frac{65}{10} = 6.5$$

$$UCL = \overline{\overline{X}} + A_2 \overline{R} = 18.2 + 0.577(6.5) = 21.95$$

$$CL = \overline{\overline{X}} = 18.2$$

$$LCL = \overline{\overline{X}} - A_2 \overline{R} = 18.2 - 0.577(6.5) = 14.5795$$

The control limits for Range chart is

$$UCL = D_4 \overline{R} = 2.114(6.5) = 13.741$$

$$CL = \overline{R} = 6.5$$

$$LCL = D_3 \overline{R} = 0(6.5) = 0$$

Example 9.22

The following data gives readings of 10 samples of size 6 each in the production of a certain product. Draw control chart for mean and range with its control limits.

Sample	1	2	3	4	5	6	7	8	9	10
Mean	383	508	505	582	557	337	514	614	707	753
Range	95	128	100	91	68	65	148	28	37	80

Solution:

Sample	1	2	3	4	5	6	7	8	9	10	**Total**
Mean	383	508	505	582	557	337	514	614	707	753	**5460**
Range	95	128	100	91	68	65	148	28	37	80	**840**

Table 9.22

$$\overline{\overline{X}} = \frac{\sum \overline{X}}{10} = \frac{5460}{10} = 546 \qquad \overline{R} = \frac{\sum R}{n} = \frac{840}{10} = 84$$

$$UCL = \overline{\overline{X}} + A_2 \overline{R} = 546 + 0.483(84) = 586.57$$

$$CL = \overline{\overline{X}} = 546$$

$$LCL = \overline{\overline{X}} - A_2 \overline{R} = 546 - 0.483(84) = 505.43$$

The control limits for Range chart is

$$UCL = D_4 \overline{R} = 2.004(84) = 168.336$$

$$CL = \overline{R} = 84$$

$$LCL = D_3 \overline{R} = 0(84) = 0$$

Example 9.23

You are given below the values of sample mean ($\overline{X}$) and the range (R) for ten samples of size 5 each. Draw mean chart and comment on the state of control of the process.

Sample number	1	2	3	4	5	6	7	8	9	10
$\overline{X}$	43	49	37	44	45	37	51	46	43	47
R	5	6	5	7	7	4	8	6	4	6

Given the following control chart constraint for : $n = 5$, $A_2 = 0.58$, $D_3 = 0$ and $D_4 = 2.115$

Solution:

$$\overline{\overline{X}} = \frac{\sum \overline{X}}{10} = \frac{442}{10} = 44.2$$

$$\overline{R} = \frac{\sum R}{n} = \frac{58}{10} = 5.8$$

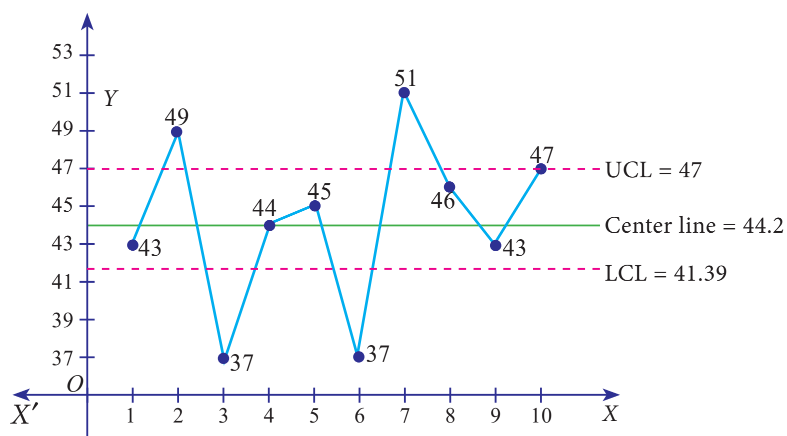


Fig 9.5

$$UCL = \overline{\overline{X}} + A_2 \overline{R}$$

$$= 44.2 + 0.483(5.8) = 47.00$$

$$CL = \overline{\overline{X}} = 44.2$$

$$LCL = \overline{\overline{X}} - A_2 \overline{R} = 44.2 - 0.483(5.8) = 41.39$$

The above diagram shows all the three control lines with the data points plotted, since four points falls out of the control limits, we can say that the process is out of control.

Exercise 9.3

1. Define Statistical Quality Control.
2. Mention the types of causes for variation in a production process.
3. Define Chance Cause.
4. Define Assignable Cause.
5. What do you mean by product control?
6. What do you mean by process control?
7. Define a control chart.
8. Name the control charts for variables.

9. Define mean chart.

10. Define R Chart.

11. What are the uses of statistical quality control?

12. Write the control limits for the mean chart.

13. Write the control limits for the R chart.

14. A machine is set to deliver packets of a given weight. Ten samples of size five each were recorded. Below are given relevant data:

Sample number	1	2	3	4	5	6	7	8	9	10
$\overline{X}$	15	17	15	18	17	14	18	15	17	16
R	7	7	4	9	8	7	12	4	11	5

Calculate the control limits for mean chart and the range chart and then comment on the state of control. (conversion factors for $n=5$, $A_2 = 0.58$, $D_3 = 0$ and $D_4 = 2.115$)

15. Ten samples each of size five are drawn at regular intervals from a manufacturing process. The sample means ($\overline{X}$) and their ranges (R) are given below:

Sample number	1	2	3	4	5	6	7	8	9	10
$\overline{X}$	49	45	48	53	39	47	46	39	51	45
R	7	5	7	9	5	8	8	6	7	6

Calculate the control limits in respect of $\overline{X}$ chart. (Given $A_2 = 0.58$, $D_3 = 0$ and $D_4 = 2.115$) Comment on the state of control.

16. Construct $\overline{X}$ and R charts for the following data:

Sample Number	Observations		
1	32	36	42
2	28	32	40
3	39	52	28
4	50	42	31
5	42	45	34
6	50	29	21
7	44	52	35
8	22	35	44

(Given for $n=3$, $A_2 = 0.58$, $D_3 = 0$ and $D_4 = 2.115$)

17. The following data show the values of sample mean ($\overline{X}$) and its range (R) for the samples of size five each. Calculate the values for control limits for mean , range chart and determine whether the process is in control.

Sample number	1	2	3	4	5	6	7	8	9	10
Mean	11.2	11.8	10.8	11.6	11.0	9.6	10.4	9.6	10.6	10.0
Range	7	4	8	5	7	4	8	4	7	9

(conversion factors for $n=5,\ A_2=0.58$, $D_3=0$ and $D_4=2.115$)

18. A quality control inspector has taken ten samples of size four packets each from a potato chips company.The contents of the sample are given below, Calculate the control limits for mean and range chart.

Sample Number	Observations			
	1	2	3	4
1	12.5	12.3	12.6	12.7
2	12.8	12.4	12.4	12.8
3	12.1	12.6	12.5	12.4
4	12.2	12.6	12.5	12.3
5	12.4	12.5	12.5	12.5
6	12.3	12.4	12.6	12.6
7	12.6	12.7	12.5	12.8
8	12.4	12.3	12.6	12.5
9	12.6	12.5	12.3	12.6
10	12.1	12.7	12.5	12.8

(Given for $n=5,\ A_2=0.58$, $D_3=0$ and $D_4=2.115$)

19. The following data show the values of sample means and the ranges for ten samples of size 4 each. Construct the control chart for mean and range chart and determine whether the process is in control.

Sample number	1	2	3	4	5	6	7	8	9	10
$\overline{X}$	29	26	37	34	14	45	39	20	34	23
R	39	10	39	17	12	20	05	21	23	15

20. In a production process, eight samples of size 4 are collected and their means and ranges are given below. Construct mean chart and range chart with control limits.

Sample number	1	2	3	4	5	6	7	8
$\overline{X}$	12	13	11	12	14	13	16	15
R	2	5	4	2	3	2	4	3

21. In a certain bottling industry the quality control inspector recorded the weight of each of the 5 bottles selected at random during each hour of four hours in the morning.

Time	Weights in ml				
8:00 AM	43	41	42	43	41
9:00 AM	40	39	40	39	44
10:00 AM	42	42	43	38	40
11:00 AM	39	43	40	39	42

Exercise 9.4

Choose the correct Answer

1. A time series is a set of data recorded

   (a) Periodically (b) Weekly

   (c) successive points of time (d) all the above

2. A time series consists of

   (a) Five components (b) Four components

   (c) Three components (d) Two components

3. The components of a time series which is attached to short term fluctuation is

   (a) Secular trend (b) Seasonal variations

   (c) Cyclic variation (d) Irregular variation

4. Factors responsible for seasonal variations are

   (a) Weather (b) Festivals

   (c) Social customs (d) All the above

5. The additive model of the time series with the components T, S, C and I is

   (a) $y=T+S+C\times I$ (b) $y=T+S\times C\times I$ (c) $y=T+S+C+I$ (d) $y=T+S\times C+I$

6. Least square method of fitting a trend is

   (a) Most exact (b) Least exact

   (c) Full of subjectivity (d) Mathematically unsolved

7. The value of 'b' in the trend line y=a+bx is

(a) Always positive (b) Always negative

(c) Either positive or negative (d) Zero

8. The component of a time series attached to long term variation is trended as

(a) Cyclic variation (b) Secular variations

(c) Irregular variation (d) Seasonal variations

9. The seasonal variation means the variations occurring with in

(a) A number of years (b) within a year

(c) within a month (d) within a week

10. Another name of consumer's price index number is:

(a) Whole-sale price index number (b) Cost of living index

(c) Sensitive (d) Composite

11. Cost of living at two different cities can be compared with the help of

(a) Consumer price index (b) Value index

(c) Volume index (d) Un-weighted index

12. Laspeyre's index = 110, Paasche's index = 108, then Fisher's Ideal index is equal to:

(a) 110 (b) 108 (c) 100 (d) 109

13. Most commonly used index number is:

(a) Volume index number (b) Value index number

(c) Price index number (d) Simple index number

14. Consumer price index are obtained by:

(a) Paasche's formula (b) Fisher's ideal formula

(c) Marshall Edgeworth formula (d) Family budget method formula

15. Which of the following Index number satisfy the time reversal test?

(a)Laspeyre's Index number (b) Paasche's Index number

(c) Fisher Index number (d) All of them.

16. While computing a weighted index, the current period quantities are used in the:

(a) Laspeyre's method (b) Paasche's method

(c) Marshall Edgeworth method (d) Fisher's ideal method

17. The quantities that can be numerically measured can be plotted on a

(a) p - chart (b) c – chart (c) x bar chart (d) np – chart

18. How many causes of variation will affect the quality of a product?

(a) 4 (b) 3 (c) 2 (d) 1

19. Variations due to natural disorder is known as

(a) random cause (b) non-random cause

(c) human cause (d) all of them

20. The assignable causes can occur due to

(a) poor raw materials (b) unskilled labour

(c) faulty machines (d) all of them

21. A typical control charts consists of

(a) CL, UCL (b) CL, LCL (c) CL, LCL, UCL (d) UCL, LCL

22. $\overline{X}$ chart is a

(a) attribute control chart (b) variable control chart

(c) neither Attribute nor variable control chart

(d) both Attribute and variable control chart

23. R is calculated using

(a) $x_{max} - x_{min}$ (b) $x_{min} - x_{max}$ (c) $\overline{x}_{max} - \overline{x}_{min}$ (d) $\overline{\overline{x}}_{max} - \overline{\overline{x}}_{min}$

24. The upper control limit for $\overline{X}$ chart is given by

(a) $\overline{X} + A_2\overline{R}$ (b) $\overline{\overline{X}} + A_2 R$ (c) $\overline{\overline{X}} + A_2\overline{R}$ (d) $\overline{\overline{X}} + A_2\overline{\overline{R}}$

25. The LCL for R chart is given by

(a) $D_2\overline{R}$ (b) $D_2\overline{\overline{R}}$ (c) $D_3\overline{\overline{R}}$ (d) $D_3\overline{R}$

Miscellaneous Problems

1. Using three yearly moving averages, Determine the trend values from the following data.

Year	Profit	Year	Profit
2001	142	2007	241
2002	148	2008	263
2003	154	2009	280
2004	146	2010	302
2005	157	2011	326
2006	202	2012	353

2. From the following data, calculate the trend values using fourly moving averages.

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998
Sales	506	620	1036	673	588	696	1116	738	663

3. Fit a straight line trend by the method of least squares to the following data.

Year	1980	1981	1982	1983	1984	1985	1986	1987
Sales	50.3	52.7	49.3	57.3	56.8	60.7	62.1	58.7

4. Calculate the Laspeyre's, Paasche's and Fisher's price index number for the following data. Interpret on the data.

Commodities	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	170	562	72	632
B	192	535	70	756
C	195	639	95	926
D	187	128	92	255
E	185	542	92	632
F	150	217	180	314
7	12.6	12.7	12.5	12.8
8	12.4	12.3	12.6	12.5
9	12.6	12.5	12.3	12.6
10	12.1	12.7	12.5	12.8

5. Using the following data, construct Fisher's Ideal Index Number and Show that it satisfies Factor Reversal Test and Time Reversal Test?

Commodities	Price		Quantity	
	Base Year	Current year	Base Year	Current year
Wheat	6	10	50	56
Ghee	2	2	100	120
Firewood	4	6	60	60
Sugar	10	12	30	24
Cloth	8	12	40	36

6. Compute the consumer price index for 2015 on the basis of 2014 from the following data.

Commodities	Quantities	Prices in 2015	Prices in 2016
A	6	5.75	6.00
B	6	5.00	8.00
C	1	6.00	9.00
D	6	8.00	10.00
E	4	2.00	1.50
F	1	20.00	15.00

7. An Enquiry was made into the budgets of the middle class families in a city gave the following information.

Expenditure	Food	Rent	Clothing	Fuel	Rice
Price(2010)	150	50	100	20	60
Price(2011)	174	60	125	25	90
Weights	35	15	20	10	20

What changes in the cost of living have taken place in the middle class families of a city?

8. From the following data, calculate the control limits for the mean and range chart.

Sample No.	1	2	3	4	5	6	7	8	9	10
Sample Observations	50	51	50	48	46	55	45	50	47	56
	55	50	53	53	50	51	48	56	53	53
	52	53	48	50	44	56	53	54	49	55
	49	50	52	51	48	47	48	53	52	54
	54	46	47	53	47	51	51	57	54	52

10. The following data gives the average life(in hours) and range of 12 samples of 5lamps each. The data are

Sample No	1	2	3	4	5	6
Sample Mean	1080	1390	1460	1380	1230	1370
Sample Range	410	670	180	320	690	450
Sample No	7	8	9	10	11	12
Sample Mean	1310	1630	1580	1510	1270	1200
Sample Range	380	350	270	660	440	310

Construct control charts for mean and range. Comment on the control limits.

11. The following are the sample means and ranges for 10 samples, each of size 5. Calculate the control limits for the mean chart and range chart and state whether the process is in control or not.

Sample number	1	2	3	4	5	6	7	8	9	10
Mean	5.10	4.98	5.02	4.96	4.96	5.04	4.94	4.92	4.92	4.98
Range	0.3	0.4	0.2	0.4	0.1	0.1	0.8	0.5	0.3	0.5

Table of Control Chart Constants

Sample Size	A2	D3	D4
2	1.880	0	3.267
3	1.023	0	2.574
4	0.729	0	2.282
5	0.577	0	2.114
6	0.483	0	2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
9	0.337	0.184	1.816
10	0.308	0.223	1.777
11	0.285	0.256	1.744
12	0.266	0.283	1.717
13	0.249	0.307	1.693
14	0.235	0.328	1.672
15	0.223	0.347	1.653
16	0.212	0.363	1.637
17	0.203	0.378	1.622
18	0.194	0.391	1.608
19	0.187	0.403	1.597
20	0.180	0.415	1.585
21	0.173	0.425	1.575
22	0.167	0.434	1.566
23	0.162	0.443	1.557
24	0.157	0.451	1.548
25	0.153	0.459	1.541

Table 9.23

Summary

In this chapter we have acquired the knowledge of

- **Method of Moving Averages**

Three year moving averages: $\frac{a+b+c}{3}, \frac{b+c+d}{3}, \frac{c+d+e}{3}, \frac{d+e+f}{3}, \ldots\ldots$

Four year moving averages: $\frac{a+b+c+d}{4}, \frac{b+c+d+e}{4}, \frac{c+d+e+f}{4}, \frac{d+e+f+g}{4}, \ldots$

- **Method of Least Squares**

The straight line equation, Y = a + bX

Two Normal Equations, ΣY = n a + b ΣX ; ΣXY = a ΣX + b ΣX2

- **Methods of measuring Seasonal Variations-Method of Simple Averages:**

$$\text{Seasonal Index (S.I)} = \frac{Seasonal\ Average}{Grand\ average} X100$$

- **If the data is given in months**

$$\text{S.I for Jan} = \frac{Monthly\ Average\ (for\ Jan)}{Grand\ average} X100$$

- **If the data is given in quarter**

$$\text{S.I for Kth Quarter} = \frac{Average\ of\ \mathrm{K}^{\mathrm{th}}\ quarter}{Grand\ average} X100$$

- Continuous distribution function

If X is a continuous random variable with the probability density function $f_X(x)$, then the function $F_X(x)$ is defined by

- Weighted Index Number

$$\text{Price Index (P01)} = \frac{\sum p_1 w}{\sum p_0 w} x100$$

Laspeyre's price index number $P_{01}^{L} = \frac{\sum p_1 q_0}{\sum p_0 q_0} X100$

Paasche's price index number $P_{01}^{P} = \frac{\sum p_1 q_1}{\sum p_0 q_1} X100$

Fisher's price index number $P_{01}^{F} = \sqrt{P_{01}^{L}\ x\ P_{01}^{P}} = \sqrt{\frac{\sum p_1 q_0\ X \sum p_1 q_1}{\sum p_0 q_0\ X \sum p_0 q_1}} X100$

- Time Reversal Test : $P_{01} \times P_{10} = 1.$ Factor Reversal Test: $P_{01}\ X\ Q_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$

- Cost of Living Index Number

  Aggregate Expenditure Method = $\dfrac{\sum p_1 q_0}{\sum p_0 q_0} X 100$.

  Family Budget Method = $\dfrac{\sum PV}{\sum V}$

- Causes of Variation

  1. Chance Causes (or) Random causes 2. Assignable Causes

- Control Charts

  (i) Centre Line (CL) indicates the desired standard level of the process.

  (ii) Upper Control Limit (UCL) indicates the upper limit of tolerance.

  (iii) Lower Control Limit (LCL) indicates the lower limit of tolerance.

  (i) Centre Line (CL) indicates the desired standard level of the process.

  (ii) Upper Control Limit (UCL) indicates the upper limit of tolerance.

  (iii) Lower Control Limit (LCL) indicates the lower limit of tolerance.

The control limits for $\overline{X}$ chart in two different cases are

case (i) when $\overline{X}$ and SD are given	case (i) when $\overline{X}$ and SD are not given
$UCL = \overline{\overline{X}} + 3\dfrac{\sigma}{\sqrt{n}}$	$UCL = \overline{\overline{X}} + A_2\overline{R}$
$CL = \overline{\overline{X}}$	$CL = \overline{\overline{X}}$
$LCL = \overline{\overline{X}} - 3\dfrac{\sigma}{\sqrt{n}}$	$LCL = \overline{\overline{X}} - A_2\overline{R}$

The control limits for R chart in two different cases are

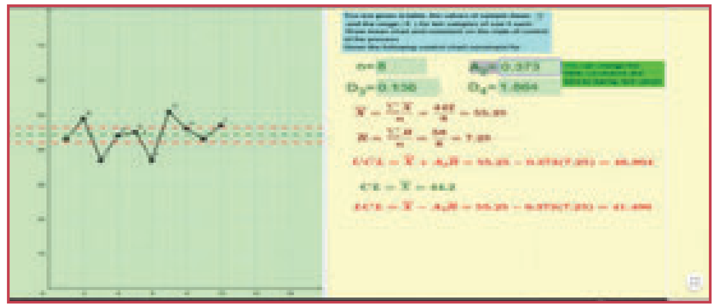
case (i) when SD are given	case (i) when SD are not given
$UCL = \overline{R} + 3\sigma_R$	$UCL = D_4\overline{R}$
$CL = \overline{R}$	$CL = \overline{R}$
$LCL = \overline{R} - 3\sigma_R$	$LCL = D_3\overline{R}$

GLOSSARY

Control charts	கட்டுப்பாட்டு விளக்கப்படங்கள்
Control limit	கட்டுப்பாட்டு எல்லை
Cyclical variation	சுழல் மாறுபாடு
Factor reversal test	காரணி மாற்று சோதனை
Family budget	குடும்ப வரவு– செலவுத் திட்டம்
Fisher's index	பிஷரின் குறியீடு
Index number	குறியீட்டெண்கள்
Irregular variation	ஒழுங்கற்ற மாறுபாடு
Laspeyre's index	லாஸ்பியரின் குறியீடு
Least square	மீச்சிறு வர்க்கம்
Lower control limit	கீழ்க் கட்டுப்பாட்டு எல்லை
Mean charts	சராசரி வரைவுகள்
Moving average	நகரும் சராசரி
Observation	கண்டறிபதிவு / கூர்நோக்கு
Paasche's index	பாசியின் குறியீடு
Process control	செயல்பாட்டு கட்டுப்பாடு (அல்லது) செயலாக்கக் கட்டுபாடு
Product control	உற்பத்தி கட்டுப்பாடு
Range charts	வீச்சு வரைகள்
Seasonal variation	பருவகால மாறுபாடு
Seasonal Index	பருவகால குறியீடு
Secular trend	நீள் காலப்போக்கு
Semi-vverage	பகுதி சராசரி
Statistical quality control	புள்ளியியல் தரக்கட்டுப்பாடு
Time reversal test	காலமாற்று சோதனை
Time series	காலம்சார் தொடர்வரிசை
Trend	போக்கு
Unweighted Index	நிறையிடா குறியீடு
Upper control limit	மேல் கட்டுப்பாட்டு எல்லை
Weighted index	நிறையிட்ட குறியீடு

ICT Corner

Expected Result is shown in this picture

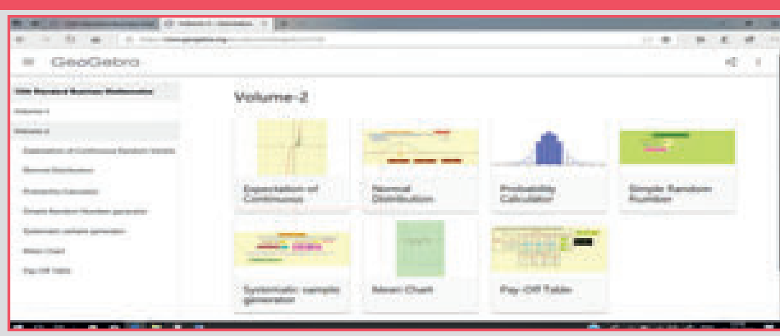


Step – 1 : Open the Browser, type the URL Link given below (or) Scan the QR Code. GeoGebra work Book named "12th Standard Business Mathematics " will open. In the work book there are two Volumes. Select "Volume-2"..

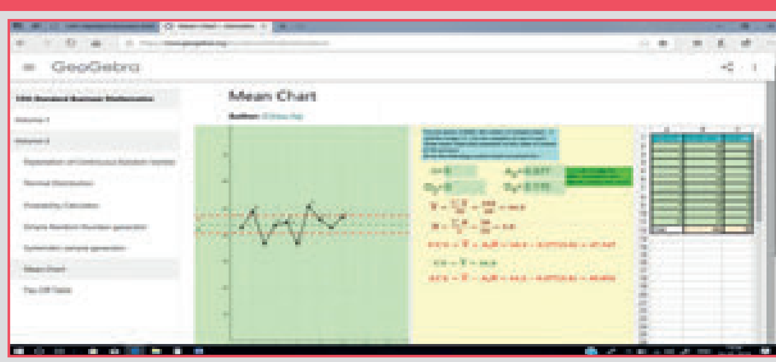
Step - 2 : Select the worksheet named "Mean Chart."

You can enter your own data in Spread Sheet given. Also enter n, A2, D3 and D4 table values in the respective boxes. You can see UCL, CL and LCL values on the right hand side and the corresponding graph on Left hand side

Step 1



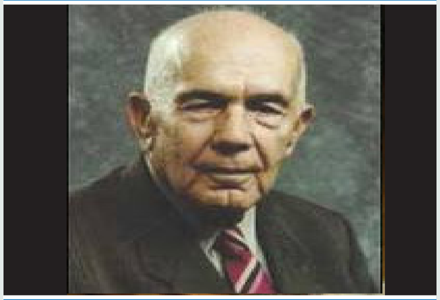
Step 2



Scan the QR Code.

B284_12_BUS_MAT_EM

10 Operations Research



Frank Lauren Hitchcock
(1875-1957)

Introduction

Operations reserach (O.R.) is an analytical method of problem solving and decision-making, that is useful in management organisations. The transportation problem involves certain origins (sources) which may represent factories where we produce homogeneous items and a number of destinations where we supply a required quantity of the products. Each factory has a certain capacity constraint and each destination (dealer or customer) has a certain requirement. The unit cost of transportation of the items from the factory to the dealer/ customer is known. American mathematician and physicist Frank Lauren Hitchcock (1875-1957) known for his formulation of transportation problem in 1941.

BKMB31

Learning Objectives

After studying this chapter students are able to understand

- formulate the transportation and assignment problems
- distinguish between transportation and assignment problems
- find an initial basic feasible solution of a transportation problem
- identify the degeneracy and non-degeneracy in a transportation problem
- find the solution of an assignment problem by Hungarian method.
- distinguish between tactic and strategic decisions
- find the best alternatives using maximin and minimax criteria

10.1 Transportation Problem

The objective of transportation problem is to determine the amount to be transported from each origin to each destinations such that the total transportation cost is minimized.

10.1.1 Definition and formulation

The Structure of the Problem

Let there be m origins and n destinations. Let the amount of supply at the *i th* origin is a_i. Let the demand at *j th* destination is b_j.

The cost of transporting one unit of an item from origin *i* to destination j is c_{ij} and is known for all combinations *(i,j)*. Quantity transported from origin i to destination *j* be x_{ij}

The objective is to determine the quantity x_{ij} to be transported over all routes (*i,j*) so as to minimize the total transportation cost. The supply limits at the origins and the demand requirements at the destinations must be satisfied.

The above transportation problem can be written in the following tabular form:

		Destinations					
		1	2	3	…	n	supply
Origins	1	(x_{11}) C_{11}	(x_{12}) C_{12}	(x_{13}) C_{13}	…	(x_{1n}) C_{1n}	a_1
	2	(x_{21}) C_{21}	(x_{22}) C_{22}	(x_{23}) C_{23}	…	(x_{2n}) C_{2n}	a_2
	⋮		⋮	⋮	⋮	⋮	⋮
	m	(x_{m1}) C_{m1}	(x_{m2}) C_{m2}	(x_{m3}) C_{m3}	…	(x_{mn}) C_{mn}	a_m
	demand	b_1	b_2	b_3	…	b_n	

Table-10.1

Now the linear programming model representing the transportation problem is given by

The objective function is Minimize Z = $\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$ Subject to the constraints

$\sum_{j=1}^{n} x_{ij} = a_i$, i=1,2,…….m (supply constraints)

$\sum_{i=1}^{m} x_{ij} = b_j$, j=1,2,…….$n$ (demand constraints)

$x_{ij} \geq 0$ for all *i,j*. (non-negative restrictions)

Some Definitions

Feasible Solution: A feasible solution to a transportation problem is a set of non-negative values $x_{ij}(i=1,2,..,m, j=1,2,\ldots n)$ that satisfies the constraints.

Basic Feasible Solution: A feasible solution is called a basic feasible solution if it contains not more than $m+n-1$ allocations, where m is the number of rows and n is the number of columns in a transportation problem.

Optimal Solution: Optimal Solution is a feasible solution (not necessarily basic) which optimizes(minimize) the total transportation cost.

Non degenerate basic feasible Solution: If a basic feasible solution to a transportation problem contains exactly $m+n-1$ allocations in independent positions, it is called a Non degenerate basic feasible solution. Here m is the number of rows and n is the number of columns in a transportation problem.

Degeneracy : If a basic feasible solution to a transportation problem contains less than $m+n-1$ allocations , it is called a degenerate basic feasible solution.Here m is the number of rows and n is the number of columns in a transportation problem.

10.1.2 Methods of finding initial Basic Feasible Solutions

There are several methods available to obtain an initial basic feasible solution of a transportation problem. We discuss here only the following three. For finding the initial basic feasible solution total supply must be equal to total demand.

$$\text{(i.e)} \quad \sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

Method:1North-West Corner Rule (NWC)

It is a simple method to obtain an initial basic feasible solution. Various steps involved in this method are summarized below.

Step 1: Choose the cell in the north-west corner of the transportation Table10.1 and allocate as much as possible in this cell so that either the capacity of first row (supply)is exhausted or the destination requirement of the first column(demand) is exhausted. (i.e) $x_{11} = \min(a_1,b_1)$

Step 2: If the demand is exhausted $(b_1 < a_1)$, move one cell right horizontally tothe second column and allocate as much as possible.(i.e) $x_{12}= \min\,(a_1 - x_{11}, b_2)$

If the supply is exhausted $(b_1 > a_1)$, move one cell down vertically to the second row and allocateas much as possible.(i.e)$x_{21} = \min\,(a_2,b_1 - x_{11})$

If both supply and demand are exhausted move one cell diagonally and allocate as much as possible.

Step 3: Continue the above procedure until all the allocations are made

Example 10.1

Obtain the initial solution for the following problem

		Destination			
		A	B	C	Supply
Sources	1	2	7	4	5
	2	3	3	1	8
	3	5	4	7	7
	4	1	6	2	14
Demand		7	9	18	

Solution:

Here total supply = 5+8+7+14=34, Total demand = 7+9+18=34

(i.e) Total supply =Total demand$\therefore$The given problem is balanced transportation problem.

$\therefore$ we can findan initial basic feasible solution to the given problem.

From the above table we can choose the cell in the North West Corner. Here the cell is $(1,A)$

Allocate as much as possible in this cell so that either the capacity of first row is exhausted or the destination requirement of the first column is exhausted.

i.e. x_{11} = min (5,7) = 5

	A	B	C	Supply (a_i)
1	(5) 2	7	4	5/0
2	3	3	1	8
3	5	4	7	7
4	1	6	2	14
Demand (b_j)	7/2	9	18	

Reduced transportation table is

	A	B	C	a_i
2	3	3	1	8
3	5	4	7	7
4	1	6	2	14
b_j	2	9	18	

Now the cell in the North west corner is (2, A)

Allocate as much as possible in the first cell so that either the capacity of second row is exhausted or the destination requirement of the first column is exhausted.

i.e. x_{12} = min (2,8) = 2

	A	B	C	a_i
2	(2) 3	3	1	8/6
3	5	4	7	7
4	1	6	2	14
b_j	2/0	9	18	

Reduced transportation table is

	B	C	a_i
2	3	1	6
3	4	7	7
4	6	2	14
b_j	9	18	

Here north west corner cell is (2,B) Allocate as much as possible in the first cell so that either the capacity of second row is exhausted or the destination requirement of the second column is exhausted.

i.e. x_{22} = min (6,9) = 6

	B	C	a_i
2	(6) 3	1	6/0
3	4	7	7
4	6	2	14
b_j	9/3	18	

Reduced transportation table is

	B	C	a_i
3	4	7	7
4	6	2	14
b_j	3	18	

Here north west corner cell is (3,B).

Allocate as much as possible in the first cell so that either the capacity of third row is exhausted or the destination requirement of the second column is exhausted.

i.e. x_{32} = min (7,3) = 3

	B	C	a_i
3	(3) 4	7	7/4
4	6	2	14
b_j	3/0	18	

Reduced transportation table is

	C	a_i
3	7	7/4
4	2	14
b_j	18	

Here north west corner cell is (3,C) Allocate as much as possible in the first cell so that either the capacity of third row is exhausted or the destination requirement of the third column is exhausted.

i.e. x_{33} = min (4,18) = 4

	C	a_i
3	(4) 7	4/0
4	2	14
b_j	18/14	

Reduced transportation table and final allocation is x_{44} = 14

	C	a_i
4	(14) 2	14/0
b_j	14/0	

Thus we have the following allocations

	A	*B*	*C*	a_i
1	(5) 2	7	4	5
2	(2) 3	(6) 3	1	8
3	5	(3) 4	(4) 7	7
4	1	6	(14) 2	14
b_j	7	9	18	

Transportation schedule : $1 \rightarrow A,\ 2 \rightarrow A,\ 2 \rightarrow B,\ 3 \rightarrow B,\ 3 \rightarrow C,\ 4 \rightarrow C$

The total transportation cost.

$= (5\times2) + (2\times3) + (6\times3) + (3\times4) + (4\times7) + (14\times2)$

= Rs.102

Example 10.2

Determine an initial basic feasible solution to the following transportation problem using North West corner rule.

	D_1	D_2	D_3	D_4	Availability
O_1	6	4	1	5	14
O_2	8	9	2	7	16
O_3	4	3	6	2	5
Requirement	6	10	15	4	35

Here O_i and D_j represent i^{th} origin and j^{th} destination.

Solution :

Given transportation table is

	D_1	D_2	D_3	D_4	Availability (a_i)
O_1	6	4	1	5	14
O_2	8	9	2	7	16
O_3	4	3	6	2	5
Requirement (b_j)	6	10	15	4	35

Total Availability = Total Requirement $\therefore$ The given problem is balanced transportation problem.

Hence there exists a feasible solution to the given problem.

First allocation:

	D_1	D_2	D_3	D_4	a_i
O_1	(6) 6	4	1	5	14/8
O_2	8	9	2	7	16
O_3	4	3	6	2	5
b_j	6/0	10	15	4	35

Second allocation:

	D_1	D_2	D_3	D_4	a_i
O_1	(6) 6	(8) 4	1	5	14/8/0
O_2	8	9	2	7	16
O_3	4	3	6	2	5
b_j	6/0	10/2	15	4	35

Third Allocation:

	D_1	D_2	D_3	D_4	a_i
O_1	(6) 6	(8) 4	1	5	14/8/0
O_2	8	9	2	7	16/14
O_3	4	3	6	2	5
b_j	6/0	10/2/0	15	4	35

Fourth Allocation:

	D_1	D_2	D_3	D_4	a_i
O_1	(6) 6	(8) 4	1	5	14/8/0
O_2	8	(2) 9	(14) 2	7	16/14/0
O_3	4	3	6	2	5
b_j	6/0	10/2/0	15/1	4	35

Fifth allocation:

	D_1	D_2	D_3	D_4	a_i
O_1	(6) 6	(8) 4	1	5	14/8/0
O_2	8	(2) 9	(14) 2	7	16/14/0
O_3	4	3	(1) 6	2	5/4
b_j	6/0	10/2/0	15/1/0	4	35

Final allocation:

	D_1	D_2	D_3	D_4	a_i
O_1	(6) 6	(8) 4	1	5	14/8/0
O_2	8	(2) 9	(14) 2	7	16/14/0
O_3	4	3	(1) 6	(4) 2	5/4/0
b_j	6/0	10/2/0	15/1/0	4/0	35

Transportation schedule : $O_1 \rightarrow D_1$, $O_1 \rightarrow D_2$, $O_2 \rightarrow D_2$, $O_2 \rightarrow D_3$, $O_3 \rightarrow D_3$, $O_3 \rightarrow D_3$.

The transportation cost

$$= (6\times6)+(8\times4)+(2\times9)+(14\times2)+(1\times6)+(4\times2) = \text{Rs.}128$$

Method:2 Least Cost Method (LCM)

The least cost method is more economical than north-west corner rule,since it starts with a lower beginning cost. Various steps involved in this method are summarized as under.

Step 1: Find the cell with the least(minimum) cost in the transportation table.

Step 2: Allocate the maximum feasible quantity to this cell.

Step:3 Eliminate the row or column where an allocation is made.

Step:4 Repeat the above steps for the reduced transportation table until all the allocations are made.

Note

If the minimum cost is not unique then the choice can be made arbitrarily.

Example 10.3

Obtain an initial basic feasible solution to the following transportation problem using least cost method.

	D_1	D_2	D_3	D_4	Supply
O_1	1	2	3	4	6
O_2	4	3	2	5	8
O_3	5	2	2	1	10
Demand	4	6	8	6	

Here O_i and D_j denote i^{th} origin and j^{th} destination respectively.

Solution:

Total Supply = Total Demand = 24

$\therefore$ The given problem is a balanced transportation problem.

Hence there exists a feasible solution to the given problem.

Given Transportation Problem is:

	D_1	D_2	D_3	D_4	Supply (a_i)
O_1	1	2	3	4	6
O_2	4	3	2	5	8
O_3	5	2	2	1	10
Demand (b_j)	4	6	8	6	

The least cost is 1 corresponds to the cells (O_1, D_1) and (O_3, D_4)

Take the Cell (O_1, D_1) arbitrarily.

Allocatemin (6,4) = 4 units to this cell.

	D_1	D_2	D_3	D_4	a_i
O_1	(4) 1	2	3	4	6/2
O_2	4	3	2	5	8
O_3	5	2	2	1	10
b_j	4/0	6	8	6	

The reduced table is

	D_2	D_3	D_4	a_i
O_1	2	3	4	2
O_2	3	2	5	8
O_3	2	2	1	10
b_j	6	8	6	

The least cost corresponds to the cell (O_3, D_4). Allocate min (10,6) = 6 units to this cell.

	D_2	D_3	D_4	a_i
O_1	2	3	4	2
O_2	3	2	5	8
O_3	2	2	(6) 1	10/4
b_j	6	8	6/0	

The reduced table is

	D_2	D_3	a_i
O_1	2	3	2
O_2	3	2	8
O_3	2	2	4
b_j	6	8	

The least costis 2 corresponds to the cells (O_1, D_2), (O_2, D_3), (O_3, D_2), (O_3, D_3)

Allocate min (2,6) = 2 units to this cell.

	D_2	D_3	a_i
O_1	(2) 2	3	2/0
O_2	3	2	8
O_3	2	2	4
b_j	6/4	8	

The reduced table is

	D_2	D_3	a_i
O_2	3	2	8
O_3	2	2	4
bj	4	8	

The least cost is 2 corresponds to the cells (O_2, D_3), (O_3, D_2), (O_3, D_3)

Allocate min (8,8) = 8 units to this cell.

	D_2	D_3	a_i
O_2	3	(8) 2	8/0
O_3	2	2	4
b_j	4	8/0	

The reduced table is

	D_2	a_i
O_3	2	4
b_j	4	

Here allocate 4 units in the cell (O_3, D_2)

	D_2	a_i
O_3	(4) 2	4/0
b_j	4/0	

Thus we have the following allocations:

	D_1	D_2	D_3	D_4	a_i
O_1	(4) 1	(2) 2	3	4	6/2/0
O_2	4	3	(8) 2	5	8/0
O_3	5	(4) 2	2	(6) 1	10/4/0
b_j	4/0	6/4/0	8/0	6/0	

Transportation schedule :

$$O_1 \rightarrow D_1,\ O_1 \rightarrow D_2,\ O_2 \rightarrow D_3,\ O_3 \rightarrow D_2,\ O_3 \rightarrow D_4$$

Total transportation cost

$= (4\times1)+ (2\times2)+(8\times2)+(4\times2)+(6\times1)$

$= 4+4+16+8+6$

=Rs. 38.

Example 10.4

Determine how much quantity should be stepped from factory to various destinations for the following transportation problem using the least cost method

		Destination				
		C	H	K	P	Capacity
Factory	T	6	8	8	5	30
	B	5	11	9	7	40
	M	8	9	7	13	50
	Demand	35	28	32	25	

Cost are expressed in terms of rupees per unit shipped.

Solution:

Total Capacity = Total Demand

$\therefore$ The given problem is balanced transportation problem.

Hence there exists a feasible solution to the given problem.

Given Transportation Problem is

		Destination				
		C	H	K	P	Capacity (a_i)
Factory	T	6	8	8	5	30
	B	5	11	9	7	40
	M	8	9	7	13	50
	Demand (b_j)	35	28	32	25	

First Allocation:

	C	H	K	P	a_i
T	6	8	8	(25) 5	30/5
B	5	11	9	7	40
M	8	9	7	13	50
b_j	35	28	32	25/0	

Second Allocation:

	C	H	K	P	a_i
T	6	8	8	(25) 5	30/5
B	(35) 5	11	9	7	40/5
M	8	9	7	13	50
b_j	35/0	28	32	25/0	

Third Allocation:

	C	H	K	P	a_i
T	6	8	8	(25) 5	30/5
B	(35) 5	11	9	7	40/5
M	8	9	(32) 7	13	50/18
b_j	35/0	28	32/0	25/0	

Fourth Allocation:

	C	H	K	P	a_i
T	6	(5) 8	8	(25) 5	30/5/0
B	(35) 5	11	9	7	40/5
M	8	9	(32) 7	13	50/18
b_j	35/0	28/23	32/0	25/0	

Fifth Allocation:

	C	H	K	P	a_i
T	6	(5) 8	8	(25) 5	30/5/0
B	(35) 5	11	9	7	40/5
M	8	(18) 9	(32) 7	13	50/18/0
b_j	35/0	28/23/5	32/0	25/0	

Sixth Allocation:

	C	H	K	P	a_i
T	6	(5) 8	8	(25) 5	30/5/0
B	(35) 5	(5) 11	9	7	40/5/0
M	8	(18) 9	(32) 7	13	50/18/0
b_j	35/0	28/23/5/0	32/0	25/0	

Transportation schedule :

$T \rightarrow H$, $T \rightarrow P$, $B \rightarrow C$, $B \rightarrow H$, $M \rightarrow H$, $M \rightarrow K$

The total Transportation cost = (5×8) + (25×5)+ (35×5) + (5×11)+ (18×9) + (32×7)

= 40+125+175+55+162+224

= ₹ 781

Method:3 Vogel's Approximation Method(VAM)

Vogel's approximation method yields an initial basic feasible solution which is very close to the optimum solution.Various steps involved in this method are summarized as under

Step 1: Calculate the penalties for each row and each column. Here penalty means the difference between the two successive least cost in a row and in a column .

Step 2: Select the row or column with the largest penalty.

Step 3: In the selected row or column, allocate the maximum feasible quantity to the cell with the minimum cost.

Step 4: Eliminate the row or column where all the allocations are made.

Step 5: Write the reduced transportation table and repeat the steps 1 to 4.

Step 6: Repeat the procedure until all the allocations are made.

DO YOU KNOW?

Optimum solution for any transportation problem can be obtained by using MODI method.

Example 10.5

Find the initial basic feasible solution for the following transportation problem by VAM

		Distribution Centers				Availability
		D_1	D_2	D_3	D_4	---
origin	S_1	11	13	17	14	250
	S_2	16	18	14	10	300
	S_3	21	24	13	10	400
	Requirement	200	225	275	250	

Solution:

Here $\sum a_i = \sum b_j = 950$

(i.e) Total Availability =Total Requirement

$\therefore$ The given problem is balanced transportation problem.

Hence there exists a feasible solution to the given problem.

First let us find the difference (penalty) between the first two smallest costs in each row and column and write them in brackets against the respective rows and columns

	D_1	D_2	D_3	D_4	a_i	Penalty
S_1	(200) 11	13	17	14	250/50	2
S_2	16	18	14	10	300	4
S_3	21	24	13	10	400	3
b_j	200/0	225	275	250		
Penalty	5	5	1	4		

Choose the largest difference. Here the difference is 5 which corresponds to column D_1 and D_2. Choose either D_1 or D_2 arbitrarily. Here we take the column D_1. In this column choose the least cost. Here the least cost corresponds to (S_1, D_1). Allocate min (250, 200) = 200units to this Cell.

The reduced transportation table is

	D_2	D_3	D_4	a_i	Penalty
S_1	(50) 13	17	14	50/0	1
S_2	18	14	10	300	4
S_3	24	13	10	400	3
b_j	225/175	275	250		
Penalty	5	1	4		

Choose the largest difference. Here the difference is 5 whichcorresponds to column D_2. In this column choose the least cost. Here the least cost corresponds to (S_1, D_2). Allocate $\min(50,175) = 50$ units to this Cell.

The reduced transportation table is

	D_2	D_3	D_4	a_i	Penalty
S_2	(175) 18	14	10	300/125	4
S_3	24	13	10	400	3
b_j	175/0	275	250		
Penalty	6	1	0		

Choose the largest difference. Here the difference is 6 which corresponds to column D_2. In this column choose the least cost. Here the least cost corresponds to (S_2, D_2). Allocate $\min(300,175) = 175$ units to this cell.

The reduced transportation table is

	D_3	D_4	a_i	Penalty
S_2	14	(125) 10	125/0	4
S_3	13	10	400	3
b_j	275	250/125		
Penalty	1	0		

Choose the largest difference. Here the difference is 4 corresponds to rowS_2. In this row choose the least cost. Here the least cost corresponds to (S_2, D_4). Allocate $\min(125,250) = 125$ units to this Cell.

The reduced transportation table is

	D_3	D_4	a_i	Penalty
S_3	13	10	400	3
b_j	275	125		
Penalty	-	-		

The Allocation is

	D_3	D_4	a_i
S_3	(275) 13	(125) 10	400/275/0
b_j	275	125/0	

Thus we have the following allocations:

	D_1	D_2	D_3	D_4	a_i
S_1	(200) 11	(50) 13	17	14	250
S_2	16	(175) 18	14	(125) 10	300
S_3	21	24	(275) 13	(125) 10	400
b_j	200	225	275	250	

Transportation schedule :

$$S_1 \rightarrow D_1,\ S_1 \rightarrow D_2,\ S_2 \rightarrow D_2,\ S_2 \rightarrow D_4,\ S_3 \rightarrow D_3,\ S_3 \rightarrow D_4$$

This initial transportation cost

$$= (200 \times 11) + (50 \times 13) + (175 \times 18) + (125 \times 10) + (275 \times 13) + (125 \times 10)$$

$$= ₹\ 12{,}075$$

Example 10.6

Obtain an initial basic feasible solution to the following transportation problem using Vogel's approximation method.

Ware houses	I	II	III	IV	Availability (a_i)
A	5	1	3	3	34
B	3	3	5	4	15
C	6	4	4	3	12
D	4	1	4	5	19
Requirement (b_j)	21	25	17	17	

Stores

Solution:

Here $\sum a_i = \sum b_j = 80$ (i.e) Total Availability =Total Requirement

$\therefore$ The given problem is balanced transportation problem.

Hence there exists a feasible solution to the given problem.

First Allocation:

	I	II	III	IV	a_i	Penalty
A	5	1	3	3	34	2
B	3	3	5	4	15	0
C	6	4	4	3	12	1
D	4	(19) 1	4	5	19/0	3
b_j	21	25/6	17	17		
Penalty	1	0	1	0		

Second Allocation:

	I	II	III	IV	a_i	Penalty
A	5	(6) 1	3	3	34/28	2
B	3	3	5	4	15	0
C	6	4	4	3	12	1
b_j	21	6/0	17	17		
Penalty	2	2	1	0		

Third Allocation:

	I	III	IV	a_i	Penalty
A	5	3	3	28	0
B	(15) 3	5	4	15/0	1
C	6	4	3	12	1
b_j	21/6	17	17		
Penalty	2	1	0		

Fourth Allocation:

	I	III	IV	a_i	Penalty
A	5	3	3	28	0
C	6	4	(12) 3	12/0	1
b_j	6	17	17/5		
Penalty	1	1	0		

Fifth Allocation:

	I	III	IV		Penalty
A	5	3	(5) 3	28/23	0
	6	17	5/0		
Penalty	-	-	-		

Sixth Allocation:

	I	III		Penalty
A	(6) 5	(17) 3	23/6/0	2
	6/0	17/0		
Penalty	-	-		

Thus we have the following allocations:

	I	II	III	IV	a_i
A	(6) 5	(6) 1	(17) 3	(5) 3	34

B	(15) 3	3	5	4	15
C	6	4	4	(12) 3	12
D	4	(19) 1	4	5	19
b_j	21	25	17	17	

Transportation schedule :

$$A \rightarrow I,\ A \rightarrow II,\ A \rightarrow III,\ A \rightarrow IV,\ B \rightarrow I,\ C \rightarrow IV,\ D \rightarrow II$$

Total transportation cost:

$$= (6\times5)+(6+1)+(17\times3)+(5\times3)+(15\times3)+(12\times3)+(19\times1)$$

$$= 30+6+51+15+45+36+19$$

$$= ₹\ 202$$

Exercise 10.1

1. What is transportation problem?
2. Write mathematical form of transportation problem.
3. what is feasible solution and non degenerate solution in transportation problem?
4. What do you mean by balanced transportation problem?
5. Find an initial basic feasible solution of the following problem using north west corner rule.

	D_1	D_2	D_3	D_4	Supply
O_1	5	3	6	2	19
O_2	4	7	9	1	37
O_3	3	4	7	5	34
Demand	16	18	31	25	

6. Determine an initial basic feasible solution of the following transportation problem by north west corner method

	Bangalore	Nasik	Bhopal	Delhi	Capacity
Chennai	6	8	8	5	30
Madurai	5	11	9	7	40
Trichy	8	9	7	13	50
Demand (Units/day	35	28	32	25	

7. Obtain an initial basic feasible solution to the following transportation problem by using least- cost method.

	D_1	D_2	D_3	Supply
O_1	9	8	5	25
O_2	6	8	4	35
O_3	7	6	9	40
demand	30	25	45	

8. Explain Vogel's approximation method by obtaining initial feasible solution of the following transportation problem

	D_1	D_2	D_3	D_4	Supply
O_1	2	3	11	7	6
O_2	1	0	6	1	1
O_3	5	8	15	9	10
Demand	7	5	3	2	

9. Consider the following transportation problem

	D_1	D_2	D_3	D_4	Availability
O_1	5	8	3	6	30
O_2	4	5	7	4	50
O_3	6	2	4	6	20
Requirement	30	40	20	10	

Determine initial basic feasible solution by VAM

10. Determine basic feasible solution to the following transportation problem using North west Corner rule.

		Sinks					Supply
		A	*B*	*C*	*D*	*E*	
	P	2	11	10	3	7	4
Origins	*Q*	1	4	7	2	1	8
	R	3	9	4	8	12	9
Demand		3	3	4	5	6	

11. Find the initial basic feasible solution of the following transportation problem:

	I	II	III	Demand
A	1	2	6	7
B	0	4	2	12
C	3	1	5	11
Supply	10	10	10	

Using (i) North West Corner rule

(ii) Least Cost method

(iii) Vogel's approximation method

12. Obtain an initial basic feasible solution to the following transportation problem by north west corner method.

	D	*E*	*F*	*C*	Available
A	11	13	17	14	250
B	16	18	14	10	300
C	21	24	13	10	400
Required	200	225	275	250	

10.2 Assignment Problem:

Introduction:

The assignment problem is a particular case of transportation problem for which more efficient (less-time consuming) solution method has been devised by KUHN (1956) and FLOOD (1956). The justification of the steps leading to the solution is based on theorems proved by Hungarian Mathematicians KONEIG (1950) and EGERVARY (1953), hence the method is named Hungarian Method.

Suppose that we have 'm' jobs to be performed on 'n' machines . The cost of assigning each job to each machine is C_{ij} (i =1,2,…,n and j = 1,2,…,n).Our objective is to assign the different jobs to the different machines(one job per machine) to minimize the overall cost.This is known as **assignment problem.**

The assignment problem is a special case of transportation problem where the number of sources and destinations are equal .Supply at each source and demand at each destination must be one. It means that there is exactly one occupied cell in each row and each column of the transportation table . Jobs represent sources and machines represent destinations.

10.2.1 Definition and formulation

Consider the problem of assigning n jobs to n machines (one job to one machine). Let C_{ij} be the cost of assigning i^{th} job to the j^{th} machine and x_{ij} represents the assignment of i^{th} job to the j^{th} machine.

$$\text{Then, } x_{ij} = \begin{cases} 1, \textit{ if } i^{th} \textit{ job is assigned to } j^{th} \textit{ machine} \\ 0, \textit{ if } i^{th} \textit{ job is not assigned to } j^{th} \textit{ machine.} \end{cases}$$

machines

Machines

Jobs		1	2	…	n	supply
	1	(x_{11}) C_{11}	(x_{12}) C_{12}	…	(x_{1n}) C_{1n}	1
	2	(x_{21}) C_{21}	(x_{22}) C_{22}	…	(x_{2n}) C_{2n}	1
	⋮		⋮	⋮	⋮	1
	m	(x_{ij}) C_{n1}	(x_{ij}) C_{n2}	…	(x_{ij}) C_{nn}	1
	demand	b_1	b_2	…	b_n	

x_{ij} is missing in any cell means that no assignment is made between the pair of job and machine.$(i.e)$ $x_{ij} = 0$.

x_{ij} presents in any cell means that an assignment is made their.In such cases $x_{ij} = 1$

The assignment model can be written in LPP as follows

$$\text{Minimize } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij}\, x_{ij}$$

Subject to the constrains

$$\sum_{i=1}^{n} x_{ij} = 1, \; j = 1, 2, \ldots..n$$

$$\sum_{j=1}^{n} x_{ij} = 1,\ i = 1, 2, \ldots\ldots n \quad \text{and} \quad x_{ij} = 0 \ \text{(or)}\ 1 \ \text{for all}\ i,j$$

Note

The optimum assignment schedule remains unaltered if we add or subtract a constant from all the elements of the row or column of the assignment cost matrix.

Note

If for an assignment problem all $C_{ij} > 0$ then an assignment schedule (x_{ij}) which satisfies $\sum C_{ij}\, x_{ij} = 0$ must be optimal.

10.2.2 Solution of assignment problems (Hungarian Method)

First check whether the number of rows is equal to the numbers of columns, if it is so, the assignment problem is said to be balanced.

Step :1 Choose the least element in each row and subtract it from all the elements of that row.

Step :2 Choose the least element in each column and subtract it from all the elements of that column. Step 2 has to be performed from the table obtained in step 1.

Step:3 Check whether there is atleast one zero in each row and each column and make an assignment as follows.

(i) Examine the rows successively until a row with exactly one zero is found. Mark that zero by $\square$, that means an assignment is made there . Cross$(\times)$ all other zeros in its column. Continue this until all the rows have been examined.

(ii) Examine the columns successively until a columns with exactly one zero is found. Mark that zero by $\square$, that means an assignment is made there . Cross $(\times)$ all other zeros in its row. Continue this until all the columns have been examined

Step :4 If each row and each column contains exactly one assignment, then the solution is optimal.

DO YOU KNOW?

Hungarian method provides optimum assignment schedule in an assignment problem.

Example 10.7

Solve the following assignment problem. Cell values represent cost of assigning job A, B, C and D to the machines I, II, III and IV.

		I	II	III	IV
		machines			
jobs	*A*	10	12	19	11
	B	5	10	7	8
	C	12	14	13	11
	D	8	15	11	9

Solution:

Here the number of rows and columns are equal.

$\therefore$ The given assignment problem is balanced.

Now let us find the solution.

Step 1: Select a smallest element in each row and subtract this from all the elements in its row.

	I	II	III	IV
A	0	2	9	1
B	0	5	2	3
C	1	3	2	0
D	0	7	3	1

Look for atleast one zero in each row and each column.Otherwise go to step 2.

Step 2: Select the smallest element in each column and subtract this from all the elements in its column.

	I	II	III	IV
A	0	0	7	1
B	0	3	0	3
C	1	1	0	0
D	0	5	1	1

Since each row and column contains atleast one zero, assignments can be made.

Step 3 (Assignment):

Examine the rows with exactly one zero. First three rows contain more than one zero. Go to row *D*. There is exactly one zero. Mark that zero by □(i.e) job *D* is assigned to machine I. Mark other zeros in its column by $\times$.

$$\begin{array}{c|cccc|} & I & II & III & IV \\ \hline A & \cancel{0} & 0 & 7 & 1 \\ B & \cancel{0} & 3 & 0 & 3 \\ C & 1 & 1 & 0 & 0 \\ D & \boxed{0} & 5 & 1 & 2 \\ \hline \end{array}$$

Step 4: Now examine the columns with exactly one zero. Already there is an assignment in column I. Go to the column II. There is exactly one zero. Mark that zero by □. Mark other zeros in its rowby ×.

$$\begin{array}{c|cccc|} & I & II & III & IV \\ \hline A & \cancel{0} & \boxed{0} & 7 & 1 \\ B & \cancel{0} & 3 & 0 & 3 \\ C & 1 & 1 & 0 & 0 \\ D & \boxed{0} & 5 & 1 & 2 \\ \hline \end{array}$$

Column III contains more than one zero. Therefore proceed to Column IV, there is exactly one zero. Mark that zero by □. Mark other zeros in its row by ×.

$$\begin{array}{c|cccc|} & I & II & III & IV \\ \hline A & \cancel{0} & \boxed{0} & 7 & 1 \\ B & \cancel{0} & 3 & 0 & 3 \\ C & 1 & 1 & \cancel{0} & \boxed{0} \\ D & \boxed{0} & 5 & 1 & 2 \\ \hline \end{array}$$

Step 5: Again examine the rows. Row B contains exactly one zero. Mark that zero by □.

$$\begin{array}{c|cccc|} & I & II & III & IV \\ \hline A & \cancel{0} & \boxed{0} & 7 & 1 \\ B & \cancel{0} & 3 & \boxed{0} & 3 \\ C & 1 & 1 & \cancel{0} & \boxed{0} \\ D & \boxed{0} & 5 & 1 & 2 \\ \hline \end{array}$$

Thus all the four assignments have been made. The optimal assignment schedule and total cost is

Job	Machine	cost
A	II	12
B	III	7
C	IV	11
D	I	8
Total cost		38

The optimal assignment (minimum) cost

= ₹ 38

Example 10.8

Consider the problem of assigning five jobs to five persons. The assignment costs are given as follows. Determine the optimum assignment schedule.

Person		Job 1	2	3	4	5
	A	8	4	2	6	1
	B	0	9	5	5	4
	C	3	8	9	2	6
	D	4	3	1	0	3
	E	9	5	8	9	5

Solution:

Here the number of rows and columns are equal.

∴ The given assignment problem is balanced.

Now let us find the solution.

Step 1: Select a smallest element in each row and subtract this from all the elements in its row.

The cost matrix of the given assignment problem is

Person		Job 1	2	3	4	5
	A	7	3	1	5	0
	B	0	9	5	5	4
	C	1	6	7	0	4
	D	4	3	1	0	3
	E	4	0	3	4	0

Column 3 contains no zero. Go to Step 2.

Step 2: Select the smallest element in each column and subtract this from all the elements in its column.

$$\begin{array}{cc|ccccc|}
 & & \multicolumn{5}{c}{\text{Job}} \\
 & & 1 & 2 & 3 & 4 & 5 \\ \cline{3-7}
 & A & 7 & 3 & 0 & 5 & 0 \\
 & B & 0 & 9 & 4 & 5 & 4 \\
\text{Person} & C & 1 & 6 & 6 & 0 & 4 \\
 & D & 4 & 3 & 0 & 0 & 3 \\
 & E & 4 & 0 & 2 & 4 & 0 \\ \cline{3-7}
\end{array}$$

Since each row and column contains atleast one zero, assignments can be made.

Step 3 (Assignment):

Examine the rows with exactly one zero. Row B contains exactly one zero. Mark that zero by □ (i.e) PersonB is assigned to Job 1. Mark other zeros in its column by × .

$$\begin{array}{cc|ccccc|}
 & & \multicolumn{5}{c}{\text{Job}} \\
 & & 1 & 2 & 3 & 4 & 5 \\ \cline{3-7}
 & A & 7 & 3 & 0 & 5 & 0 \\
 & B & \boxed{0} & 9 & 4 & 5 & 4 \\
\text{Person} & C & 1 & 6 & 6 & 0 & 4 \\
 & D & 4 & 3 & 0 & 0 & 3 \\
 & E & 4 & 0 & 2 & 4 & 0 \\ \cline{3-7}
\end{array}$$

Now, Row C contains exactly one zero. Mark that zero by □. Mark other zeros in its column by ×.

$$\begin{array}{cc|ccccc|}
 & & \multicolumn{5}{c}{\text{Job}} \\
 & & 1 & 2 & 3 & 4 & 5 \\ \cline{3-7}
 & A & 7 & 3 & 0 & 5 & 0 \\
 & B & \boxed{0} & 9 & 4 & 5 & 4 \\
\text{Person} & C & 1 & 6 & 6 & \boxed{0} & 4 \\
 & D & 4 & 3 & 0 & \xcancel{0} & 3 \\
 & E & 4 & 0 & 2 & 4 & 0 \\ \cline{3-7}
\end{array}$$

Now, Row D contains exactly one zero. Mark that zero by □. Mark other zeros in its column by ×.

Job

Person	1	2	3	4	5
A	7	3	$\cancel{0}$	5	0
B	$\boxed{0}$	9	4	5	4
C	1	6	6	$\boxed{0}$	4
D	4	3	$\boxed{0}$	$\cancel{0}$	3
E	4	0	2	4	0

Row E contains more than one zero, now proceed column wise. In column 1, there is an assignment. Go to column 2. There is exactly one zero. Mark that zero by □. Mark other zeros in its row by ×.

Job

Person	1	2	3	4	5
A	7	3	$\cancel{0}$	5	0
B	$\boxed{0}$	9	4	5	4
C	1	6	6	$\boxed{0}$	4
D	4	3	$\boxed{0}$	$\cancel{0}$	3
E	4	$\boxed{0}$	2	4	$\cancel{0}$

There is an assignment in Column 3 and column 4. Go to Column 5. There is exactly one zero. Mark that zero by □. Mark other zeros in its row by ×.

Job

Person	1	2	3	4	5
A	7	3	$\cancel{0}$	5	$\boxed{0}$
B	$\boxed{0}$	9	4	5	4
C	1	6	6	$\boxed{0}$	4
D	4	3	$\boxed{0}$	$\cancel{0}$	3
E	4	$\boxed{0}$	2	4	$\cancel{0}$

Thus all the five assignments have been made. The Optimal assignment schedule and total cost is

Person	Job	cost
A	5	1
B	1	0
C	4	2
D	3	1
E	2	5
Total cost		9

The optimal assignment (minimum) cost = ₹ 9

Example 10.9

Solve the following assignment problem.

$$\begin{array}{cc} & \text{Men} \\ & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \text{Task} \begin{array}{c} P \\ Q \\ R \\ S \end{array} & \left[\begin{array}{ccc} 9 & 26 & 15 \\ 13 & 27 & 6 \\ 35 & 20 & 15 \\ 18 & 30 & 20 \end{array}\right] \end{array}$$

Solution:

Since the number of columns is less than the number of rows, given assignment problem is unbalanced one. To balance it , introduce a dummy column with all the entries zero. The revised assignment problem is

$$\begin{array}{cc} & \text{Men} \\ & \begin{array}{cccc} 1 & 2 & 3 & d \end{array} \\ \text{Task} \begin{array}{c} P \\ Q \\ R \\ S \end{array} & \left[\begin{array}{cccc} 9 & 26 & 15 & 0 \\ 13 & 27 & 6 & 0 \\ 35 & 20 & 15 & 0 \\ 18 & 30 & 20 & 0 \end{array}\right] \end{array}$$

Here only 3 tasks can be assigned to 3 men.

Step 1: is not necessary, since each row contains zero entry. Go to Step 2.

Step 2 :

$$\begin{array}{cc} & \text{Men} \\ & \begin{array}{cccc} 1 & 2 & 3 & d \end{array} \\ \text{Task} \begin{array}{c} P \\ Q \\ R \\ S \end{array} & \left[\begin{array}{cccc} 0 & 6 & 9 & 0 \\ 4 & 7 & 0 & 0 \\ 26 & 0 & 9 & 0 \\ 9 & 10 & 14 & 0 \end{array}\right] \end{array}$$

Step 3 (Assignment) :

$$
\begin{array}{cc|cccc|}
 & & & \text{Men} & & \\
 & & 1 & 2 & 3 & d \\
\hline
 & P & \boxed{0} & 6 & 9 & \cancel{0} \\
\text{Task} & Q & 4 & 7 & \boxed{0} & \cancel{0} \\
 & R & 26 & \boxed{0} & 9 & \cancel{0} \\
 & S & 9 & 10 & 14 & \boxed{0} \\
\hline
\end{array}
$$

Since each row and each columncontains exactly one assignment,all the three men have been assigned a task. But task S is not assigned to any Man. The optimal assignment schedule and total cost is

Task	Men	cost
P	1	9
Q	3	6
R	2	20
S	d	0
Total cost		35

The optimal assignment (minimum) cost = ₹ 35

Exercise 10.2

1. What is the Assignment problem?
2. Give mathematical form of assignment problem.
3. What is the difference between Assignment Problem and Transportation Problem?
4. Three jobs A, B and C one to be assigned to three machines U, V and W. The processing cost for each job machine combination is shown in the matrix given below. Determine the allocation that minimizes the overall processing cost.

$$
\begin{array}{cc|ccc|}
 & & & \text{Machine} & \\
 & & U & V & W \\
\hline
 & A & 17 & 25 & 31 \\
\text{Job} & B & 10 & 25 & 16 \\
 & C & 12 & 14 & 11 \\
\hline
\end{array}
$$

(cost is in ₹ per unit)

5. A computer centre has got three expert programmers. The centre needs three application programmes to be developed. The head of the computer centre, after studying carefully the programmes to be developed, estimates the computer time in minitues required by the experts to the application programme as follows.

		Programmes		
		P	Q	R
Programmers	1	120	100	80
	2	80	90	110
	3	110	140	120

Assign the programmers to the programme in such a way that the total computer time is least.

6. A departmental head has four subordinates and four tasks to be performed. The subordinates differ in efficiency and the tasks differ in their intrinsic difficulty. His estimates of the time each man would take to perform each task is given below

		Tasks			
		1	2	3	4
Subordinates	*P*	8	26	17	11
	Q	13	28	4	26
	R	38	19	18	15
	S	9	26	24	10

How should the tasks be allocated to subordinates so as to minimize the total man-hours?

7. Find the optimal solution for the assignment problem with the following cost matrix.

		Area			
		1	2	3	4
Salesman	*P*	11	17	8	16
	Q	9	7	12	6
	R	13	16	15	12
	S	14	10	12	11

8. Assign four trucks 1, 2, 3 and 4 to vacant spaces A, B, C, D, E and F so that distance travelled is minimized. The matrix below shows the distance.

	1	2	3	4
A	4	7	3	7
B	8	2	5	5
C	4	9	6	9
D	7	5	4	8
E	6	3	5	4
F	6	8	7	3

10. 3 DECISION THEORY

Introduction:

Decision theory is primarily concerned with helping people and organizations in making decisions. It provides a meaningful conceptual frame work for important decision making. The decision making refers to the selection of an act from amongst various alternatives, the one which is judged to be the best under given circumstances.

The management has to consider phases like planning, organization, direction, command and control. While performing so many activities, the management has to face many situations from which the best choice is to be taken. This choice making is technically termed as "decision making" or decision taking. A decision is simply a selection from two or more courses of action. Decision making may be defined as - " a process of best selection from a set of alternative courses of action, that course of action which is supposed to meet objectives upto satisfaction of the decision maker."

The knowledge of statistical techniques helps to select the best action. The statistical decision theory refers to an optimal choice under condition of uncertainty. In this case probability theory has a vital role, as such, this probability theory will be used more frequently in the decision making theory under uncertainty and risk.

The statistical decision theory tries to reveal the logical structure of the problem into alternative action, states of nature, possible outcomes and likely pay-offs from each such outcome. Let us explain the concepts associated with the decision theory approach to problem solving.

10.3.1 Meaning

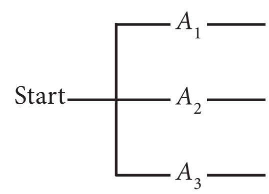
The decision maker: The decision maker refers to individual or a group of individual responsible for making the choice of an appropriate course of action amongst the available courses of action.

Acts (or courses of action): Decision making problems deals with the selection of a single act from a set of alternative acts. If two or more alternative courses of action occur in a problem, then decision making is necessary to select only one course of action. Let the acts or action be $a_1, a_2, a_3, \ldots$ then the totality of all these actions is known as action space denoted by A. For three actions $a_1, a_2\ a_3$; A = action space = (a_1, a_2, a_3) or A = (A_1, A_2, A_3). Acts may be also represented in the following matrix form.

Acts
A_1
A_2
$\ldots$
A_n

(or)

Acts	A_1	A_2	$\ldots$	A_n

In a tree diagram the acts or action are shown as 

Events (or States of nature): The events identify the occurrences, which are outside of the decision maker' s control and which determine the level of success for a given act. These events are often called ' States of nature' or outcomes. An example of an event or states of nature is the level of market demand for a particular item during a stipulated time period.

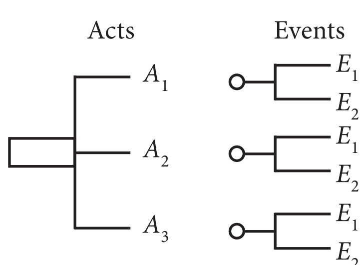
A set of states of nature may be represented in any one of the following ways:

$S = \{S_1, S_2, \ldots, S_n\}$ or $E = \{E_1, E_2, \ldots, E_n\}$ or $\Omega = \{\theta_1, \theta_2, \theta_3\}$

For example, if a washing powder is marketed, it may be highly liked by outcomes (outcome θ1) or it may not appeal at all (outcome θ2) or it may satisfy only a small fraction, say 25% (outcome θ3)

$\therefore \Omega = \{\theta_1, \theta_2, \theta_3\}$

In a tree diagram the places are next to acts. We may also get another act on the happening of events as follows:



In a matrix form, they may be represented as either of the two ways.

States of nature $\longrightarrow$ Acts $\downarrow$	S_1	S_2
A_1 A_2		

or

Acts $\longrightarrow$ States of nature $\downarrow$	$A_1 A_2, \ldots A_n$
S_1 S_2	

Pay-off: The result of combinations of an act with each of the states of nature is the outcome and monetary gain or loss of each such outcome is the pay-off. This means that the expression pay-off should be in quantitative form.

Pay -off may be also in terms of cost saving or time saving. In general, if there are k alternatives and n states of nature, there will be $k \times n$ outcomes or pay-offs. These k × n payoffs can be very conveniently represented in the form of a $k \times n$ pay -off table.

States of nature	Decision alternative			
	A_1	A_2	....	A_k
E_1	a_{11}	a_{12}	....	a_{1k}
E_2	a_{21}	a_{22}	....	a_{2k}
.	.	.	....	.
.	. .	. .	....	. .
E_n	a_{n1}	a_{n2}	....	a_{nk}

Where a_{ij} = conditional outcome (pay-off) of the i^{th} event when j^{th} alternative is chosen. The above pay-off table is called pay-off matrix.

10.3.2 Situations- Certainty and uncertainty

Types of decision making: Decisions are made based upon the information data available about the occurrence of events as well as the decision situation. There are two types of decision making situations: certainty and uncertainty

Decision making under certainty: In this case the decision maker has the complete knowledge of consequence of every decision choice with certainty. In this decision model, assumed certainty means that only one possible state of nature exists.

Decision making under uncertainty: Under conditions of uncertainty, only pay-offs are known and nothing is known about the lilkelihood of each state of nature. Such situations arise when a new product is introduced in the market or a new plant is set up. The number of different decision criteria available under the condition of uncertainty is given below.

10.3.3 Maximin and Minimax strategy

Maximin criteria

This criterion is the decision to take the course of action which maximizes the minimum possible pay-off. Since this decision criterion locates the alternative strategy that has the least possible loss, it is also known as a pessimistic decision criterion. The working method is:

(i) Determine the lowest outcome for each alternative.

(ii) Choose the alternative associated with the maximum of these.

Minimax criteria

This criterion is the decision to take the course of action which minimizes the maximum possible pay-off. Since this decision criterion locates the alternative strategy that has the greatest possible gain. The working method is:

(i) Determine the highest outcome for each alternative.

(ii) Choose the alternative associated with the minimum of these.

Example 10.10

Consider the following pay-off (profit) matrix Action States

Action	States			
	(S_1)	(S_2)	(S_3)	(S_4)
A_1	5	10	18	25
A_2	8	7	8	23
A_3	21	18	12	21
A_4	30	22	19	15

Determine best action using maximin principle.

Solution:

Action	States				Minimum
	(S_1)	(S_2)	(S_3)	(S_4)	
A_1	5	10	18	25	5
A_2	8	7	8	23	7
A_3	21	18	12	21	12
A_4	30	22	19	15	15

Max (5,7,12,15)=15 $\therefore$ Action A4 is the best

Example 10.11

A business man has three alternatives open to him each of which can be followed by any of the four possible events. The conditional pay offs for each action - event combination are given below:

Alternative	Pay – offs (Conditional events)			
	A	B	C	D
X	8	0	–10	6
Y	–4	12	18	–2
Z	14	6	0	8

Determine which alternative should the businessman choose, if he adopts the maximin principle.

Solution:

Alternative	Pay – offs (Conditional events)				Minimum pay off
	A	B	C	D	
X	8	0	–10	6	–10
Y	–4	12	18	-2	–4
Z	14	6	0	8	**0**

Max (–10,–4, 0) = 0. Since the maximum payoff is 0, the alternative Z is selected by the businessman.

Example 10.12

Consider the following pay-off matrix

Alternative	Pay – offs (Conditional events)			
	A_1	A_2	A_3	A_4
E_1	7	12	20	27
E_2	10	9	10	25
E_3	23	20	14	23
E_4	32	24	21	17

Using minmax principle, determine the best alternative.

Solution:

Alternative	Pay – offs (Conditional events)				Minimum pay off
	A_1	A_2	A_3	A_4	
E_1	7	12	20	27	27
E_2	10	9	10	25	25
E_3	23	20	14	23	23
E_4	32	24	21	17	32

min(27, 25, 23, 32) = 23. Since the minimum cost is 23, the best alternative is E_3 according to minimax principle.

Exercise 10.3

1. Given the following pay-off matrix(in rupees) for three strategies and two states of nature.

Strategy	States-of-nature	
	E_1	E_2
S_1	40	60
S_2	10	–20
S_3	–40	150

Select a strategy using each of the following rule (i) Maximin (ii) Minimax

2. A farmer wants to decide which of the three crops he should plant on his 100-acre farm. The profit from each is dependent on the rainfall during the growing season. The farmer has categorized the amount of rainfall as high medium and low. His estimated profit for each is shown in the table.

Rainfall	Estimated Conditional Profit(Rs.)		
	crop A	Crop B	Crop C
High	8000	3500	5000
Medium	4500	4500	5000
Low	2000	5000	4000

If the farmer wishes to plant only crop, decide which should be his best crop using

(i) Maximin (ii)Minimax

3. The research department of Hindustan Ltd. has recommended to pay marketing department to launch a shampoo of three different types. The marketing types of shampoo to be launched under the following estimated pay-offs for various level of sales.

Types of shampoo	Estimated Sales (in Units)		
	15000	10000	5000
Egg shampoo	30	10	10
Clinic Shampoo	40	15	5
Deluxe Shampoo	55	20	3

What will be the marketing manager's decision if (i) Maximin and (ii) Minimax principle applied?

4. Following pay-off matrix, which is the optimal decision under each of the following rule (i) maxmin (ii) minimax

Act	States of nature			
	S_1	S_2	S_3	S_4
A_1	14	9	10	5
A_2	11	10	8	7
A_3	9	10	10	11
A_4	8	10	11	13

Exercise 10.4

HIXQEY

Choose the correct Answer

1. The transportation problem is said to be unbalanced if _________

   a) Total supply ≠ Total demand (b) Total supply = Total demand

   (c) $m = n$ (d) $m+n-1$

2. In a non – degenerate solution number of allocations is

   (a) Equal to $m+n-1$ (b) Equal to $m+n+1$

   (c) Not equal to $m+n-1$ (d) Not equal to $m+n+1$

3. In a degenerate solution number of allocations is

   (a) equal to $m+n-1$ (b) not equal to $m+n-1$

   (c) less than $m+n-1$ (d) greather than $m+n-1$

4. The Penalty in VAM represents difference between the first ________

   (a) Two largest costs (b) Largest and Smallest costs

   (c) Smallest two costs (d) None of these

5. Number of basic allocation in any row or column in an assignment problem can be

   (a) Exactly one (b) at least one (c) at most one (d) none of these

6. North-West Corner refers to ________

   a) top left corner (b) top right corner

   (c) bottom right corner (d) bottom left corner

7. Solution for transportation problem using ___________method is nearer to an optimal solution.

a)NWCM (b) LCM (c) VAM (d) Row Minima

8. In an assignment problem the value of decision variable x_{ij} is __________.

(a) 1 (b) 0 c) 1 or 0 (d) none of them

9. If number of sources is not equal to number of destinations, the assignment problem is called_________________

(a) balanced (b) unsymmetric (c) symmetric (d) unbalanced

10. The purpose of a dummy row or column in an assignment problem is to

(a) prevent a solution from becoming degenerate

(b) balance between total activities and total resources

(c) provide a means of representing a dummy problem

(d) none of the above

11. The solution for an assignment problem is optimal if

(a) each row and each column has no assignment

(b) each row and each column has atleast one assignment

(c) each row and each column has atmost one assignment

(d) each row and each column has exactly one assignment

12. In an assignment problem involving four workers and three jobs, total number of assignments possible are

(a) 4 (b) 3 (c) 7 (d) 12

13. Decision theory is concerned with

(a) analysis of information that is available

(b) decision making under certainty

(c) selecting optimal decisions in sequential problem

(d) All of the above

14. A type of decision –making environment is

(a) certainty (b) uncertainty (c) risk (d) all of the above

Miscellaneous Problems

1. The following table summarizes the supply, demand and cost information for four factors $S_{1,}$ $S_{2,}$ $S_{3,}$ $S_{4.}$ shipping goods to three warehouses $D_{1,}$ $D_{2,}$ D_3

	D_1	D_2	D_3	Supply
S_1	2	7	14	5
S_2	3	3	1	8
S_3	5	4	7	7
S_4	1	6	2	14
Demand	7	9	18	

Find an initial solution by using north west corner rule. What is the total cost for this solution?

2. Consider the following transportation problem

	Destination				
	D_1	D_2	D_3	D_4	Availability
O_1	5	8	3	6	30
O_2	4	5	7	4	50
O_3	6	2	4	6	20
Requirement	30	40	20	10	

Determine an initial basic feasible solution using (a) Least cost method (b) Vogel's approximation method.

3. Determine an initial basic feasible solution to the following transportation problem by using (i)North West Corner rule (ii) least cost method.

		Destination			
		D_1	D_2	D_3	Supply
Source	S_1	9	8	5	25
	S_2	6	8	4	35
	S_3	7	6	9	40
	Requirement	30	25	45	

4. Explain Vogel's approximation method by obtaining initial basic feasible solution of the following transportation problem.

		Destination				
		D_1	D_2	D_3	D_4	Supply
Origin	O_1	2	3	11	7	6
	O_2	1	0	6	1	1
	O_3	5	8	15	9	10
	Demand	7	5	3	2	

5. A car hire company has one car at each of five depots a,b,c,d and e. A customer in each of the fine towers A,B,C,D and E requires a car. The distance (in miles) between the depots (origins) and the towers(destinations) where the customers are given in the following distance matrix.

	a	b	c	d	e
A	160	130	175	190	200
B	135	120	130	160	175
C	140	110	155	170	185
D	50	50	80	80	110
E	55	35	70	80	105

How should the cars be assigned to the customers so as to minimize the distance travelled?

6. A natural truck-rental service has a surplus of one truck in each of the cities 1,2,3,4,5 and 6 and a deficit of one truck in each of the cities 7,8,9,10,11 and 12. The distance(in kilometers) between the cities with a surplus and the cities with a deficit are displayed below:

		To					
		7	8	9	10	11	12
From	1	31	62	29	42	15	41
	2	12	19	39	55	71	40
	3	17	29	50	41	22	22
	4	35	40	38	42	27	33
	5	19	30	29	16	20	33
	6	72	30	30	50	41	20

How should the truck be dispersed so as to minimize the total distance travelled?

7. A person wants to invest in one of three alternative investment plans: Stock, Bonds and Debentures. It is assumed that the person wishes to invest all of the funds in a

plan. The pay-off matrix based on three potential economic conditions is given in the following table:

Alternative	Economic conditions		
	High growth(Rs.)	Normal growth(Rs.)	Slow growth (Rs.)s
Stocks	10000	7000	3000
Bonds	8000	6000	1000
Debentures	6000	6000	6000

Determine the best investment plan using each of following criteria i) Maxmin ii) Minimax.

Summary

- In a transportation problem if the total supply equals the total demand, it is said to be balanced transportation problem. Otherwise it is said to be unbalanced transportation problem
- **Feasible Solution**: A feasible solution to a transportation problem is a set of non-negative values x_{ij}(i=1,2,..,m, j=1,2,…n) that satisfies the constraints.
- **Basic Feasible Solution**: A feasible solution is called a basic feasible solution if it contains not more than m+n-1 allocations, where m is the number of rows and n is the number of columns in a transportation table.
- **Optimal Solution**: Optimal Solution is a feasible solution (not necessarily basic) which optimizes(minimize) the total transportation cost.
- **Non degenerate basic feasible Solution**: If a basic feasible solution to a transportation problem contains exactly m+n-1 allocations in independent positions, it is called a Non degenerate basic feasible solution.
- **Degeneracy : If** : If a basic feasible solution to a transportation problem contains less than m+n-1 allocations , it is called a degenerate basic feasible solution.
- In an assignment problems number of rows and columns must be equal
- The optimum assignment schedule remains unaltered if we add or subtract a constant from all the elements of the row or column of the assignment cost matrix.
- If for an assignment problem all $C_{ij} > 0$ then an assignment schedule (x_{ij}) which satisfies $\sum C_{ij}\, x_{ij} = 0$ must be optimal.

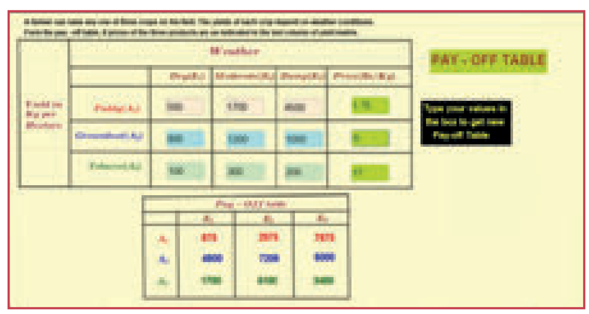
GLOSSARY

Approximation	தோராயமாக
Assignment problems	ஒதுக்கீடு கணக்குகள்
Decision theory	முடிவு கோட்பாடுகள்
Degenerate	சிதைந்த

Destination	சேருமிடம்
Feasible solution	ஏற்புடையத்தீர்வு
Initial basic feasible solution	ஆரம்பஅடிப்படைஏற்புடையத்தீர்வு
Least cost method	மீச்சிறு செலவு முறை
Non-degenerate	சிதைவற்ற
North west- Conner method	வட மேற்கு மூலைமுறை
Optimum solution	உகந்ததீர்வு
Pay off	இழப்பு ஈட்டியப்பு
Strategy	உத்தி
Transportation cost	போக்குவரத்து செலவு
Transportation problems	போக்குவரத்து கணக்குகள்

ICT Corner

Expected Result is shown in this picture

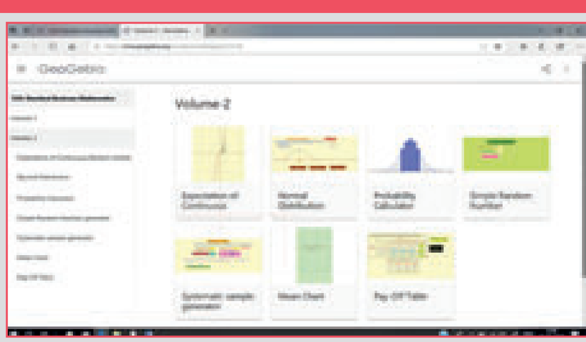


Step – 1 : Open the Browser, type the URL Link given below (or) Scan the QR Code. GeoGebra work Book named "12th Standard Business Mathematics" will open. In the work book there are two Volumes. Select "Volume-2".

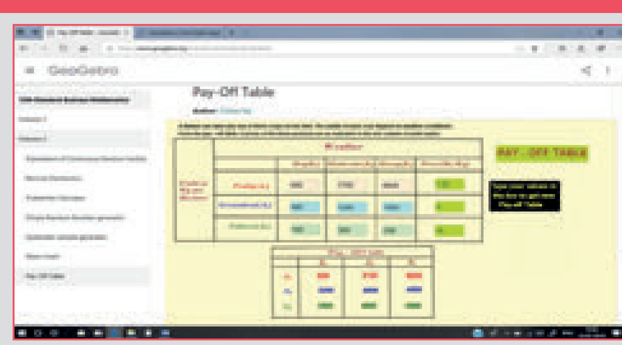
Step - 2 : Select the worksheet named "Pay-Off Table."

You can enter your own data in the respective boxes. Your Pay-Off table will be generated below. Calculate and check your answer.

Step 1



Step 2



Scan the QR Code.

B284_12_BUS_MAT_EM

Answers

6. Random Variable and Mathematical Expectation

Exercise 6.1

1.

$P(X \le k)$	0.3	0.5	0.9	1

2. $F_X(x) = \begin{cases} 0 & for < 3 \\ P_X(3) = 0.3 & for\, 3 \le x < 5 \\ P_X(3) + P_X(5) = 0.5 & for\, 5 \le x < 8 \\ P_X(3) + P_X(5) + P_X(8) = 0.8 & for\, 8 \le x < 10 \\ 1 & for\, x \ge 10 \end{cases}$

5.

$X = x_1$	0	1	2
$P(X = x_1)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

6. (i) 1/10, (ii) 81/100, 19/100, 8/10 (iii) 4

7. Hint: $\int_{-3}^{3} f(x)dx = \int_{-3}^{-1} f(x)dx + \int_{-1}^{1} f(x)dx + \int_{1}^{3} f(x)dx$

8. $(i)\, k = \frac{1}{16}, (ii)\ f(x) = \frac{1}{4}(x-1)^3, 1 \le x \le 3$

9. $A = \frac{1}{5}, (i) \frac{1}{e^2}, (ii) \frac{e-1}{e}, (iii) \frac{e-1}{e^2}$

10. $(a) yes, f(x) = \begin{cases} 0,\ for\ x < 0 \\ \frac{1}{2},\ for\ 0 \le x < 1 \\ 0,\ for\ 1 \le x < 2 \\ \frac{1}{4},\ for\ 2 \le x < 4 \\ 0\ ,\ for\ x \ge 4 \end{cases}$ $(b)(i) \frac{1}{4}, (ii) \frac{3}{4}, (iii) \frac{1}{4}$

Exercise 6.2

1. 3.5 2. 1.8 3. 2.15 4. $\frac{2}{3}$ 5. $\frac{3}{2}, \frac{3}{4}$ 6. ₹ 60

12. Expectation: ₹ 200; Variance: ₹ 21, 60,000; Standard Deviation: ₹ 1,469.69

13. 30 (or 30,000 miles) 14. Expectation: 1; Variance: 9 15. 20

Exercise 6.3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
(c)	(d)	(b)	(c)	(b)	(c)	(d)	(d)	(d)	(d)	(d)	(a)	(c)	(c)	(a)
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
(b)	(b)	(a)	(d)	(b)	(b)	(c)	(c)	(a)	(b)	(c)	(b)	(b)	(b)	(b)

Miscellaneous problems

1. (i) ½ (ii) ¼ (iii) ½ (iv) ¾

2. (a) (i) 13/24 (ii) 0 (b) X is NOT discrete since F is not a step function.

3. $\frac{1}{4};\frac{1}{2}$ 4. (a) $\frac{1}{9}$ (b) $\frac{7}{9}$ 5. (i) $\frac{3}{5},\frac{6}{5}$ (ii) $\frac{2}{25}$ 7. 2 9. $\frac{3}{4},\frac{27}{80}$ 10. $\frac{1}{2}$

7. Probability distributions

Exercise: 7.1

6. (a) 0.059 (b) 0.2642 (c) 0.0133 (d) mean = 1 and variance = 0.95

7. (i) 0.01008 (ii) 0.00262 (iii) 0.09935 8. 0.375 9. 0.5767

10. (i) 0.3969 (ii) 0.45212 (iii) 0.9797 11. 5 or more trials 12. 0.7530

13. (i) 703 (ii) 516 (iii) 656 14. (i) 0.0634 (ii) 0.0634 (iii) 0.9729

15. 25/216 16. $\binom{25}{x}\left(\frac{1}{5}\right)^{x}\left(\frac{4}{5}\right)^{(25-x)}$ 17. $\frac{3}{4^{14}}$ 18. 0.2626 19. 0.8743

20. (i) $\frac{80}{243}$ (ii) $\frac{192}{243}$

Exercise: 7.2

6. 0.2352 7. 0.0025 8. (i) 0.2231 (ii) 0.1912

9.(i) 0.08208 (ii) 0.2138 (iii) 0.1089 10. (i) 2 days (ii) 91 days (iiii) 80 days

11. 0.0265 12. (i) 0.1353 (ii) 0.3235

Exercise: 7.3

5. (i) 67 (ii) 134 (iii) 1637 6 (i) mean = 60.48 (ii) standard deviation = 19.78

7. (i) 0.9772 (ii) 0.49865 8. (a) 46 (b) 46 (c) 342

9. 0.719 10. (i) 0.2420 (ii) 0.8413

Exercise 7.4

1	2	3	4	5	6	7	8	9	10	11	12	13	14
(b)	(c)	(c)	(c)	(c)	(a)	(b)	(a)	(c)	(d)	(a)	(a)	(d)	(b)
15	16	17	18	19	20	21	22	23	24	25	26	27	28
(d)	(b)	(d)	(d)	(d)	(d)	(a)	(a)	(b)	(b)	(a)	(c)	(d)	(d)

Miscellaneous problems

1. (i) 0.89131 (ii) 0.34173 2. 0.03295 3. 0.98981

4. 0.0067379 or 6.7379×10^{-3} 5. 80.33%

6. a) 0.4013 (b) 0.3413 7. a) 30.85% b) 37.20% c) 10.56%

8. a) 0.9938 (b) 0.9878 (c) 0.3944 9. 0.2119 10. 7

Exercise 8.1

17. 0.008 18. 0.9487 19. 0.2739 20. 0.6 21. 0.025

Exercise 8.2

14. $|z| = 1.667$ 15. 1.2308 16. $|z| = 5$ 17. 3.536

Exercise 8.3

1	2	3	4	5	6	7	8	9	10
(a)	(b)	(a)	(b)	(b)	(a)	(c)	(b)	(c)	(a)
11	12	13	14	15	16	17	18	19	20
(a)	(a)	(a)	(d)	(b)	(b)	(a)	(c)	(a)	(c)

Miscellaneous problems

5. 0.015 6. (a) (66.86, 68.04) (b) (66.67,68.22) 7. $|z| = 2.67$

9. Applied Statistics

Exercise 9.1

13. Seasonal Indices

	I	II	III	IV
Total	18.6	20.8	18.8	20.8
Average	3.72	4.16	3.76	4.16
Seasonal indices	94.1772	105.3165	95.1899	105.3165

The Grand Average = 3.95

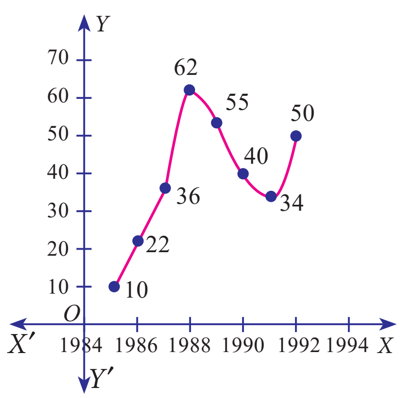
14. Three yearly moving average

Year	1987	1988	1989	1990	1991	1992
Three yearly moving total	46410	52010	63040	79470	94050	102450
Three yearly moving average	15470	17336.666	21013.333	26490	31350	34150

15. Five yearly moving average

Year	1981	1982	1983	1984	1985	1986	1987	1988
Four yearly moving total	619	617	624	621	615	619	613	606
Four yearly moving average	123.8	123.4	124.8	124.2	123	123.8	122.6	121.2

16. Free hand method



17. $a = 169.428$; $b = 3.285$; $Y = 169.428 + 3.285\ X$

18. $a = 54$; $b = 5.4$; $Y = 54 + 5.4\ X$

When $X = 2000$, $\hat{Y} = 54 + 5.4\ (2000\text{-}2002) = 43.2$

When $X = 2001$, $\hat{Y} = 54 + 5.4\ (2001\text{-}2002) = 48.6$

When $X = 2002$, $\hat{Y} = 54 + 5.4\ (2002\text{-}2002) = 54$

When $X = 2003$, $\hat{Y} = 54 + 5.4\ (2003\text{-}2002) = 59.4$

When $X = 2004$, $\hat{Y} = 54 + 5.4\ (2004\text{-}2002) = 64.8$

19. Semi- Average I = 276.666

Semi- Average II = 213.333

20. Monthly Indices

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
Monthly Total	53	61	54	43	42	51	62	54	52	51	49	67
Monthly Average	17.6	20.3	18	14.3	14	17	20.6	18	17.3	17	16.3	22.3
Seasonal Indices	99.4	114.5	101.4	80.7	78.8	97.1	116.4	101.4	97.6	95.7	92	125.8

Grand Average = 17.74

21.

	I	II	III	IV
Total	372	358	362	364
Average	74.4	71.6	72.4	72.8
Seasonal indices	102.19	98.35	99.45	100

Grand Average = 72.8

22. $a = 46.8$; $b = 3$; $Y = 46.8 + 3\,X$

When $X = 1992$, $\hat{Y} = 46.8 + 3\,(1992\text{-}1994) = 40.8$

When $X = 1993$, $\hat{Y} = 46.8 + 3\,(1993\text{-}1994) = 43.8$

When $X = 1994$, $\hat{Y} = 46.8 + 3\,(1994\text{-}1994) = 46.8$

When $X = 1995$, $\hat{Y} = 46.8 + 3\,(1995\text{-}1994) = 49.8$

When $X = 1996$, $\hat{Y} = 46.8 + 3\,(1996\text{-}1994) = 52.8$

When $X = 1997$, $\hat{Y} = 46.8 + 3\,(1997\text{-}1994) = 55.8$

Exercise 9.2

14. Laspeyre's IN = 144.8   Paasche's IN = 144.4
15. Laspeyre's IN = 228.2   Paasche's IN = 225.4
16. Laspeyre's IN = 106.6   Paasche's IN = 106.8   Fisher's IN= 106.7
17. Fisher's IN = 138.5   TRT= 1 FRT = 1880/1560
18. Fisher's IN = 83
19. Fisher's IN = 103   TRT = 1
20. Cost of Living Index =2662.38
21. Cost of Living Index = 117.31
22. Cost of Living Index = 131.49

Exercise 9.3

14.	$\overline{\overline{X}} = 16.2$,	UCL = 20.49,	CL = 16.2,	LCL = 11.91
	$\overline{R} = 7.4$,	UCL = 15.65,	CL = 7.4,	LCL = 0
15.	$\overline{\overline{X}} = 46.2$,	UCL = 50.14,	CL = 46.2,	LCL = 42.26
	$\overline{R} = 6.8$,	UCL = 14.38,	CL = 6.8,	LCL = 0
16.	$\overline{\overline{X}} = 37.71$,	UCL = 56.12,	CL = 37.71,	LCL = 19.29
	$\overline{R} = 18$,	UCL = 46.25,	CL = 18,	LCL = 0
17.	$\overline{\overline{X}} = 10.66$,	UCL = 14.31,	CL = 10.66,	LCL = 7.006
	$\overline{R} = 6.3$,	UCL = 13.32,	CL = 6.3,	LCL = 0
18.	$\overline{\overline{X}} = 12.5$,	UCL = 12.71,	CL = 12.5,	LCL = 12.28
	$\overline{R} = 0.37$,	UCL = 0.78,	CL = 0.37,	LCL = 0
19.	$\overline{\overline{X}} = 30.1$,	UCL = 44.77,	CL = 30.1,	LCL = 15.43
	$\overline{R} = 20.1$,	UCL = 45.83,	CL = 20.1,	LCL = 0
20.	$\overline{\overline{X}} = 13.25$,	UCL = 15.53,	CL = 13.25,	LCL = 10.97
	$\overline{R} = 3.12$,	UCL = 7.12,	CL = 3.12,	LCL = 0
21.	$\overline{\overline{X}} = 24.8$,	UCL = 27.12,	CL = 24.8,	LCL = 22.48
	$\overline{R} = 4$,	UCL = 8.44,	CL = 4,	LCL = 0

Exercise 9.4

1	2	3	4	5	6	7	8	9	10	11	12	13
(d)	(b)	(d)	(d)	(c)	(a)	(c)	(b)	(b)	(b)	(a)	(d)	(c)
14	15	16	17	18	19	20	21	22	23	24	25	
(d)	(c)	(b)	(c)	(c)	(a)	(d)	(c)	(b)	(a)	(c)	(d)	

Miscellaneous problems

1. Three yearly moving Average
   148, 149.33, 152.33, 168.33, 253.33, 261.33, 281.67, 302.67, 327.
2. Four yearly moving Average
   708.75, 729.25, 748.25, 768.25, 784.5

3. Y = 55.975+ 0.825X
4. L= 49.9 P= 50.32 Fisher's = 50.09
5. Fisher =139.8
6. Consumer price index = 118.77
7. CLI = 126·10. The cost of living has increased upto 26.10% in 2011 as compared to 2010.
8. Control chart for Mean — Control chart for Range

   LCL = 47.56 — LCL = 0

   CL = 51.2 — CL = 6.3

   UCL = 54.84 — UCL = 13.32
9. Control chart for Mean — Control chart for Range

   LCL = 1120.83 — LCL = 0

   CL = 1367.5 — CL = 427.5

   UCL = 1614.17 — UCL = 904.16
10. Control chart for Mean — Control chart for Range

   LCL = 4.774 — LCL = 0

   CL = 4.982 — CL = 0.36

   UCL = 5.19 — UCL = 0.7614

10. Operations Research

Exercise 10.1

5. $x_{11} = 16,\ x_{12} = 3, x_{22} = 15,\ x_{23} = 9,\ x_{33} = 9,\ x_{34} = 25$

   Total Cost =580

6. $x_{11} = 30,\ x_{21} = 5, x_{22} = 28,\ x_{23} = 7,\ x_{33} = 25,\ x_{34} = 25$

   Total Cost =1076

7. $x_{11} = 15,\ x_{13} = 10, x_{23} = 35,\ x_{31} = 15,\ x_{32} = 25,$

   Total Cost = 580

8. $x_{11} = 1,\ x_{12} = 5, x_{24} = 1,\ x_{31} = 6,\ x_{33} = 3,\ x_{34} = 1,$

   Total Cost =102

9. $x_{11} = 10,\ x_{13} = 20, x_{21} = 20,\ x_{22} = 20,\ x_{24} = 10,\ x_{32} = 20$

   Total Cost = 370

10. $x_{11} = 3,\ x_{12} = 1, x_{22} = 2,\ x_{23} = 4,\ x_{24} = 2,\ x_{34} = 3, x_{35} = 6$

   Total Cost = 153

11. (i) $x_{11} = 7,\ x_{21} = 3, x_{22} = 9,\ x_{32} = 1,\ x_{33} = 10,$

   Total Cost = 94

(ii) $x_{13} = 7, x_{21} = 10, x_{23} = 2, x_{32} = 10, x_{33} = 1,$

Total Cost =61

(iii) $x_{11} = 7, x_{21} = 2, x_{23} = 10, x_{31} = 1, x_{32} = 10,$

Total Cost 40

12. $x_{11} = 200, x_{21} = 50, x_{22} = 175, x_{23} = 125, x_{32} = 150, x_{33} = 250$

Total Cost 12,200

Exercise 10.2

4. 46 5. 280 6. 41 Hours 7. 37 8. 12

Exercise 10.3

1. (i) S_1 (ii) S_2 2. (a) Crop C (b) Crop B and Crop C

3. (i) Egg shampoo (ii) Egg Shampoo 4. (i) A_1 and A_3 (ii) A_2 and A_3

Exercise 10.4

1	2	3	4	5	6	7	8	9	10	11	12	13	14
(a)	(a)	(c)	(c)	(a)	(a)	(c)	(c)	(d)	(b)	(d)	(b)	(d)	(d)

Miscellaneous Problems

1. $x_{11} = 5\ x_{21} = 2, x_{22} = 6, x_{32} = 3, x_{33} = 4, x_{43} = 14$

Total Cost =102

2. (a) $x_{12} = 10\ x_{13} = 20, x_{21} = 30, x_{22} = 20, x_{24} = 10, x_{32} = 20$

(b) $x_{11} = 10\ x_{13} = 20, x_{21} = 20, x_{22} = 20, x_{24} = 10, x_{32} = 20$

Total cost =370

3. $x_{11} = 15\ x_{13} = 10, x_{23} = 35, x_{31} = 15, x_{32} = 25, x_{32} = 20$

Total Cost =560

4. $x_{12} = 1\ x_{12} = 5, x_{24} = 1, x_{31} = 6, x_{33} = 3, x_{34} = 1$

Total Cost =102

5. $A \to e, B \to c, C \to b, D \to a, E \to d$

Minimum Distance =570 miles

6. $1 \to 11, 2 \to 8, 3 \to 7, 4 \to 9, 5 \to 10, 6 \to 12$

Minimum distance =125 kms

7. (i) Debenture : 6000

(ii) Stocks : 1000

LOGARITHM TABLE

											Mean Difference								
	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
1.0	0.0000	0.0043	0.0086	0.0128	0.0170	0.0212	0.0253	0.0294	0.0334	0.0374	4	8	12	17	21	25	29	33	37
1.1	0.0414	0.0453	0.0492	0.0531	0.0569	0.0607	0.0645	0.0682	0.0719	0.0755	4	8	11	15	19	23	26	30	34
1.2	0.0792	0.0828	0.0864	0.0899	0.0934	0.0969	0.1004	0.1038	0.1072	0.1106	3	7	10	14	17	21	24	28	31
1.3	0.1139	0.1173	0.1206	0.1239	0.1271	0.1303	0.1335	0.1367	0.1399	0.1430	3	6	10	13	16	19	23	26	29
1.4	0.1461	0.1492	0.1523	0.1553	0.1584	0.1614	0.1644	0.1673	0.1703	0.1732	3	6	9	12	15	18	21	24	27
1.5	0.1761	0.1790	0.1818	0.1847	0.1875	0.1903	0.1931	0.1959	0.1987	0.2014	3	6	8	11	14	17	20	22	25
1.6	0.2041	0.2068	0.2095	0.2122	0.2148	0.2175	0.2201	0.2227	0.2253	0.2279	3	5	8	11	13	16	18	21	24
1.7	0.2304	0.2330	0.2355	0.2380	0.2405	0.2430	0.2455	0.2480	0.2504	0.2529	2	5	7	10	12	15	17	20	22
1.8	0.2553	0.2577	0.2601	0.2625	0.2648	0.2672	0.2695	0.2718	0.2742	0.2765	2	5	7	9	12	14	16	19	21
1.9	0.2788	0.2810	0.2833	0.2856	0.2878	0.2900	0.2923	0.2945	0.2967	0.2989	2	4	7	9	11	13	16	18	20
2.0	0.3010	0.3032	0.3054	0.3075	0.3096	0.3118	0.3139	0.3160	0.3181	0.3201	2	4	6	8	11	13	15	17	19
2.1	0.3222	0.3243	0.3263	0.3284	0.3304	0.3324	0.3345	0.3365	0.3385	0.3404	2	4	6	8	10	12	14	16	18
2.2	0.3424	0.3444	0.3464	0.3483	0.3502	0.3522	0.3541	0.3560	0.3579	0.3598	2	4	6	8	10	12	14	15	17
2.3	0.3617	0.3636	0.3655	0.3674	0.3692	0.3711	0.3729	0.3747	0.3766	0.3784	2	4	6	7	9	11	13	15	17
2.4	0.3802	0.3820	0.3838	0.3856	0.3874	0.3892	0.3909	0.3927	0.3945	0.3962	2	4	5	7	9	11	12	14	16
2.5	0.3979	0.3997	0.4014	0.4031	0.4048	0.4065	0.4082	0.4099	0.4116	0.4133	2	3	5	7	9	10	12	14	15
2.6	0.4150	0.4166	0.4183	0.4200	0.4216	0.4232	0.4249	0.4265	0.4281	0.4298	2	3	5	7	8	10	11	13	15
2.7	0.4314	0.4330	0.4346	0.4362	0.4378	0.4393	0.4409	0.4425	0.4440	0.4456	2	3	5	6	8	9	11	13	14
2.8	0.4472	0.4487	0.4502	0.4518	0.4533	0.4548	0.4564	0.4579	0.4594	0.4609	2	3	5	6	8	9	11	12	14
2.9	0.4624	0.4639	0.4654	0.4669	0.4683	0.4698	0.4713	0.4728	0.4742	0.4757	1	3	4	6	7	9	10	12	13
3.0	0.4771	0.4786	0.4800	0.4814	0.4829	0.4843	0.4857	0.4871	0.4886	0.4900	1	3	4	6	7	9	10	11	13
3.1	0.4914	0.4928	0.4942	0.4955	0.4969	0.4983	0.4997	0.5011	0.5024	0.5038	1	3	4	6	7	8	10	11	12
3.2	0.5051	0.5065	0.5079	0.5092	0.5105	0.5119	0.5132	0.5145	0.5159	0.5172	1	3	4	5	7	8	9	11	12
3.3	0.5185	0.5198	0.5211	0.5224	0.5237	0.5250	0.5263	0.5276	0.5289	0.5302	1	3	4	5	6	8	9	10	12
3.4	0.5315	0.5328	0.5340	0.5353	0.5366	0.5378	0.5391	0.5403	0.5416	0.5428	1	3	4	5	6	8	9	10	11
3.5	0.5441	0.5453	0.5465	0.5478	0.5490	0.5502	0.5514	0.5527	0.5539	0.5551	1	2	4	5	6	7	9	10	11
3.6	0.5563	0.5575	0.5587	0.5599	0.5611	0.5623	0.5635	0.5647	0.5658	0.5670	1	2	4	5	6	7	8	10	11
3.7	0.5682	0.5694	0.5705	0.5717	0.5729	0.5740	0.5752	0.5763	0.5775	0.5786	1	2	3	5	6	7	8	9	10
3.8	0.5798	0.5809	0.5821	0.5832	0.5843	0.5855	0.5866	0.5877	0.5888	0.5899	1	2	3	5	6	7	8	9	10
3.9	0.5911	0.5922	0.5933	0.5944	0.5955	0.5966	0.5977	0.5988	0.5999	0.6010	1	2	3	4	5	7	8	9	10
4.0	0.6021	0.6031	0.6042	0.6053	0.6064	0.6075	0.6085	0.6096	0.6107	0.6117	1	2	3	4	5	6	8	9	10
4.1	0.6128	0.6138	0.6149	0.6160	0.6170	0.6180	0.6191	0.6201	0.6212	0.6222	1	2	3	4	5	6	7	8	9
4.2	0.6232	0.6243	0.6253	0.6263	0.6274	0.6284	0.6294	0.6304	0.6314	0.6325	1	2	3	4	5	6	7	8	9
4.3	0.6335	0.6345	0.6355	0.6365	0.6375	0.6385	0.6395	0.6405	0.6415	0.6425	1	2	3	4	5	6	7	8	9
4.4	0.6435	0.6444	0.6454	0.6464	0.6474	0.6484	0.6493	0.6503	0.6513	0.6522	1	2	3	4	5	6	7	8	9
4.5	0.6532	0.6542	0.6551	0.6561	0.6571	0.6580	0.6590	0.6599	0.6609	0.6618	1	2	3	4	5	6	7	8	9
4.6	0.6628	0.6637	0.6646	0.6656	0.6665	0.6675	0.6684	0.6693	0.6702	0.6712	1	2	3	4	5	6	7	7	8
4.7	0.6721	0.6730	0.6739	0.6749	0.6758	0.6767	0.6776	0.6785	0.6794	0.6803	1	2	3	4	5	5	6	7	8
4.8	0.6812	0.6821	0.6830	0.6839	0.6848	0.6857	0.6866	0.6875	0.6884	0.6893	1	2	3	4	4	5	6	7	8
4.9	0.6902	0.6911	0.6920	0.6928	0.6937	0.6946	0.6955	0.6964	0.6972	0.6981	1	2	3	4	4	5	6	7	8
5.0	0.6990	0.6998	0.7007	0.7016	0.7024	0.7033	0.7042	0.7050	0.7059	0.7067	1	2	3	3	4	5	6	7	8
5.1	0.7076	0.7084	0.7093	0.7101	0.7110	0.7118	0.7126	0.7135	0.7143	0.7152	1	2	3	3	4	5	6	7	8
5.2	0.7160	0.7168	0.7177	0.7185	0.7193	0.7202	0.7210	0.7218	0.7226	0.7235	1	2	2	3	4	5	6	7	7
5.3	0.7243	0.7251	0.7259	0.7267	0.7275	0.7284	0.7292	0.7300	0.7308	0.7316	1	2	2	3	4	5	6	6	7
5.4	0.7324	0.7332	0.7340	0.7348	0.7356	0.7364	0.7372	0.7380	0.7388	0.7396	1	2	2	3	4	5	6	6	7

LOGARITHM TABLE

											Mean Difference								
	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
5.5	0.7404	0.7412	0.7419	0.7427	0.7435	0.7443	0.7451	0.7459	0.7466	0.7474	1	2	2	3	4	5	5	6	7
5.6	0.7482	0.7490	0.7497	0.7505	0.7513	0.7520	0.7528	0.7536	0.7543	0.7551	1	2	2	3	4	5	5	6	7
5.7	0.7559	0.7566	0.7574	0.7582	0.7589	0.7597	0.7604	0.7612	0.7619	0.7627	1	2	2	3	4	5	5	6	7
5.8	0.7634	0.7642	0.7649	0.7657	0.7664	0.7672	0.7679	0.7686	0.7694	0.7701	1	1	2	3	4	4	5	6	7
5.9	0.7709	0.7716	0.7723	0.7731	0.7738	0.7745	0.7752	0.7760	0.7767	0.7774	1	1	2	3	4	4	5	6	7
6.0	0.7782	0.7789	0.7796	0.7803	0.7810	0.7818	0.7825	0.7832	0.7839	0.7846	1	1	2	3	4	4	5	6	6
6.1	0.7853	0.7860	0.7868	0.7875	0.7882	0.7889	0.7896	0.7903	0.7910	0.7917	1	1	2	3	4	4	5	6	6
6.2	0.7924	0.7931	0.7938	0.7945	0.7952	0.7959	0.7966	0.7973	0.7980	0.7987	1	1	2	3	3	4	5	6	6
6.3	0.7993	0.8000	0.8007	0.8014	0.8021	0.8028	0.8035	0.8041	0.8048	0.8055	1	1	2	3	3	4	5	5	6
6.4	0.8062	0.8069	0.8075	0.8082	0.8089	0.8096	0.8102	0.8109	0.8116	0.8122	1	1	2	3	3	4	5	5	6
6.5	0.8129	0.8136	0.8142	0.8149	0.8156	0.8162	0.8169	0.8176	0.8182	0.8189	1	1	2	3	3	4	5	5	6
6.6	0.8195	0.8202	0.8209	0.8215	0.8222	0.8228	0.8235	0.8241	0.8248	0.8254	1	1	2	3	3	4	5	5	6
6.7	0.8261	0.8267	0.8274	0.8280	0.8287	0.8293	0.8299	0.8306	0.8312	0.8319	1	1	2	3	3	4	5	5	6
6.8	0.8325	0.8331	0.8338	0.8344	0.8351	0.8357	0.8363	0.8370	0.8376	0.8382	1	1	2	3	3	4	4	5	6
6.9	0.8388	0.8395	0.8401	0.8407	0.8414	0.8420	0.8426	0.8432	0.8439	0.8445	1	1	2	2	3	4	4	5	6
7.0	0.8451	0.8457	0.8463	0.8470	0.8476	0.8482	0.8488	0.8494	0.8500	0.8506	1	1	2	2	3	4	4	5	6
7.1	0.8513	0.8519	0.8525	0.8531	0.8537	0.8543	0.8549	0.8555	0.8561	0.8567	1	1	2	2	3	4	4	5	5
7.2	0.8573	0.8579	0.8585	0.8591	0.8597	0.8603	0.8609	0.8615	0.8621	0.8627	1	1	2	2	3	4	4	5	5
7.3	0.8633	0.8639	0.8645	0.8651	0.8657	0.8663	0.8669	0.8675	0.8681	0.8686	1	1	2	2	3	4	4	5	5
7.4	0.8692	0.8698	0.8704	0.8710	0.8716	0.8722	0.8727	0.8733	0.8739	0.8745	1	1	2	2	3	4	4	5	5
7.5	0.8751	0.8756	0.8762	0.8768	0.8774	0.8779	0.8785	0.8791	0.8797	0.8802	1	1	2	2	3	3	4	5	5
7.6	0.8808	0.8814	0.8820	0.8825	0.8831	0.8837	0.8842	0.8848	0.8854	0.8859	1	1	2	2	3	3	4	5	5
7.7	0.8865	0.8871	0.8876	0.8882	0.8887	0.8893	0.8899	0.8904	0.8910	0.8915	1	1	2	2	3	3	4	4	5
7.8	0.8921	0.8927	0.8932	0.8938	0.8943	0.8949	0.8954	0.8960	0.8965	0.8971	1	1	2	2	3	3	4	4	5
7.9	0.8976	0.8982	0.8987	0.8993	0.8998	0.9004	0.9009	0.9015	0.9020	0.9025	1	1	2	2	3	3	4	4	5
8.0	0.9031	0.9036	0.9042	0.9047	0.9053	0.9058	0.9063	0.9069	0.9074	0.9079	1	1	2	2	3	3	4	4	5
8.1	0.9085	0.9090	0.9096	0.9101	0.9106	0.9112	0.9117	0.9122	0.9128	0.9133	1	1	2	2	3	3	4	4	5
8.2	0.9138	0.9143	0.9149	0.9154	0.9159	0.9165	0.9170	0.9175	0.9180	0.9186	1	1	2	2	3	3	4	4	5
8.3	0.9191	0.9196	0.9201	0.9206	0.9212	0.9217	0.9222	0.9227	0.9232	0.9238	1	1	2	2	3	3	4	4	5
8.4	0.9243	0.9248	0.9253	0.9258	0.9263	0.9269	0.9274	0.9279	0.9284	0.9289	1	1	2	2	3	3	4	4	5
8.5	0.9294	0.9299	0.9304	0.9309	0.9315	0.9320	0.9325	0.9330	0.9335	0.9340	1	1	2	2	3	3	4	4	5
8.6	0.9345	0.9350	0.9355	0.9360	0.9365	0.9370	0.9375	0.9380	0.9385	0.9390	1	1	2	2	3	3	4	4	5
8.7	0.9395	0.9400	0.9405	0.9410	0.9415	0.9420	0.9425	0.9430	0.9435	0.9440	0	1	1	2	2	3	3	4	4
8.8	0.9445	0.9450	0.9455	0.9460	0.9465	0.9469	0.9474	0.9479	0.9484	0.9489	0	1	1	2	2	3	3	4	4
8.9	0.9494	0.9499	0.9504	0.9509	0.9513	0.9518	0.9523	0.9528	0.9533	0.9538	0	1	1	2	2	3	3	4	4
9.0	0.9542	0.9547	0.9552	0.9557	0.9562	0.9566	0.9571	0.9576	0.9581	0.9586	0	1	1	2	2	3	3	4	4
9.1	0.9590	0.9595	0.9600	0.9605	0.9609	0.9614	0.9619	0.9624	0.9628	0.9633	0	1	1	2	2	3	3	4	4
9.2	0.9638	0.9643	0.9647	0.9652	0.9657	0.9661	0.9666	0.9671	0.9675	0.9680	0	1	1	2	2	3	3	4	4
9.3	0.9685	0.9689	0.9694	0.9699	0.9703	0.9708	0.9713	0.9717	0.9722	0.9727	0	1	1	2	2	3	3	4	4
9.4	0.9731	0.9736	0.9741	0.9745	0.9750	0.9754	0.9759	0.9763	0.9768	0.9773	0	1	1	2	2	3	3	4	4
9.5	0.9777	0.9782	0.9786	0.9791	0.9795	0.9800	0.9805	0.9809	0.9814	0.9818	0	1	1	2	2	3	3	4	4
9.6	0.9823	0.9827	0.9832	0.9836	0.9841	0.9845	0.9850	0.9854	0.9859	0.9863	0	1	1	2	2	3	3	4	4
9.7	0.9868	0.9872	0.9877	0.9881	0.9886	0.9890	0.9894	0.9899	0.9903	0.9908	0	1	1	2	2	3	3	4	4
9.8	0.9912	0.9917	0.9921	0.9926	0.9930	0.9934	0.9939	0.9943	0.9948	0.9952	0	1	1	2	2	3	3	4	4
9.9	0.9956	0.9961	0.9965	0.9969	0.9974	0.9978	0.9983	0.9987	0.9991	0.9996	0	1	1	2	2	3	3	3	4

ANTI LOGARITHM TABLE

											Mean Difference								
	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
0.00	1.000	1.002	1.005	1.007	1.009	1.012	1.014	1.016	1.019	1.021	0	0	1	1	1	1	2	2	2
0.01	1.023	1.026	1.028	1.030	1.033	1.035	1.038	1.040	1.042	1.045	0	0	1	1	1	1	2	2	2
0.02	1.047	1.050	1.052	1.054	1.057	1.059	1.062	1.064	1.067	1.069	0	0	1	1	1	1	2	2	2
0.03	1.072	1.074	1.076	1.079	1.081	1.084	1.086	1.089	1.091	1.094	0	0	1	1	1	1	2	2	2
0.04	1.096	1.099	1.102	1.104	1.107	1.109	1.112	1.114	1.117	1.119	0	1	1	1	1	2	2	2	2
0.05	1.122	1.125	1.127	1.130	1.132	1.135	1.138	1.140	1.143	1.146	0	1	1	1	1	2	2	2	2
0.06	1.148	1.151	1.153	1.156	1.159	1.161	1.164	1.167	1.169	1.172	0	1	1	1	1	2	2	2	2
0.07	1.175	1.178	1.180	1.183	1.186	1.189	1.191	1.194	1.197	1.199	0	1	1	1	1	2	2	2	2
0.08	1.202	1.205	1.208	1.211	1.213	1.216	1.219	1.222	1.225	1.227	0	1	1	1	1	2	2	2	3
0.09	1.230	1.233	1.236	1.239	1.242	1.245	1.247	1.250	1.253	1.256	0	1	1	1	1	2	2	2	3
0.10	1.259	1.262	1.265	1.268	1.271	1.274	1.276	1.279	1.282	1.285	0	1	1	1	1	2	2	2	3
0.11	1.288	1.291	1.294	1.297	1.300	1.303	1.306	1.309	1.312	1.315	0	1	1	1	2	2	2	2	3
0.12	1.318	1.321	1.324	1.327	1.330	1.334	1.337	1.340	1.343	1.346	0	1	1	1	2	2	2	2	3
0.13	1.349	1.352	1.355	1.358	1.361	1.365	1.368	1.371	1.374	1.377	0	1	1	1	2	2	2	3	3
0.14	1.380	1.384	1.387	1.390	1.393	1.396	1.400	1.403	1.406	1.409	0	1	1	1	2	2	2	3	3
0.15	1.413	1.416	1.419	1.422	1.426	1.429	1.432	1.435	1.439	1.442	0	1	1	1	2	2	2	3	3
0.16	1.445	1.449	1.452	1.455	1.459	1.462	1.466	1.469	1.472	1.476	0	1	1	1	2	2	2	3	3
0.17	1.479	1.483	1.486	1.489	1.493	1.496	1.500	1.503	1.507	1.510	0	1	1	1	2	2	2	3	3
0.18	1.514	1.517	1.521	1.524	1.528	1.531	1.535	1.538	1.542	1.545	0	1	1	1	2	2	2	3	3
0.19	1.549	1.552	1.556	1.560	1.563	1.567	1.570	1.574	1.578	1.581	0	1	1	1	2	2	3	3	3
0.20	1.585	1.589	1.592	1.596	1.600	1.603	1.607	1.611	1.614	1.618	0	1	1	1	2	2	3	3	3
0.21	1.622	1.626	1.629	1.633	1.637	1.641	1.644	1.648	1.652	1.656	0	1	1	2	2	2	3	3	3
0.22	1.660	1.663	1.667	1.671	1.675	1.679	1.683	1.687	1.690	1.694	0	1	1	2	2	2	3	3	3
0.23	1.698	1.702	1.706	1.710	1.714	1.718	1.722	1.726	1.730	1.734	0	1	1	2	2	2	3	3	4
0.24	1.738	1.742	1.746	1.750	1.754	1.758	1.762	1.766	1.770	1.774	0	1	1	2	2	2	3	3	4
0.25	1.778	1.782	1.786	1.791	1.795	1.799	1.803	1.807	1.811	1.816	0	1	1	2	2	2	3	3	4
0.26	1.820	1.824	1.828	1.832	1.837	1.841	1.845	1.849	1.854	1.858	0	1	1	2	2	3	3	3	4
0.27	1.862	1.866	1.871	1.875	1.879	1.884	1.888	1.892	1.897	1.901	0	1	1	2	2	3	3	3	4
0.28	1.905	1.910	1.914	1.919	1.923	1.928	1.932	1.936	1.941	1.945	0	1	1	2	2	3	3	4	4
0.29	1.950	1.954	1.959	1.963	1.968	1.972	1.977	1.982	1.986	1.991	0	1	1	2	2	3	3	4	4
0.30	1.995	2.000	2.004	2.009	2.014	2.018	2.023	2.028	2.032	2.037	0	1	1	2	2	3	3	4	4
0.31	2.042	2.046	2.051	2.056	2.061	2.065	2.070	2.075	2.080	2.084	0	1	1	2	2	3	3	4	4
0.32	2.089	2.094	2.099	2.104	2.109	2.113	2.118	2.123	2.128	2.133	0	1	1	2	2	3	3	4	4
0.33	2.138	2.143	2.148	2.153	2.158	2.163	2.168	2.173	2.178	2.183	0	1	1	2	2	3	3	4	4
0.34	2.188	2.193	2.198	2.203	2.208	2.213	2.218	2.223	2.228	2.234	1	1	2	2	3	3	4	4	5
0.35	2.239	2.244	2.249	2.254	2.259	2.265	2.270	2.275	2.280	2.286	1	1	2	2	3	3	4	4	5
0.36	2.291	2.296	2.301	2.307	2.312	2.317	2.323	2.328	2.333	2.339	1	1	2	2	3	3	4	4	5
0.37	2.344	2.350	2.355	2.360	2.366	2.371	2.377	2.382	2.388	2.393	1	1	2	2	3	3	4	4	5
0.38	2.399	2.404	2.410	2.415	2.421	2.427	2.432	2.438	2.443	2.449	1	1	2	2	3	3	4	4	5
0.39	2.455	2.460	2.466	2.472	2.477	2.483	2.489	2.495	2.500	2.506	1	1	2	2	3	3	4	5	5
0.40	2.512	2.518	2.523	2.529	2.535	2.541	2.547	2.553	2.559	2.564	1	1	2	2	3	4	4	5	5
0.41	2.570	2.576	2.582	2.588	2.594	2.600	2.606	2.612	2.618	2.624	1	1	2	2	3	4	4	5	5
0.42	2.630	2.636	2.642	2.649	2.655	2.661	2.667	2.673	2.679	2.685	1	1	2	2	3	4	4	5	6
0.43	2.692	2.698	2.704	2.710	2.716	2.723	2.729	2.735	2.742	2.748	1	1	2	3	3	4	4	5	6
0.44	2.754	2.761	2.767	2.773	2.780	2.786	2.793	2.799	2.805	2.812	1	1	2	3	3	4	4	5	6
0.45	2.818	2.825	2.831	2.838	2.844	2.851	2.858	2.864	2.871	2.877	1	1	2	3	3	4	5	5	6
0.46	2.884	2.891	2.897	2.904	2.911	2.917	2.924	2.931	2.938	2.944	1	1	2	3	3	4	5	5	6
0.47	2.951	2.958	2.965	2.972	2.979	2.985	2.992	2.999	3.006	3.013	1	1	2	3	3	4	5	5	6
0.48	3.020	3.027	3.034	3.041	3.048	3.055	3.062	3.069	3.076	3.083	1	1	2	3	4	4	5	6	6
0.49	3.090	3.097	3.105	3.112	3.119	3.126	3.133	3.141	3.148	3.155	1	1	2	3	4	4	5	6	6

ANTI LOGARITHM TABLE

											Mean Difference								
	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
0.50	3.162	3.170	3.177	3.184	3.192	3.199	3.206	3.214	3.221	3.228	1	1	2	3	4	4	5	6	7
0.51	3.236	3.243	3.251	3.258	3.266	3.273	3.281	3.289	3.296	3.304	1	2	2	3	4	5	5	6	7
0.52	3.311	3.319	3.327	3.334	3.342	3.350	3.357	3.365	3.373	3.381	1	2	2	3	4	5	5	6	7
0.53	3.388	3.396	3.404	3.412	3.420	3.428	3.436	3.443	3.451	3.459	1	2	2	3	4	5	6	6	7
0.54	3.467	3.475	3.483	3.491	3.499	3.508	3.516	3.524	3.532	3.540	1	2	2	3	4	5	6	6	7
0.55	3.548	3.556	3.565	3.573	3.581	3.589	3.597	3.606	3.614	3.622	1	2	2	3	4	5	6	7	7
0.56	3.631	3.639	3.648	3.656	3.664	3.673	3.681	3.690	3.698	3.707	1	2	3	3	4	5	6	7	8
0.57	3.715	3.724	3.733	3.741	3.750	3.758	3.767	3.776	3.784	3.793	1	2	3	3	4	5	6	7	8
0.58	3.802	3.811	3.819	3.828	3.837	3.846	3.855	3.864	3.873	3.882	1	2	3	4	4	5	6	7	8
0.59	3.890	3.899	3.908	3.917	3.926	3.936	3.945	3.954	3.963	3.972	1	2	3	4	5	5	6	7	8
0.60	3.981	3.990	3.999	4.009	4.018	4.027	4.036	4.046	4.055	4.064	1	2	3	4	5	6	6	7	8
0.61	4.074	4.083	4.093	4.102	4.111	4.121	4.130	4.140	4.150	4.159	1	2	3	4	5	6	7	8	9
0.62	4.169	4.178	4.188	4.198	4.207	4.217	4.227	4.236	4.246	4.256	1	2	3	4	5	6	7	8	9
0.63	4.266	4.276	4.285	4.295	4.305	4.315	4.325	4.335	4.345	4.355	1	2	3	4	5	6	7	8	9
0.64	4.365	4.375	4.385	4.395	4.406	4.416	4.426	4.436	4.446	4.457	1	2	3	4	5	6	7	8	9
0.65	4.467	4.477	4.487	4.498	4.508	4.519	4.529	4.539	4.550	4.560	1	2	3	4	5	6	7	8	9
0.66	4.571	4.581	4.592	4.603	4.613	4.624	4.634	4.645	4.656	4.667	1	2	3	4	5	6	7	9	10
0.67	4.677	4.688	4.699	4.710	4.721	4.732	4.742	4.753	4.764	4.775	1	2	3	4	5	7	8	9	10
0.68	4.786	4.797	4.808	4.819	4.831	4.842	4.853	4.864	4.875	4.887	1	2	3	4	6	7	8	9	10
0.69	4.898	4.909	4.920	4.932	4.943	4.955	4.966	4.977	4.989	5.000	1	2	3	5	6	7	8	9	10
0.70	5.012	5.023	5.035	5.047	5.058	5.070	5.082	5.093	5.105	5.117	1	2	4	5	6	7	8	9	11
0.71	5.129	5.140	5.152	5.164	5.176	5.188	5.200	5.212	5.224	5.236	1	2	4	5	6	7	8	10	11
0.72	5.248	5.260	5.272	5.284	5.297	5.309	5.321	5.333	5.346	5.358	1	2	4	5	6	7	9	10	11
0.73	5.370	5.383	5.395	5.408	5.420	5.433	5.445	5.458	5.470	5.483	1	3	4	5	6	8	9	10	11
0.74	5.495	5.508	5.521	5.534	5.546	5.559	5.572	5.585	5.598	5.610	1	3	4	5	6	8	9	10	12
0.75	5.623	5.636	5.649	5.662	5.675	5.689	5.702	5.715	5.728	5.741	1	3	4	5	7	8	9	10	12
0.76	5.754	5.768	5.781	5.794	5.808	5.821	5.834	5.848	5.861	5.875	1	3	4	5	7	8	9	11	12
0.77	5.888	5.902	5.916	5.929	5.943	5.957	5.970	5.984	5.998	6.012	1	3	4	5	7	8	10	11	12
0.78	6.026	6.039	6.053	6.067	6.081	6.095	6.109	6.124	6.138	6.152	1	3	4	6	7	8	10	11	13
0.79	6.166	6.180	6.194	6.209	6.223	6.237	6.252	6.266	6.281	6.295	1	3	4	6	7	9	10	11	13
0.80	6.310	6.324	6.339	6.353	6.368	6.383	6.397	6.412	6.427	6.442	1	3	4	6	7	9	10	12	13
0.81	6.457	6.471	6.486	6.501	6.516	6.531	6.546	6.561	6.577	6.592	2	3	5	6	8	9	11	12	14
0.82	6.607	6.622	6.637	6.653	6.668	6.683	6.699	6.714	6.730	6.745	2	3	5	6	8	9	11	12	14
0.83	6.761	6.776	6.792	6.808	6.823	6.839	6.855	6.871	6.887	6.902	2	3	5	6	8	9	11	13	14
0.84	6.918	6.934	6.950	6.966	6.982	6.998	7.015	7.031	7.047	7.063	2	3	5	6	8	10	11	13	15
0.85	7.079	7.096	7.112	7.129	7.145	7.161	7.178	7.194	7.211	7.228	2	3	5	7	8	10	12	13	15
0.86	7.244	7.261	7.278	7.295	7.311	7.328	7.345	7.362	7.379	7.396	2	3	5	7	8	10	12	13	15
0.87	7.413	7.430	7.447	7.464	7.482	7.499	7.516	7.534	7.551	7.568	2	3	5	7	9	10	12	14	16
0.88	7.586	7.603	7.621	7.638	7.656	7.674	7.691	7.709	7.727	7.745	2	4	5	7	9	11	12	14	16
0.89	7.762	7.780	7.798	7.816	7.834	7.852	7.870	7.889	7.907	7.925	2	4	5	7	9	11	13	14	16
0.90	7.943	7.962	7.980	7.998	8.017	8.035	8.054	8.072	8.091	8.110	2	4	6	7	9	11	13	15	17
0.91	8.128	8.147	8.166	8.185	8.204	8.222	8.241	8.260	8.279	8.299	2	4	6	8	9	11	13	15	17
0.92	8.318	8.337	8.356	8.375	8.395	8.414	8.433	8.453	8.472	8.492	2	4	6	8	10	12	14	15	17
0.93	8.511	8.531	8.551	8.570	8.590	8.610	8.630	8.650	8.670	8.690	2	4	6	8	10	12	14	16	18
0.94	8.710	8.730	8.750	8.770	8.790	8.810	8.831	8.851	8.872	8.892	2	4	6	8	10	12	14	16	18
0.95	8.913	8.933	8.954	8.974	8.995	9.016	9.036	9.057	9.078	9.099	2	4	6	8	10	12	15	17	19
0.96	9.120	9.141	9.162	9.183	9.204	9.226	9.247	9.268	9.290	9.311	2	4	6	8	11	13	15	17	19
0.97	9.333	9.354	9.376	9.397	9.419	9.441	9.462	9.484	9.506	9.528	2	4	7	9	11	13	15	17	20
0.98	9.550	9.572	9.594	9.616	9.638	9.661	9.683	9.705	9.727	9.750	2	4	7	9	11	13	16	18	20
0.99	9.772	9.795	9.817	9.840	9.863	9.886	9.908	9.931	9.954	9.977	2	5	7	9	11	14	16	18	20

EXPONENTIAL FUNCTION TABLE

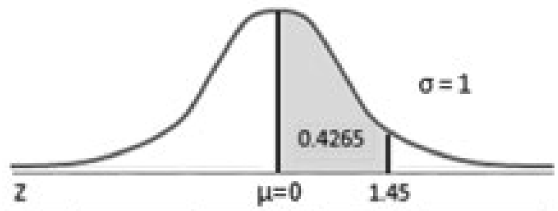
	0	1	2	3	4	5	6	7	8	9
0.00	1.00000000	2.71828183	7.38905610	20.08553692	54.59815003	148.41315910	403.42879349	1096.63315843	2980.95798704	8103.08392758
0.01	1.01005017	2.74560102	7.46331735	20.28739993	55.14687056	149.90473615	407.48332027	1107.65450490	3010.91711288	8184.52127494
0.02	1.02020134	2.77319476	7.53832493	20.49129168	55.70110583	151.41130379	411.57859573	1118.78661775	3041.17733294	8266.77708126
0.03	1.03045453	2.80106583	7.61408636	20.69723259	56.26091125	152.93301270	415.71502938	1130.03061019	3071.74167327	8349.85957218
0.04	1.04081077	2.82921701	7.69060920	20.90524324	56.82634281	154.47001503	419.89303489	1141.38760663	3102.61319033	8433.77705601
0.05	1.05127110	2.85765112	7.76790111	21.11534442	57.39745705	156.02246449	424.11303004	1152.85874278	3133.79497129	8518.53792457
0.06	1.06183655	2.88637099	7.84596981	21.32755716	57.97431108	157.59051632	428.37543686	1164.44516577	3165.29013436	8604.15065402
0.07	1.07250818	2.91537950	7.92482312	21.54190268	58.55696259	159.17432734	432.68068157	1176.14803425	3197.10182908	8690.62380571
0.08	1.08328707	2.94467955	8.00446891	21.75840240	59.14546985	160.77405593	437.02919472	1187.96851851	3229.23323664	8777.96602703
0.09	1.09417428	2.97427407	8.08491516	21.97707798	59.73989170	162.38986205	441.42141115	1199.90780061	3261.68757023	8866.18605226
0.10	1.10517092	3.00416602	8.16616991	22.19795128	60.34028760	164.02190730	445.85777008	1211.96707449	3294.46807528	8955.29270348
0.11	1.11627807	3.03435839	8.24824128	22.42104440	60.94671757	165.67035487	450.33871517	1224.14754609	3327.57802989	9045.29489144
0.12	1.12749685	3.06485420	8.33113749	22.64637964	61.55924226	167.33536962	454.86469450	1236.45043347	3361.02074508	9136.20161642
0.13	1.13882838	3.09565650	8.41486681	22.87397954	62.17792293	169.01711804	459.43616068	1248.87696691	3394.79956514	9228.02196918
0.14	1.15027380	3.12676837	8.49943763	23.10386686	62.80282145	170.71576832	464.05357086	1261.42838910	3428.91786799	9320.76513183
0.15	1.16183424	3.15819291	8.58485840	23.33606458	63.43400030	172.43149032	468.71738678	1274.10595517	3463.37906548	9414.44037876
0.16	1.17351087	3.18993328	8.67113766	23.57059593	64.07152260	174.16445561	473.42807483	1286.91093291	3498.18660376	9509.05707757
0.17	1.18530485	3.22199264	8.75828404	23.80748436	64.71545211	175.91483748	478.18610609	1299.84460280	3533.34396362	9604.62469001
0.18	1.19721736	3.25437420	8.84630626	24.04675355	65.36585321	177.68281099	482.99195635	1312.90825825	3568.85466082	9701.15277293
0.19	1.20924960	3.28708121	8.93521311	24.28842744	66.02279096	179.46855293	487.84610621	1326.10320561	3604.72224646	9798.65097920
0.20	1.22140276	3.32011692	9.02501350	24.53253020	66.68633104	181.27224188	492.74904109	1339.43076439	3640.95030733	9897.12905874
0.21	1.23367806	3.35348465	9.11571639	24.77908622	67.35653981	183.09405819	497.70125129	1352.89226737	3677.54246627	9996.59685944
0.22	1.24607673	3.38718773	9.20733087	25.02812018	68.03348429	184.93418407	502.70323202	1366.48906071	3714.50238251	10097.06432815
0.23	1.25860001	3.42122954	9.29986608	25.27965697	68.71723217	186.79280352	507.75548350	1380.22250409	3751.83375209	10198.54151171
0.24	1.27124915	3.45561346	9.39333129	25.53372175	69.40785184	188.67010241	512.85851094	1394.09397087	3789.54030817	10301.03855791
0.25	1.28402542	3.49034296	9.48773584	25.79033992	70.10541235	190.56626846	518.01282467	1408.10484820	3827.62582144	10404.56571656
0.26	1.29693009	3.52542149	9.58308917	26.04953714	70.80998345	192.48149130	523.21894011	1422.25653720	3866.09410048	10509.13334045
0.27	1.30996445	3.56085256	9.67940081	26.31133934	71.52163562	194.41596245	528.47737788	1436.55045304	3904.94899215	10614.75188643
0.28	1.32312981	3.59663973	9.77668041	26.57577270	72.24044001	196.36987535	533.78866383	1450.98802511	3944.19438198	10721.43191645
0.29	1.33642749	3.63278656	9.87493768	26.84286366	72.96646850	198.34342541	539.15332908	1465.57069720	3983.83419453	10829.18409859
0.30	1.34985881	3.66929667	9.97418245	27.11263892	73.69979370	200.33680997	544.57191013	1480.29992758	4023.87239382	10938.01920817
0.31	1.36342511	3.70617371	10.07442466	27.38512547	74.44048894	202.35022839	550.04494881	1495.17718919	4064.31298371	11047.94812878
0.32	1.37712776	3.74342138	10.17567431	27.66035056	75.18862829	204.38388199	555.57299245	1510.20396976	4105.16000827	11158.98185341
0.33	1.39096813	3.78104339	10.27794153	27.93834170	75.94428657	206.43797416	561.15659385	1525.38177199	4146.41755226	11271.13148552
0.34	1.40494759	3.81904351	10.38123656	28.21912671	76.70753934	208.51271029	566.79631138	1540.71211367	4188.08974147	11384.40824018
0.35	1.41906755	3.85742553	10.48556972	28.50273364	77.47846293	210.60829787	572.49270901	1556.19652784	4230.18074313	11498.82344515
0.36	1.43332941	3.89619330	10.59095145	28.78919088	78.25713442	212.72494645	578.24635639	1571.83656296	4272.69476640	11614.38854204
0.37	1.44773461	3.93535070	10.69739228	29.07852706	79.04363170	214.86286770	584.05782889	1587.63378304	4315.63606270	11731.11508747
0.38	1.46228459	3.97490163	10.80490286	29.37077111	79.83803341	217.02227542	589.92770766	1603.58976783	4359.00892620	11849.01475419
0.39	1.47698079	4.01485005	10.91349394	29.66595227	80.64041898	219.20338555	595.85657969	1619.70611293	4402.81769423	11968.09933225
0.40	1.49182470	4.05519997	11.02317638	29.96410005	81.45086866	221.40641620	601.84503787	1635.98443000	4447.06674770	12088.38073022
0.41	1.50681779	4.09595540	11.13396115	30.26524426	82.26946350	223.63158768	607.89368106	1652.42634686	4491.76051155	12209.87097633
0.42	1.52196156	4.13712044	11.24585931	30.56941502	83.09628536	225.87912250	614.00311413	1669.03350774	4536.90345519	12332.58221972
0.43	1.53725752	4.17869919	11.35888208	30.87664275	83.93141691	228.14924542	620.17394801	1685.80757337	4582.50009296	12456.52673161
0.44	1.55270722	4.22069582	11.47304074	31.18695817	84.77494167	230.44218346	626.40679981	1702.75022115	4628.55498456	12581.71690655
0.45	1.56831219	4.26311452	11.58834672	31.50039231	85.62694400	232.75816591	632.70229281	1719.86314538	4675.07273551	12708.16526367
0.46	1.58407398	4.30595953	11.70481154	31.81697651	86.48750910	235.09742437	639.06105657	1737.14805735	4722.05799763	12835.88444790
0.47	1.59999419	4.34923514	11.82244685	32.13674244	87.35672301	237.46019276	645.48372697	1754.60668558	4769.51546949	12964.88723127
0.48	1.61607440	4.39294568	11.94126442	32.45972208	88.23467268	239.84670737	651.97094627	1772.24077593	4817.44989687	13095.18651418
0.49	1.63231622	4.43709552	12.06127612	32.78594771	89.12144588	242.25720686	658.52336322	1790.05209184	4865.86607325	13226.79532664
0.50	1.64872127	4.48168907	12.18249396	33.11545196	90.01713130	244.69193226	665.14163304	1808.04241446	4914.76884030	13359.72682966

EXPONENTIAL FUNCTION TABLE

	0	1	2	3	4	5	6	7	8	9
0.51	1.66529119	4.52673079	12.30493006	33.44826778	90.92181851	247.15112707	671.82641759	1826.21354282	4964.16308832	13493.99431650
0.52	1.68202765	4.57222520	12.42859666	33.78442846	91.83559798	249.63503719	678.57838534	1844.56729405	5014.05375679	13629.61121401
0.53	1.69893231	4.61817682	12.55350614	34.12396761	92.75856108	252.14391102	685.39821149	1863.10550356	5064.44583482	13766.59108401
0.54	1.71600686	4.66459027	12.67967097	34.46691919	93.69080012	254.67799946	692.28657804	1881.83002516	5115.34436165	13904.94762458
0.55	1.73325302	4.71147018	12.80710378	34.81331749	94.63240831	257.23755591	699.24417382	1900.74273134	5166.75442718	14044.69467150
0.56	1.75067250	4.75882125	12.93581732	35.16319715	95.58347983	259.82283632	706.27169460	1919.84551337	5218.68117245	14185.84619960
0.57	1.76826705	4.80664819	13.06582444	35.51659315	96.54410977	262.43409924	713.36984313	1939.14028156	5271.12979019	14328.41632413
0.58	1.78603843	4.85495581	13.19713816	35.87354085	97.51439421	265.07160579	720.53932925	1958.62896539	5324.10552531	14472.41930224
0.59	1.80398842	4.90374893	13.32977160	36.23407593	98.49443016	267.73561971	727.78086990	1978.31351375	5377.61367541	14617.86953434
0.60	1.82211880	4.95303242	13.46373804	36.59823444	99.48431564	270.42640743	735.09518924	1998.19589510	5431.65959136	14764.78156558
0.61	1.84043140	5.00281123	13.59905085	36.96605281	100.48414964	273.14423800	742.48301872	2018.27809772	5486.24867780	14913.17008727
0.62	1.85892804	5.05309032	13.73572359	37.33756782	101.49403213	275.88938323	749.94509711	2038.56212982	5541.38639368	15063.04993840
0.63	1.87761058	5.10387472	13.87376990	37.71281662	102.51406411	278.66211763	757.48217064	2059.05001984	5597.07825281	15214.43610708
0.64	1.89648088	5.15516951	14.01320361	38.09183673	103.54434758	281.46271848	765.09499302	2079.74381657	5653.32982444	15367.34373205
0.65	1.91554083	5.20697983	14.15403865	38.47466605	104.58498558	284.29146582	772.78432554	2100.64558942	5710.14673375	15521.78810420
0.66	1.93479233	5.25931084	14.29628910	38.86134287	105.63608216	287.14864256	780.55093713	2121.75742858	5767.53466250	15677.78466809
0.67	1.95423732	5.31216780	14.43996919	39.25190586	106.69774243	290.03453439	788.39560446	2143.08144525	5825.49934952	15835.34902351
0.68	1.97387773	5.36555597	14.58509330	39.64639407	107.77007257	292.94942992	796.31911202	2164.61977185	5884.04659134	15994.49692704
0.69	1.99371553	5.41948071	14.73167592	40.04484696	108.85317981	295.89362064	804.32225214	2186.37456223	5943.18224271	16155.24429358
0.70	2.01375271	5.47394739	14.87973172	40.44730436	109.94717245	298.86740097	812.40582517	2208.34799189	6002.91221726	16317.60719802
0.71	2.03399126	5.52896148	15.02927551	40.85380653	111.05215991	301.87106828	820.57063945	2230.54225819	6063.24248804	16481.60187677
0.72	2.05443321	5.58452846	15.18032224	41.26439411	112.16825267	304.90492296	828.81751148	2252.95958057	6124.17908811	16647.24472945
0.73	2.07508061	5.64065391	15.33288702	41.67910816	113.29556235	307.96926838	837.14726595	2275.60220079	6185.72811120	16814.55232047
0.74	2.09593551	5.69734342	15.48698510	42.09799016	114.43420168	311.06441098	845.56073585	2298.47238312	6247.89571226	16983.54138073
0.75	2.11700002	5.75460268	15.64263188	42.52108200	115.58428453	314.19066029	854.05876253	2321.57241461	6310.68810809	17154.22880929
0.76	2.13827622	5.81243739	15.79984295	42.94842598	116.74592590	317.34832892	862.64219579	2344.90460528	6374.11157799	17326.63167502
0.77	2.15976625	5.87085336	15.95863401	43.38006484	117.91924196	320.53773265	871.31189399	2368.47128836	6438.17246436	17500.76721836
0.78	2.18147227	5.92985642	16.11902095	43.81604174	119.10435004	323.75919042	880.06872411	2392.27482054	6502.87717335	17676.65285301
0.79	2.20339643	5.98945247	16.28101980	44.25640028	120.30136866	327.01302438	888.91356183	2416.31758219	6568.23217547	17854.30616767
0.80	2.22554093	6.04964746	16.44464677	44.70118449	121.51041752	330.29955991	897.84729165	2440.60197762	6634.24400628	18033.74492783
0.81	2.24790799	6.11044743	16.60991822	45.15043887	122.73161752	333.61912567	906.87080695	2465.13043529	6700.91926702	18214.98707751
0.82	2.27049984	6.17185845	16.77685067	45.60420832	123.96509078	336.97205363	915.98501008	2489.90540804	6768.26462527	18398.05074107
0.83	2.29331874	6.23388666	16.94546082	46.06253823	125.21096065	340.35867907	925.19081248	2514.92937342	6836.28681562	18582.95422504
0.84	2.31636698	6.29653826	17.11576554	46.52547444	126.46935173	343.77934066	934.48913473	2540.20483383	6904.99264036	18769.71601992
0.85	2.33964685	6.35981952	17.28778184	46.99306323	127.74038985	347.23438048	943.88090667	2565.73431683	6974.38897011	18958.35480204
0.86	2.36316069	6.42373677	17.46152694	47.46535137	129.02420211	350.72414402	953.36706749	2591.52037541	7044.48274457	19148.88943544
0.87	2.38691085	6.48829640	17.63701820	47.94238608	130.32091690	354.24898027	962.94856581	2617.56558819	7115.28097317	19341.33897375
0.88	2.41089971	6.55350486	17.81427318	48.42421507	131.63066389	357.80924171	972.62635979	2643.87255970	7186.79073580	19535.72266207
0.89	2.43512965	6.61936868	17.99330960	48.91088652	132.95357405	361.40528437	982.40141722	2670.44392068	7259.01918349	19732.05993893
0.90	2.45960311	6.68589444	18.17414537	49.40244911	134.28977968	365.03746787	992.27471561	2697.28232827	7331.97353916	19930.37043823
0.91	2.48432253	6.75308880	18.35679857	49.89895197	135.63941441	368.70615541	1002.24724229	2724.39046634	7405.66109828	20130.67399118
0.92	2.50929039	6.82095847	18.54128746	50.40044478	137.00261319	372.41171388	1012.31999453	2751.77104573	7480.08922969	20332.99062831
0.93	2.53450918	6.88951024	18.72763050	50.90697767	138.37951234	376.15451382	1022.49397962	2779.42680452	7555.26537625	20537.34058145
0.94	2.55998142	6.95875097	18.91584631	51.41860130	139.77024956	379.93492954	1032.77021496	2807.36050830	7631.19705565	20743.74428576
0.95	2.58570966	7.02868758	19.10595373	51.93536683	141.17496392	383.75333906	1043.14972818	2835.57495047	7707.89186111	20952.22238178
0.96	2.61169647	7.09932707	19.29797176	52.45732595	142.59379590	387.61012424	1053.63355724	2864.07295251	7785.35746218	21162.79571750
0.97	2.63794446	7.17067649	19.49191960	52.98453084	144.02688737	391.50567075	1064.22275054	2892.85736422	7863.60160548	21375.48535043
0.98	2.66445624	7.24274299	19.68781664	53.51703423	145.47438165	395.44036816	1074.91836700	2921.93106408	7942.63211550	21590.31254971
0.99	2.69123447	7.31553376	19.88568249	54.05488936	146.93642350	399.41460993	1085.72147619	2951.29695948	8022.45689535	21807.29879823

STANDARD NORMAL DISTRIBUTION TABLE

This table provides the area between the mean and some Z score.
For example, when Z score = 1.45 the area = 0.4265.



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Books for Reference

1. Introduction to Matrices, S.P.Gupta , S.Chand & Company
2. Matrices, Shanthi Narayanan,S.Chand & Company
3. Matrices and Determinants,P.N.Arora, S.Chand & Company
4. Stochastic Processes, J. Medhi, New age international publishers
5. A Text Book on differential Calculus – S. K. Goyal -Jai Prakash Nath Publications.
6. A Text Book on Integral Calculus – S. K. Goyal - Jai Prakash Nath Publications.
7. Mathematics for Economics – Mehta, Madnani – Sultan Chand & Sons.
8. Differential and Integral Calculus - N.Piskunov - Mir Publishers,Moscow.
9. Differential and Integral Calculus - Schamum's Outline Series - Frank Ayres.
10. Calculus - S. Narayanan, T. K. Manicavachagon Pillay - S. Viswanathan - Printers and Publishers Pvt. Ltd.
11. Differential Equations and Its Applications - S. Narayanan, T. K. Manicavachagon Pillay
12. Calculus (Volume I & II) - Tom. M. Apostol - John Wiley Publications.
13. Numerical Methods- P.Kandasamy, K.Thilagvathy, K.Gunavathi- S.Chand & Company
14. Finite differences and numerical analysis-H.C.Saxena, S.Chand & Company
15. Applied statistics by A. Chandrasekaran and A. Latha
16. Basic Statistics : B L Agarwal, New Age International Publishers.
17. Business Statistics Problems and solutions: J K Sharma, Vikas Publishing House Pvt Ltd.
18. Comprehensive Statistical Methods : P.N. Arora,Sumeet Arora and S. Arora
19. Elements of statistical methods by P. N. Arora, sumeet arora.
20. Fundamentals of Statistics: S.C.Gupta, Himalaya Publishing House.
21. Fundamentals of Applied Statistics: S.C.Gupta and V.K.Kapoor, Sultan Chand & Sons.
22. Goon, A.M. Gupta M.K. and Das Gupta B (1977) An Outline of Statistical Theory, Vol I, 6/e, World Press, Calcutta.
23. Gupta S.C, Kapoor V.K (2009) Fundamentals of Mathematical Statistics. Sultan Chand & Sons, New Delhi.
24. Handbook of basic c statistical concepts for scientists and pharmacists by shubha rani
25. Hogg. R.V. Craig. A.T. (1978): Introduction to Mathematical Statistics, McGraw Hill Publishing Co. Inc. New York.
26. Introduction to Statistical Quality Control: Douglas C. Montgomery, Wiley Publications.
27. Mood A. M, Graybill F. A, Boes D. C (1983) Introduction to the Theory of Statistics. Third edition, McGraw-Hill International Book Company.
28. Sanjay Arora and Bansilal (1989): New Mathematical Statistics, Satyaprakashan, New Delhi.
29. Statistical Methods: S.P Gupta, Sultan Chand & Sons.
30. Statistical Quality Control: Douglas C. Montgomery, Wiley Publications.
31. Statistical Quality Control: M. Mahajan, Dhanpat Rai & Co Publications
32. Statistics Theory and Practice: R S N Pillai and Bagavathi: S Chand.
33. Operations Research, Dr.S.P.Gupta, P.K. Gupta, Dr.Manmohan, Sultan Chand & Sons
34. Operations Research, A.Ravindran, James J.Solberg, Willey Student Edition
35. Operations Research, Frederick S.Hilton, Gerald J.Lieberman,Mc Graw Hill Education
36. Operations Research –Dr.S.J.Venkatesan, Sri Krishna Publications, Chennai
37. Business Mathematics and Statistics, HSC First Year, Tamilnadu Text Book Corporation.
38. Mathematics, HSC First & Second Year, Tamilnadu Text Book Corporation.
39. Statistics - HSC First & Second Year – Tamil nadu Text Book Corporation.

Business Mathematics and Statistics - Higher Secondary Second Year
Text Book Development Team

Domain Expert

- **Mr.N. RAMESH,**
  Associate Professor (Retd),
  Department of Mathematics,
  Government Arts College(Men), Nandanam,
  Chennai-600 035

Reviewers

- **Dr. M.R. Sreenivasan**
  Professor & HOD,
  Department of Statistics
  University of Madras, Chennai-600 005.
- **Dr. D. Arivudainambi**
  Professor in Mathematics
  Department of Mathematics,
  Anna University, Chennai-600 025

Content Experts

- **Dr. VENU PRAKASH**
  Associate Professor & HOD,
  Department of Statistics
  Presidency College, Chennai-600 005.
- **Dr. R. THIRUMALAISAMY**
  Associate Professor & HOD,
  Department of Mathematics,
  Government Arts College(Men), Nandanam,
  Chennai-600 035
- **Dr. S. J. VENKATESAN**
  Associate Professor,
  Department of Mathematics,
  Government Arts College(Men), Nandanam,
  Chennai-600 035
- **Mrs. K. GOKILA**
  Assistant Professor & HOD
  Department of Statistics
  Dr.Ambedkar Government Arts College
  Vyasarpadi, Chennai
- **Dr. N. SUNDARAM**
  Assistant Professor
  Department of Statistics
  Presidency College, Chennai-600 005.
- **Dr. L.PARIDAYAL**
  Assistant Professor
  Department of Statistics
  MCC, Tambaram,Chennai
- **Mrs. M. THILAGAM**
  Assistant Professor
  Department of Statistics
  Presidency College, Chennai-600 005.

Text Book Group In-Charge

- **Mr.A.RAVI KUMAR**
  Principal, DIET
  Mayanoor, Karur Dist.

Academic Co-ordinator

- **Mr. S.BABU**
  Assistant Professor
  State Council of Educational Research and
  Training, Chennai-600 006

Authors

- **Mr.R.VILLAVANKOTHAI**
  Head Master, GHSS, Korukkai, Nagapattinam District.
- **Mr. T.P. SWAMINATHAN**
  Post Graduate Teacher
  MMA Govt. Hr. Sec. School, Pallavaram, Chennai-43
- **Mr. H. VENKATESH,**
  Post Graduate Teacher,
  Sir Ramaswami Mudaliar Hr. Sec. School,
  Ambattur, Chennai 600 053.
- **Mrs. C.BAGAVATHY**
  Post Graduate Teacher
  Ahobila Math Oriental HSS., West Mambalam,
  Chennai - 600 033.
- **Mr. V.KALAI SELVAN**
  Post Graduate Teacher
  GHSS,Tiruninravur, Thiruvallur Dist
- **Mr. S.F. SULAIMAN**
  Post Graduate Teacher
  MMA Govt. Hr. Sec. School, Pallavaram, Chennai-43
- **Mr. T. RAJA SEKHAR**
  Post Graduate Teacher
  Govt. Boys Hr. Sec. School, Chrompet, Chennai-44
- **Mrs. A.SUGANYA**
  Post Graduate Teacher
  GHSS, Kovilambakkam, Kancheepuram District
- **Mr. V.GANESAN**
  Post Graduate Teacher
  Got. Boys Hr. Sec. School, Nanganallur, Chennai-114
- **Mr. M.K.THRILOGACHANDRAN**
  Post Graduate Teacher, GHSS, Meyyur, Thiruvallur Dist.

ICT Co-ordinator

- **Mr. D. VASURAJ**
  PGT - Maths Teacher & HOD
  KRM Public School, Chennai- 600 011

QR Management Team

- **V. Padmavathi, , B.T. Asst.,**
  G.H.S, Vetriyur, Thirumanur, Ariyalur.
- **S. Albert Valavan Babu, B.T Asst.**
  G.H.S, Perumal Kovil, Paramakudi, Ramanathapuram.
- **A. Devi Jesintha, B.T Asst.**
  G.H.S, N.M.kovil, Vellore.

Layout Designer

- **JOY GRAPHICS,**
  Chindatripet, Chennai-600 002.

In House QC

- **T. Rajesh**
  C. Jerald Wilson
  R. Mathan Raj
  P. Arun Kamaraj

Wrapper Design

- **Kathir Arumugam**

Typing

- **K. Nagavelu.**

This book has been printed on 80 GSM
Elegant Maplitho paper
Printed by offset at :

NOTES

NOTES