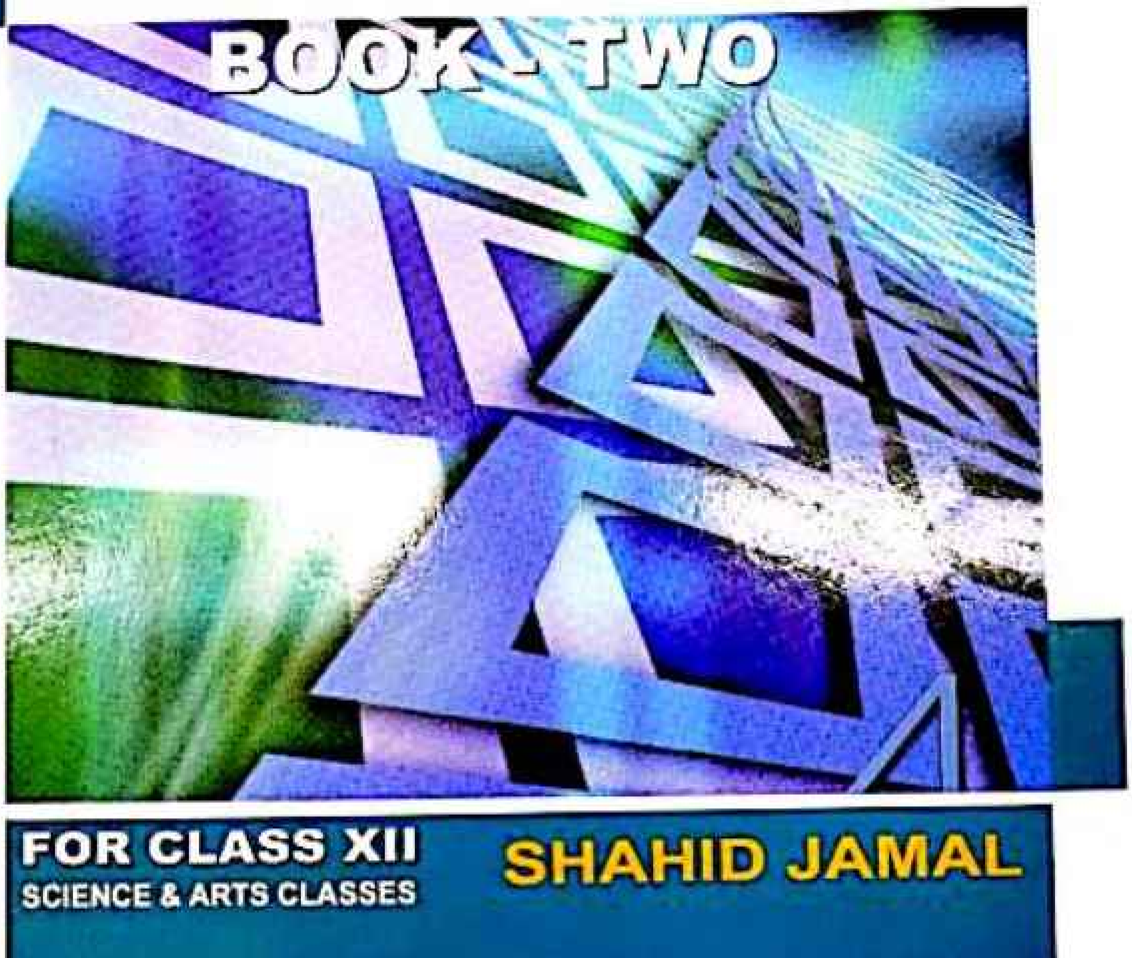
INTRODUCTORY STATISTICS



FOR CLASS XII
SCIENCE & ARTS CLASSES

SHAHID JAMAL

AHMED ACADEMY

CS CamScanner

STATISTICS PAPER – II
(PRACTICAL)

Max. Marks: 15

1. Tossing of Coin and die Throwing experiments in generating a binomial distribution. Calculation of mean and variance of binomial distribution.

2. Problems on the use of Normal Probability Integral Table.

3. Calculations of simple Correlation Coefficient, Regression Coefficients, Regression lines and estimates from these lines, drawing of scatter diagram, Rank correlation.

4. Drawing all possible samples of size 3 from a finite population (without replacement only) computation of sample means, mean and variance of all possible sample means.

5. Drawing of a simple random sample from a given population, computation of sample mean and variance and comparison of these values with Population values.

6. Drawing a stratified random sample from the given 3 strata by selecting 10% observations from each stratum.

7. Computation of χ^2 (Chi-Square) statistic for testing the independence of attributes in contingency table.

CamScanner

CONTENTS

1 REGRESSION AND CORRELATION

2 IMPORTANT TOOLS OF PROBABILITY

CS CamScanner

3 PROBABILITY

CamScanner

4 PROBABILITY DISTRIBUTIONS

5 SAMPLING

CamScanner

6 THEORY OF ATTRIBUTES & TEST OF INDEPENDENCE

CamScanner

Chapter – 1

Regression and Correlation

Regression

The dependence of one variable over the other variable is termed as regression. Regression is a statistical device which helps us in estimating (or predicting) the unknown value of one variable provided the value of other variable is given to us. The variable whose value is to be estimated is called dependent variable whereas the variable whose value is given is called independent variable.

For example, sale is the dependent variable and advertising is the independent variable, weight is the dependent variable and height is the independent variable, demand is the dependent variable and price is the independent variable, etc.

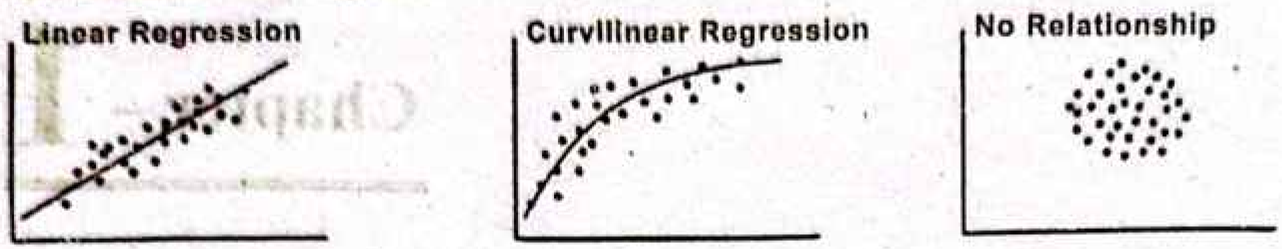
When two variables are found to be strongly related, we may be further interested in obtaining a mathematical equation describing the relationship, such an equation is known as regression equation. To determine the regression equation, a first step is the preparation of a scatter diagram.

Scatter Diagram

The first step is to determine the type of relationship between the two variables is the Scatter Diagram.

According to this method, we first plot the paired vales of the two variables X and Y on a graph paper and do not join the plotted points by any way. If all the plotted points tend to lie near a straight line the relationship is said to be linear (or regression is said to be linear)

CS CamScanner



If the plotted points tend to lie near a curve (not a straight line), the relationship is said to be curvilinear (or regression is said to be curvilinear)

In this text we shell only be concerned with the linear regression (i.e. a straight line relationship).

Regression Equation

If there is a linear relationship between the two variables X and Y, then the equation $Y = a + bX$ is called the regression equation of Y on X where 'a' and 'b' are some constants which determine the line.

The constant 'a' is called the regression constant and the constant 'b' is called the regression coefficient. The regression coefficient 'b' represents the change in Y due to a one unit increase in the value of X.

The values of 'a' and 'b' are computed by the following formulas:

$$b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} \qquad \text{and} \qquad a = \overline{Y} - b\overline{X}$$

Similarly the regression equation of X on Y is $X = c + dY$, where the values of 'c' and 'd' are computed by the following formulas:

$$d = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum Y^2 - (\sum Y)^2} \qquad \text{and} \qquad c = \overline{X} - d\overline{Y}$$

Note:

1) The regression lines of Y on X and X on Y are also called Least Square Lines of regression.

2) The regression line of Y on X (i.e. $Y = a + bX$) is used for estimating Y when a value of X is given.

CamScanner

3) The regression line of X on Y (i.e. X = c + dY) is used for estimating X when a value of Y is given.

4) The regression coefficient 'b' may also be written as b_{yx} and is called regression coefficient of Y on X, the regression coefficient 'd' may also be written as b_{xy} and is called regression coefficient of X on Y.

Example 1.1

For the following data

X	1	2	3	4	5
Y	10	12	15	14	15

(i) Find line of regression of Y on X.

(ii) Estimate Y for X = 8

Solution:

X	Y	X^2	XY
1	10	1	10
2	12	4	24
3	15	9	45
4	14	16	56
5	15	25	75
15	66	55	210

Regression equation of Y on X is Y = a + bX, the values of 'b' and 'a' are computed as

$$b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \frac{(5)(210) - (15)(66)}{(5)(55) - (15)^2} = \frac{60}{50} = 1.2 \quad \text{and}$$

$a = \frac{66}{5} - 1.2\left(\frac{15}{5}\right)$

$$a = \overline{Y} - b\overline{X} \quad \text{where } \overline{Y} = \frac{\sum y}{n} = \frac{66}{5} = 13.2 \text{ and } \overline{X} = \frac{\sum x}{n} = \frac{15}{5} = 3$$

then a = 13.2 − (1.2)(3) = 9.6. Therefore the regression line of y on x is

Y = 9.6 + 1.2X

(ii) Now estimate y by putting x = 8 in the equation of regression line

$\hat{y} = 9.6 + 1.2X$

then $\hat{y} = 9.6 + 12(8) = 19.2$.

Hence for $X = 8$, $\hat{y} = 19.2$

Example 1.2

The following table gives the age of cars of certain make and the annual maintenance costs. Obtain the regression equation for costs related to age. Also predict the maintenance cost, if the age of car is 12 years. X

Y

X	*Age of cars in years*	2	4	6	8
Y	*Maintenance cost in Rs. hundreds*	10	20	25	30

$y = a + bx$

Solution:

Since the maintenance cost depends upon age, therefore we denote age by X and the maintenance cost by Y and fit a regression of Y on X.

The regression equation of Y on X is: $Y = a + bX$

where $b = \dfrac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2}$ and $a = \bar{Y} - b\bar{X}$

then $\Sigma X, \Sigma Y, \Sigma X^2, \Sigma Y^2$ and ΣXY are computed as

X	Y	X^2	Y^2	XY
2	10	4	100	20
4	20	16	400	80
6	25	36	625	150
8	30	64	900	240
20	85	120	2025	490

Now $b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \frac{(4)(490)-(20)(85)}{(4)(120)-(20)^2} = 3.25$ and

$a = \bar{Y} - b\bar{X} = 21.25 - (3.25)(5) = 5$ where $\bar{X} = 5$ and $\bar{Y} = 21.25$

Then, $\hat{Y} = 5 + 3.25X$ is the regression equation of Y on X.

Now predict the value of Y for X = 12

Therefore, $\hat{Y} = 5 + 3.25(12) = 44$ (in rupees hundred)

Therefore at age 12 years, the maintenance cost is 44 (Rs. hundred).

Example 1.3

Obtain Lines of Regression (Y on X and X on Y) for the following data:

X	1	2	3	4	5	6	7	8	9
Y	9	8	10	12	11	13	14	16	15

Solution:

X	Y	X^2	Y^2	XY
1	9	1	81	9
2	8	4	64	16
3	10	9	100	30
4	12	16	144	48
5	11	25	121	55
6	13	36	169	78
7	14	49	196	98
8	16	64	256	128
9	15	81	225	135
45	108	285	1356	597

Regression of Y on X is Y = a + bX

CS CamScanner

where $b = \dfrac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \dfrac{(9)(597)-(45)(108)}{(9)(285)-(45)^2} = 0.95$

and $a = \overline{Y} - b\overline{X} = 12 - (0.95)(5) = 7.25$ where $\overline{X} = \dfrac{\sum x}{n} = \dfrac{45}{9} = 5$

and $\overline{Y} = \dfrac{\sum y}{n} = \dfrac{108}{9} = 12$, Then regression line of y on x is:

$Y = 7.25 + 0.95X$ Y = a + b(x)

Now find Regression of x on y i.e. (X = c + dY)

where $d = \dfrac{n\sum XY - (\sum X)(\sum Y)}{n\sum Y^2 - (\sum Y)^2} = \dfrac{(9)(597)-(45)(108)}{(9)(1356)-(108)^2} = \dfrac{513}{540} = 0.95$

and $c = \overline{X} - d\overline{Y}$ where $\overline{Y} = \dfrac{\sum y}{n} = \dfrac{108}{9} = 12$ and $\overline{X} = \dfrac{\sum x}{n} = \dfrac{45}{9} = 5$

$c = 5 - (0.95)(12) = -6.4$, then the regression line of X on Y is:

$X = -6.4 + 0.95Y$

Example 1.4

Find the regression of yield on fertilizer using Least Square method from the following data

X	*Fertilizer (units)*	0	2	4	6	8	10
Y	*Yield (units)*	110	113	118	119	120	118

Estimate the yield when fertilizer used is 3 units.

Solution:

Let X = amount of fertilizer

Y = amount of yield

X	Y	X^2	XY
0	110	0	0
2	113	4	226
4	118	16	472
6	119	36	714
8	120	64	960
10	118	100	1180
30	698	220	3552

Regression of Y on X is Y = a + bX

where $b = \dfrac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \dfrac{(6)(3552)-(30)(698)}{(6)(220)-(30)^2} = 0.89$

and $a = \bar{Y} - b\bar{X}$ where $\bar{Y} = \dfrac{\sum y}{n} = \dfrac{698}{6} = 116.33$ and $\bar{X} = \dfrac{\sum x}{n} = \dfrac{30}{6} = 5$

$a = 116.33 - (0.89)(5) = 111.333$, then the regression line of Y on X is:

$$Y = 111.88 + 0.89X$$

Now estimate the yield when the amount of fertilizer is 3 units, therefore put X = 3 in regression equation we get

$$\hat{Y} = 111.88 + (0.89)(3) = 114.6$$

Example 1.5

A researcher wants to find out if there is a relationship between the heights of the sons and the heights of their fathers. In other words do tall fathers have tall sons? He took a random sample of 6 fathers and their 6 sons. Their heights in inches is given below in an ordered array

Heights of Fathers	63	65	66	67	67	68
Heights of sons	66	68	65	67	69	70

Predict the height of the son, if the father height is 70 inches.

CamScanner

Solution:

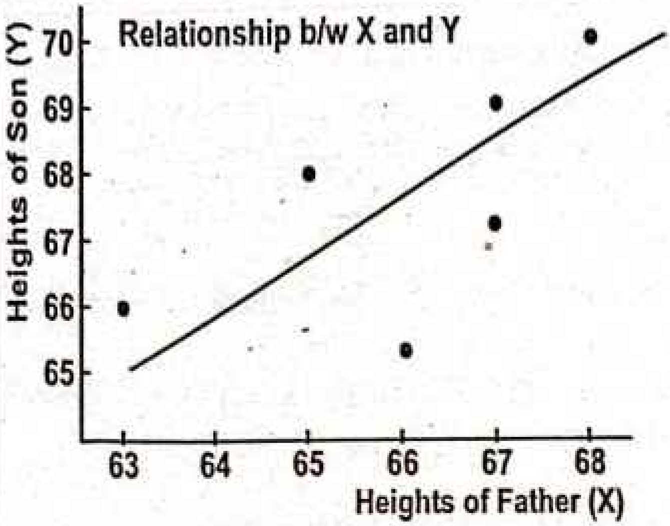
Let X = heights of fathers

Y = heights of sons

We can start with showing the scatter diagram for this data to study the relationship between X and Y.

The scatter diagram shows an increasing linear trend. Now find line of regression of Y on X (i.e Y = a + bX)

X	Y	X^2	XY
63	66	3969	4158
65	68	4225	4420
66	65	4356	4290
67	67	4489	4489
67	69	4489	4623
68	70	4624	4760
396	405	26152	26740



Regression Y on X is Y = a + bX

$$\text{where } b=\frac{n\sum XY-(\sum X)(\sum Y)}{n\sum X^2-(\sum X)^2}=\frac{(6)(26740)-(396)(405)}{(6)(26152)-(396)^2}=0.625$$

$$\text{and} \quad a=\bar{Y}-b\bar{X} \quad \text{where } \bar{Y}=\frac{\sum y}{n}=\frac{405}{6}=67.5 \quad \text{and} \quad \bar{X}=\frac{\sum x}{n}=\frac{396}{6}=66$$

$a = 67.5 - (0.625)(66) = 26.25$ then the regression line of Y on X is:

$$Y = 26.25 + 0.625X$$

If the father's height is 70 inches i.e. if X = 70, then the son's height (i.e. estimated value of Y) would be

$$\hat{Y} = 26.25 + (0.625)(70) = 70 \text{ inches}$$

CS CamScanner

Example 1.6

A company selling household appliances wants to determine if there is any relationship between advertising expenditures and sales. The following data was compiled for 6 major sales regions. The expenditure is in thousands of rupees and the sales are in millions of rupees.

Region	1	2	3	4	5	6
Expenditure (x)	40	45	80	20	15	50
Sales (y)	25	30	45	20	20	40

a) Compute the line of regression.

b) Compute the expected sales for a region where Rs. 72000 is being spent on advertising.

Solution:

X	Y	X^2	XY	Expected values $\hat{Y} = 12.5 + 0.42X$
40	25	1600	1000	29.3
45	30	2025	1350	31.4
80	45	6400	3600	46.1
20	20	400	400	20.9
15	20	225	300	18.8
50	40	2500	2000	33.5
250	180	13150	8650	180

Regression Y on X i.e. $Y = a + bX$ is computed as

$$\text{where } b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \frac{(6)(8650) - (250)(180)}{(6)(13150) - (250)^2} = 0.42$$

$$\text{and} \quad a = \bar{Y} - b\bar{X} \quad \text{where } \bar{Y} = \frac{\sum y}{n} = \frac{180}{6} = 30 \text{ and } \bar{X} = \frac{\sum x}{n} = \frac{250}{6} = 41.67$$

CS CamScanner

$a = 30 - (0.42)(41.67) = 12.50$, then the regression line of Y on X is

$$Y = 12.50 + 0.42X$$

Now, if expenses on advertising is Rs. 72000 then estimate the sales by putting $X = 72$ in the regression equation then

$$\hat{Y} = 12.50 + (0.42)(72) = 42.74 \text{ (millions of rupees)}$$

Correlation

If two sets of variables vary in such a way that the changes of one set are related by the changes in the other then these sets are said to be correlated. For example, there is a relation between income and expenditure, height and weight, rainfall and production, supply and price, etc. Such a relation between any two variables is termed as correlation.

Generally speaking correlation measures the degree of relationship between the two variables.

Nature of Correlation

The nature of correlation, when it exists between two variables may be positive or may be negative.

If with an increase in the values of one variable, the values of the other variable also increases or with a decrease in the values of one variable, the values of the other variable also decreases, i.e. the two series are moving in the same direction both increasing or both decreasing, the correlation will be direct or positive. For example an increase in income is generally associated with an increase in expenditures. Hence there is generally a positive correlation between income and expenditures.

If both the series are moving in opposite directions i.e. an increase in the values of one variable, the values of other variable decreases or vice - versa, the correlation is said to be inverse or negative. For example an increase in supply of a commodity may be associated with a decrease in the price of that commodity and hence the correlation in this case will be negative.

If there is no association between the two variables, the Correlation is said to be zero.

CS CamScanner

Methods of Studying Correlation

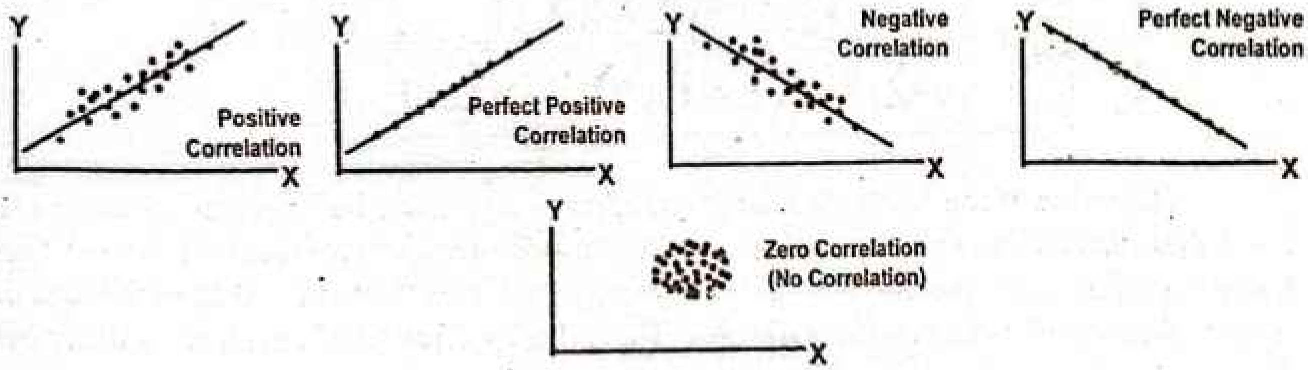
Following are two main methods of studying correlation

(i) Graphical Method (Scatter Diagram)

(ii) Mathematical method (Coefficient of Correlation)

Scatter Diagram

This is the simplest and the easiest method to investigate the nature of correlation between the two variables. According to this method we first plot the paired values of the two variables X and Y on a graph paper. If all the plotted points tend to lie near a straight line, the correlation is said to be linear. If all the points lie exactly on the line, the linear correlation is said to be perfect. If the points tend to lie near an upward sloping line, the linear correlation is said to be positive. If the points tend to lie near a downward sloping line, the linear correlation is said to be negative. If all the points do not show a definite movement in any direction, then there is no relationship between the two variables. The following Scatter Diagram, illustrate the different cases:



This method is mainly used when we are interested in finding out whether there is correlation and only in getting a rough idea about its nature and degree. It does not give us any measure of correlation.

Therefore, Karl Pearson a great biometrician and statistician has suggested a mathematical method or coefficient, known as Karl Pearson's Coefficient of Correlation or simply Coefficient of Correlation.

CamScanner

Coefficient of Correlation

The coefficient of correlation is the numerical measure of strength of the linear relationship between the two variables. The correlation coefficient (sometimes referred to as the Pearson's product moment correlation coefficient) is denoted by r and is defined as

$$r = r_{XY} = \frac{Covariance\,(X,Y)}{S.D.(X)\cdot S.D.(Y)} = \frac{S_{XY}}{S_X \cdot S_Y} \quad \text{where } S_{XY} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{n}$$

$$S_X = \sqrt{\frac{\sum(X-\bar{X})^2}{n}} \quad \text{and} \quad S_Y = \sqrt{\frac{\sum(Y-\bar{Y})^2}{n}}$$

then $$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2}\sqrt{\sum(Y-\bar{Y})^2}}$$

May also be written in the following simple computational form

$$\boxed{r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}}$$

The value of the coefficient of correlation shall always be between -1 and $+1$. If $r = 1$ then there is a perfect positive correlation between the variables. If $r = -1$ then there is perfect negative correlation between the variables, but if $r = 0$ then there is no linear relationship between the variables. Thus, the coefficient of correlation describes the magnitude and direction of correlation.

Note: r^2 is called "Coefficient of Determination".

Strength of Coefficient of Correlation

Coefficient of Correlation	*Degree of Association*
± 0.8 to ± 1	Strong
± 0.5 to ± 0.8	Moderate
± 0.2 to ± 0.5	Weak
0 to ± 0.2	Negligible

CamScanner

Example 1.7

Calculate correlation coefficient between X and Y from the following sample data.

X	1	2	3	4	5	6	7
Y	2	4	5	3	8	6	7

Solution:

Since $r = \dfrac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$

therefore we compute $\Sigma X, \Sigma Y, \Sigma X^2, \Sigma Y^2$ and ΣXY in the following table

X	Y	XY	X^2	Y^2
1	2	2	1	4
2	4	8	4	16
3	5	15	9	25
4	3	12	16	9
5	8	40	25	64
6	6	36	36	36
7	7	49	49	49
28	35	162	140	203

$$r = \frac{(7)(162)-(28)(35)}{\sqrt{(7)(140)-(28)^2}\sqrt{(7)(203)-(35)^2}} = \frac{154}{\sqrt{196}\sqrt{196}} = +0.78$$

Example 1.8

Following data given the marks obtained by 8 students in Accounting (X) and Statistics (Y).

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

Calculate Coefficient of Correlation.

CamScanner

Solution:

X	Y	XY	X^2	Y^2
65	67	4355	4225	4489
66	68	4488	4356	4624
67	65	4355	4489	4225
67	68	4556	4489	4624
68	72	4896	4624	5184
69	72	4968	4761	5184
70	69	4830	4900	4761
72	71	5112	5184	5041
544	552	37560	37028	38132

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}} = \frac{(8)(37560) - (544)(552)}{\sqrt{(8)(37028) - (544)^2}\sqrt{(8)(38132) - (552)^2}}$$

$r = 0.60$

Example 1.9

Given $\sum(x-\bar{x})(y-\bar{y}) = 15$, $\sum(x-\bar{x})^2 = 20$ and $\sum(y-\bar{y})^2 = 14$

Calculate 'r'

Solution:

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2}\sqrt{\sum(y-\bar{y})^2}} = \frac{15}{\sqrt{30}\sqrt{14}} = 0.9$$

Properties of Coefficient of Correlation

1. The coefficient of correlation always varies from -1 to $+1$, i.e. $-1 \leq r \leq 1$
2. The coefficient of correlation is unaffected by change of origin and scale i.e. $r_{xy} = r_{uv}$ where $u = \frac{x-a}{h}$, $v = \frac{y-b}{k}$ and a, b, h, k are some constants.

CS CamScanner

$\sqrt{b_{yx} \cdot b_{xy}}$

3. The coefficient of correlation is the square root of the product of two regression coefficients i.e. $r = \sqrt{b \times d}$ or $r = \sqrt{b_{yx} \cdot b_{xy}}$ and r has the same sign (+ or –) as b_{yx} or b_{xy}

4. The regression coefficients and the regression equations may be expressed as

(i) $b_{yx} = \dfrac{rs_y}{s_x}$ and $b_{xy} = \dfrac{rs_x}{s_y}$

(ii) Regression of y on x is : $y - \bar{y} = \dfrac{rs_y}{s_x}(x - \bar{x})$

Regression of x on y is : $x - \bar{x} = \dfrac{rs_x}{s_y}(y - \bar{y})$

Example 1.10

Calculate Coefficient of Correlation between X and Y for the values given below. Also verify that $r_{xy} = r_{uv}$ if $u = X - 8$ and $v = Y - 10$

X	12	9	8	10	11	13	7
Y	14	8	6	9	11	12	3

$V = Y - 4$ $u = X - 6$ $12 - 6$

Solution:

X	Y	XY	X^2	Y^2
12	14	168	144	196
9	8	72	81	64
8	6	48	64	36
10	9	90	100	81
11	11	121	121	121
13	12	156	169	144
7	13	21	49	9
70	63	676	728	651

$$r=\frac{n\sum XY-(\sum X)(\sum Y)}{\sqrt{n\sum X^2-(\sum X)^2}\sqrt{n\sum Y^2-(\sum Y)^2}}=\frac{(7)(676)-(70)(63)}{\sqrt{(7)(728)-(70)^2}\sqrt{(7)(651)-(63)^2}}=0.95$$

$r_{xy}=0.95$

Now find r_{uv}, where $u = X-8$ and $v = Y-10$ (i.e. by change of origin)

$u=X-8$	$v=Y-10$	uv	u^2	v^2
4	4	16	16	16
1	–2	–2	–1	4
0	–4	0	0	16
2	–1	–2	4	1
3	1	3	9	1
5	2	10	25	4
–1	–7	7	1	49
14	–7	32	56	91

$$r=\frac{n\sum uv-(\sum u)(\sum v)}{\sqrt{n\sum u^2-(\sum u)^2}\sqrt{n\sum v^2-(\sum v)^2}}=\frac{(7)(32)-(14)(-7)}{\sqrt{(7)(56)-(14)^2}\sqrt{(7)(91)-(-7)^2}}=0.95$$

$r_{uv}=0.95$

Hence $r_{xy}=r_{uv}$

Example 1.11

Calculate Karl Pearson's Coefficient of Correlation from the following data and verify that $r_{xy}=r_{uv}$ if $u=\frac{x-6}{2}$ and $v=\frac{y-16}{2}$

(X)	0	4	6	8	10
(Y)	12	14	16	18	16

CamScanner

Solution:

X	Y	XY	X^2	Y^2
0	12	0	0	144
4	14	56	16	196
6	16	96	36	256
8	18	144	64	324
10	16	160	100	256
28	76	456	216	1176

$$r=\frac{n\sum XY-(\sum X)(\sum Y)}{\sqrt{n\sum X^2-(\sum X)^2}\sqrt{n\sum Y^2-(\sum Y)^2}}=\frac{(5)(456)-(28)(76)}{\sqrt{(5)(216)-(28)^2}\sqrt{(5)(1176)-(76)^2}}=0.87$$

$r_{xy}=0.87$

Now find r_{uv} where $u=\dfrac{x-6}{2}$ and $v=\dfrac{y-16}{2}$ (i.e. by change of origin and scale)

$u=\dfrac{X-6}{2}$	$v=\dfrac{Y-16}{2}$	uv	u^2	v^2
–3	–2	6	9	4
–1	–1	1	1	1
0	0	0	0	0
1	1	1	1	1
2	0	0	4	0
–1	–2	8	15	6

$$r=\frac{n\sum uv-(\sum u)(\sum v)}{\sqrt{n\sum u^2-(\sum u)^2}\sqrt{n\sum v^2-(\sum v)^2}}=\frac{(5)(-8)-(-1)(-2)}{\sqrt{(5)(15)-(-1)^2}\sqrt{(5)(6)-(-2)^2}}=0.87$$

$r_{uv}=0.87$

Hence $r_{xy}=r_{uv}$

Example 1.12

Prove that $r_{xy} = r_{uv}$, if $u = \dfrac{x-a}{h}$ and $v = \dfrac{y-b}{k}$ where a, b, h and k are the new origins and units of measurement.

Solution:

$$\text{LHS} = r_{xy} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2}\sqrt{\sum(Y-\bar{Y})^2}}$$

$$u = \frac{x-a}{h} \qquad \text{then } x = hu + a \quad \text{and} \quad \bar{x} = h\bar{u} + a$$

$$v = \frac{y-b}{k} \qquad \text{then } y = kv + b \quad \text{and} \quad \bar{y} = k\bar{v} + b$$

$$\text{LHS} = r_{xy} = \frac{\sum(hu + a - h\bar{u} - a)(kv + b - k\bar{v} - b)}{\sqrt{\sum(hu + a - h\bar{u} - a)^2}\sqrt{\sum(kv + b - k\bar{v} - b)^2}}$$

$$= \frac{hk\sum(u-\bar{u})(v-\bar{v})}{\sqrt{h^2\sum(u-\bar{u})^2}\sqrt{k^2\sum(v-\bar{v})^2}}$$

$$= \frac{\sum(u-\bar{u})(v-\bar{v})}{\sqrt{\sum(u-\bar{u})^2}\sqrt{\sum(v-\bar{v})^2}} = r_{uv} = \text{RHS}$$

Hence LHS = RHS

i.e. $\boxed{r_{xy} = r_{uv}}$

Example 1.13

From the data given below:

(x)	1	5	3	2	1	1	7	3
(y)	6	1	0	0	1	2	1	5

(a) Calculate coefficient of Correlation r_{xy}

(b) Find regression Coefficients Y on X and X on Y i.e b_{yx} and b_{xy} and then verify that

$$r = \sqrt{b_{yx} \cdot b_{xy}}$$

CamScanner

Solution:

X	Y	XY	X^2	Y^2
1	6	6	1	36
5	1	5	25	1
3	0	0	9	0
2	0	0	4	0
1	1	1	1	1
1	2	2	1	4
7	1	7	49	1
3	5	15	9	25
23	16	36	99	68

$$r=\frac{n\sum XY-(\sum X)(\sum Y)}{\sqrt{n\sum X^2-(\sum X)^2}\sqrt{n\sum Y^2-(\sum Y)^2}}=\frac{(8)(36)-(23)(16)}{\sqrt{(8)(99)-(23)^2}\sqrt{(8)(68)-(16)^2}}=-0.29$$

$r_{xy}=-0.29$

(b) Regression Coefficient of Y on X

$$b_{yx}=\frac{n\sum XY-(\sum X)(\sum Y)}{n\sum X^2-(\sum X)^2}=\frac{(8)(36)-(23)(16)}{(8)(99)-(23)^2}=-0.304$$

Regression Coefficient of X on Y

$$b_{xy}=\frac{n\sum XY-(\sum X)(\sum Y)}{n\sum Y^2-(\sum Y)^2}=\frac{(8)(36)-(23)(16)}{(8)(68)-(16)^2}=-0.278$$

Now $\sqrt{b_{yx}\cdot b_{xy}}=\sqrt{(-0.304)(-0.278)}=-0.29$

Hence $r_{xy}=\sqrt{b_{yx}\cdot b_{xy}}$

Example 1.14

The regression equations of two variables are $3Y-2X-10=0$ and $2Y-X-50=0$. Find coefficient of correlation.

Solution: (2.)

$2Y = 50 + x$

$Y = \frac{50}{2} + \frac{x}{2}$

We assume that

$2Y - X = 50$ is the regression equation of Y on X and

$3Y - 2X = 10$ is the regression equation of X on Y.

$2X = 3Y - 10$

$X = \frac{3}{2}Y - \frac{10}{2}$

Then regression Y on X is $Y = \frac{X}{2} + \frac{50}{2}$ where $b_{YX} = \frac{1}{2}$ and

the regression equation of X on Y is $X = \frac{3}{2}Y - \frac{10}{2}$ where $b_{XY} = \frac{3}{2}$

Since $r = \sqrt{b_{yx} \cdot b_{xy}}$ therefore $r = \sqrt{\frac{1}{2} \cdot \frac{3}{2}} = \sqrt{\frac{3}{4}} = 0.87$

Example 1.15

Prove that $r = \sqrt{b_{yx} \cdot b_{xy}}$

Solution:

Since $b_{yx} = \frac{rs_y}{s_x}$ and $b_{xy} = \frac{rs_x}{s_y}$

then $r_{xy} = \sqrt{\frac{rs_y}{s_x} \cdot \frac{rs_x}{s_y}} = r$

Rank Correlation

When there is no unit of measurement in the magnitudes of two variables or in other words, when the direct measurements of the variables are not possible. They are then ranked in order of the quality possessed. The correlation between two such sets of rankings is called Rank Correlation. The measure of correlation computed by this method is called the Spearman's coefficient of rank correlation and is designated by r_s. The formula for finding r_s is as under:

$$\boxed{r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}}$$

CamScanner

where d is the difference of ranks and n is the number of paired observations and also $-1 \le r_s \le +1$

Example 1.16

Eight competitors in a voice test are ranked by two judges in the following order

Competitors	A	B	C	D	E	F	G	H
Ranking Judge-I X	5	2	8	1	4	6	3	7
Ranking Judge-II Y	4	5	7	3	2	8	1	6

Find Spearman's Coefficient of Rank Correlation between rankings of two judges.

Solution:

Let, Rankings Judge – I = X and Rankings Judge – II = Y then $\boxed{d = X - Y}$

X	Y	$d = X - Y$	d^2
5	4	+1	1
2	5	−3	9
8	7	+1	1
1	3	−2	4
4	2	+2	4
6	8	−2	4
3	1	+2	4
7	6	+1	1
			28

Therefore, $\sum d^2 = 28$

Now $r_s = 1 - \dfrac{6\sum d^2}{n(n^2 - 1)} = 1 - \dfrac{(6)(28)}{8(64-1)} = 0.67$

$r_s = 0.67$

EXERCISE – 1

1.1 Determine the regression equation of Y on X

x	1	1	5	5
y	1	3	2	4

1.2 Determine the line of regression of Y on X from the following data

x	0	1	2	3	4	5	6
y	2	1	3	2	4	3	5

1.3 Using the method of least squares to find the regression equation Y on X for the data in the following table:

x	1	2	3	4	5	6	7
y	1	2	2	3	4	6	7

1.4 Find the line of regression of Y on X and predict the average value of Y when X is given to be 9

x	3	6	5	4	4	6	7	5
y	3	2	3	5	3	6	6	4

1.5 Find the regression line of Y on X for the following data and estimate Y at X=12

x	6	2	10	4	8
y	9	11	5	8	7

1.6 Find regression equation Y on X and then predict the value of Y for X = 13

CS CamScanner

x	2	4	6	8	10	12	14
y	4	2	5	10	4	11	12

1.7 Estimate the value of Y corresponding X = 5.5

x	1	2	3	4	5
y	2	3	5	4	6

1.8 Find the regression of Y on X from the data and predict the value of Y when X= 10

x	5	6	7	8	9
y	2	4	6	7	11

1.9 Fit a Linear regression of Y on X to the following data and then estimate Y for X = 7.5

x	1	2	3	4	5	6	7	8	9
y	3	4	6	5	10	9	10	12	14

1.10 Given

x	11	13	17	11	15
y	11	8	10	6	11

Required:

(a) Apply the method of least squares to estimate the regression equation.

(b) Estimate the value of Y when X = 14

1.11 Given

x	1	2	5	6	8	9
y	3	5	11	13	17	19

(i) Using method of least square, fit a linear regression equation Y = a + bX

(ii) Estimate the value of Y if X = 15

CamScanner

1.12 Fit a regression line of X on Y

x	1	5	3	2	1	1	7	3
y	6	1	0	0	1	2	1	5

1.13 Find the line of regression of X on Y and estimate X for Y = 7.5

x	6	2	10	4	8
y	9	11	5	8	7

1.14 Find the regression equations of Y on X and X on Y from the following data

x	6	2	10	4	8
y	9	11	5	8	7

1.15 Find the two regression lines from the following data

$n = 5$, $\Sigma x = 0$, $\Sigma y = 26$, $\Sigma xy = 21$

$\Sigma x^2 = 10$ and $\Sigma y^2 = 189$

1.16 Given the following data

x	1	2	5	4	1	6
y	0	3	0	6	2	1

(a) Find regression line of Y on X and then estimate Y at X = 10

(b) Find regression line of X on Y then estimate X for Y = 4

1.17 Given

x	1	3	5	8	5
y	3	4	5	5	3

Required:

(i) Estimate the value of X if Y = 6 and (ii) Estimate the value of Y if X = 4

1.18 An economist gives the following estimates of sales price and demand for a product

CS CamScanner

Price (in dollars)	1	2	3	4	5
Demand (in tons)	9	7	6	3	1

(a) What demand would be predicted if the sales price is $ 1.50? $\hat{X} = c + dY$

(b) What price should be set for the product if the demand is to be 8? $\hat{X} = 1 + 3Y$

1.19 A social researcher wants to determine if there is any correlation between the age in years and height in inches. A random sample of 8 students from a school was taken and their ages and heights were recorded as follows

Age	6	7	8	8	10	11	12	12
Height	46	47	50	51	54	54	56	57

Compute the expected height of the student who will be 15 years of age

1.20 The following table shows the chart of price and demand for an item at different periods of time.

Price Rs. 'X'	12	15	18	25	22	18	30
Demand Kg 'Y'	65	60	50	41	40	56	45

Predict the demand for the price Rs. 35.

Karachi Board 1998, 2010

1.21 Define the regression coefficient of (i) Y on X and (ii) X on Y

1.22 Given

x	0	20	40	60	80
y	51	65	75	85	96

(i) Find the regression line of y on x

(ii) Estimate y for x = 50

Karachi Board 1988

1.23 (a) Define Simple Regression, Linear Regression and Regression Coefficient. Also write the use of Regression.

CS CamScanner

(b) The following table gives the cloth manufactured in a textile mill and the number of persons employed in it.

No. of persons employed	137	209	113	189	176	200	219
Cloth manufactured (000 yards)	23	47	22	40	39	51	49

Estimate how many persons are required to manufacture 30,000 yards of cloth.

Karachi Board 1997

1.24 Calculate Coefficient of Correlation from the following data

x	1	2	3	4	5
y	3	2	5	4	6

1.25 Calculate 'r' from the following data.

x	1	2	3	4	5	6
y	6	4	3	5	4	2

1.26 Calculate Karl Pearson's Coefficient of Correlation from following data.

x	1	2	3	4	5
y	4	5	7	6	8

1.27 Calculate the value of Karl Pearson's Coefficient of Correlation from the following results

$n = 20$, $\Sigma x = 240$, $\Sigma y = 400$, $\Sigma xy = 6960$

$\Sigma x^2 = 4560$, $\Sigma y^2 = 11020$

1.28 If a random sample of size 20 yields the following calculations

$\Sigma x = 200$, $\Sigma y = 287$, $\Sigma xy = 3739$

$\Sigma x^2 = 3568$, $\Sigma y^2 = 4956$

Find Karl Pearson's Correlation.

CS CamScanner

1.29 Calculate the correlation coefficient between the height of father and son from the given data

X	*Height of father (in inches)*	64	65	67	68	69	70	$u = x - 67$
Y.	*Height of Son (in inches)*	66	67	68	70	68	72	$v = y - 68$

1.30 From the following table, calculate the coefficient of correlation by Karl Pearson's method

x	6	2	10	4	8
y	9	11	5	8	7

$\gamma_{xy} = \sqrt{b_{yx} \cdot b_{xy}}$

1.31 Calculate coefficient of linear correlation for the following data

x	12	9	8	10	11	13	7
y	14	8	6	9	11	12	3

and interpret the meaning when r = –1, 0, +1

1.32 Calculate the coefficient of correlation for the following data

x	3	4	5	2	1	3	1	1
y	8	9	11	6	6	9	7	5

1.33 Marks out of 10 obtained on a certain test by 10 Matric students in Mathematics (X) and English (Y) are:

Marks in X	6	8	4	7	5	3	9	8	4	6
Marks in Y	3	4	6	4	5	6	6	5	5	6

Compute: Karl Pearson's Coefficient of Correlation

1.34 For the following data calculate 'r' by using Karl Pearson's correlation coefficient formula.

x	1	2	3	4	5	6
y	6	4	3	5	4	2

CamScanner

1.35 (a) Distinguish between Correlation Coefficient and Regression Coefficient.

(b) Find out the Regression Equation from the following data and estimate Y for X = 8.5

x	6	2	10	4	8
y	6	11	5	8	7

Karachi Board 2002

1.36 The following data shows the test score and the hours of study of 10 students.

Hours studied (x)	14	4	8	11	5	6	13	22	1	18
Test Score (y)	75	32	58	67	39	40	61	90	15	86

(i) Find the Coefficient of Correlation between x and y.

(ii) Determine the line of regression of y on x.

(iii) Estimate the test score of a student who spent 20 hours on his/her study.

Karachi Board 2003

1.37 (a) Define Correlation Coefficient and Regression Coefficient.

(b) Find the Regression Equation of y on x for the given data:

x	1	3	5	7	9	11
y	0	2	4	6	8	10

If x = 13; predict the value of y. Also find coefficient of correlation.

Karachi Board 2001, 2008 (Supp.)

1.38 (a) Differentiate Regression and Correlation and state the properties of Correlation.

(b) Given the following data on the amount of Fertilizer 'X' and Yield 'Y' Wheat:

Fertilizer in Pound 'X'	2	4	5	7	10	11	12	15
Wheat Yield in Bushels 'Y'	8	9	11	11	12	14	15	16

CS CamScanner

(i) Find the Regression Line Y on X.

(ii) Estimate the Wheat Yield, when 13 pound of Fertilizer is used.

Karachi Board 2000

1.39 (a) Explain the difference between positive and negative correlation.

(b) The following table shows the heights and weights of 6 students.

Height (x) (in inches)	60	66	65	70	69	72
Weight (y) (in Kg)	48	54	56	60	58	62

(i) Find the Karl Pearson's coefficient of correlation.

(ii) Find the equation of regression line of weight on height.

(iii) If $u = \frac{x}{4}$ and $v = y - 20$ what would be the coefficient of correlation between u and v (calculation is not required).

Karachi Board 2000, 2006

1.40 (a) Differentiate Regression and Correlation and state the properties of Correlation.

(b) Fit a Regression Line of Y on X to the following data.

X	0	1	2	3	4	5	6
Y	2	1	3	2	4	3	5

Karachi Board 1999

1.41 (a) Define Simple Regression and Simple Correlation and give their examples.

(b) Using the following data verify the relationship between coefficient of correlation and the regression coefficient.

Marks in Mathematics	60	54	72	65	80	75
Marks in Statistics	68	59	69	72	82	70

Karachi Board 1997

CamScanner

1.42 **(a)** Define Correlation and explain positive and negative correlations.

(b) Define the sketches of scatter diagram to explain the various types of correlation.

(c) Calculate the correlation coefficient for the following data and comment on the correlation.

Income (x) (000)	10	20	30	40	50	60
Expenditure (y) (000)	7	21	23	24	36	53

Karachi Board 1996

1.43 **(a)** Define correlation. What are its types? Explain each of them.

(b) State the properties of correlation coefficient.

(c) The data are given below – weight of fertilizer (in hundreds) and tons of corn per acre:

Fertilizer	8.3	8.7	9.2	7.7	8.4	8.8
Yield	13.6	15.4	12.8	13.4	14.6	15.8

Calculate the coefficient of correlation.

Karachi Board 1995

1.44 **(a)** State three important properties of correlation coefficient.

(b) While calculating the correlation coefficient, a computer obtained the following constants:

$$n = 25 \ , \ \Sigma x = 125 \ , \ \Sigma y = 100 \ , \ \Sigma xy = 508$$

$$\Sigma x^2 = 650 \ , \ \Sigma y^2 = 460$$

Later on it was discovered that two pairs were copied down as (x, y): (6, 14), (8, 6) while their correct values were (x, y): (8, 12), (6, 8). Find out the correct value of the coefficient of correlation.

Karachi Board 1994

1.45 The following table gives the quantity of production (x) and manufacturing expenses (y) of 10 randomly selected firm.

CamScanner

x	40	42	50	55	79	88	100	120	140	65
y	150	140	160	155	162	185	165	190	185	150

(i) Determine the regression line of x on y.

(ii) Estimate the manufacturing expenses, given the production to be equal to 85.

(iii) Calculate the coefficient of correlation 'r'.

Karachi Board 1993

1.46 (a) What are the properties of Correlation Coefficient?

(b) Compute and Interpret the Coefficient of Correlation for the following grades of six students selected as random.

Maths grade	70	92	80	74	65	83
English grade	74	84	63	87	78	90

Karachi Board 1990

1.47 (a) Define Regression, Regression Coefficient and Regression lines.

(b) The quantity produced and the total cost of production are given below:

Quantity produced (000 tons)	40	15	2	50	30	12	15
Total cost (in thousand rupees)	70	9	11	20	14	8	9

Estimate the cost when the quantity produced is 26000 tons.

Karachi Board 1992

1.48 (a) Define the coefficient of correlation. Interpret the meaning when $r = -1, 0$ and $+1$.

(b) An economist prepared the following supply schedule, where x represents the price of tea in Rs. per pound and y represents the supply in thousand pounds:

CamScanner

x (Rs. per lb.)	20	22	24	26	28	30	32	34
y (in thousand pounds)	40	45	50	50	60	65	75	80

(i) Determine the equation of regression to estimate y.

(ii) Estimate the supply when the price of tea is Rs. 45 per pound.

Karachi Board 1989

1.49 **(a)** What is the difference between Regression and Correlation.

(b) Fit a Line of Regression of y on x to the following data.

x	1	2	3	4	5	6
y	4	8	12	18	22	22

Karachi Board 1989

1.50 **(a)** Define Correlation and Regression and their relationship.

(b) Find Regression line of y on x and then find y for x = 7

x	2	3	4	5	6
y	8	5	4	3	2

Karachi Board 1988

1.51 **(a)** Give some properties of the coefficient of correlation.

(b) Calculate the coefficient of correlation for the following data and comment.

x	1	3	4	5	6	8
y	5	13	17	21	25	33

Karachi Board 1987

1.52 **(a)** Define Karl Pearson's Coefficient of Correlation: Interpret Positive Correlation, Negative Correlation and Perfect Correlation.

CamScanner

(b) Given:

x	11	7	9	5	8	6	10
y	17	15	13	12	16	14	18

(i) Find the Regression line of y on x (y = a + bx)

(ii) Estimate y for x = 15

Karachi Board 1986

1.53 (a) Define Simple Correlation and explain the different types of correlation with the help of scatter diagram.

(b) State the properties of the coefficient of correlation.

Karachi Board 1985

1.54 (a) Differentiate between Regression and Correlation.

(b) Find the correlation coefficient between x and y from the given data and interpret the relationship.

x	2	4	6	8	10
y	1	3	2	4	6

Karachi Board 1984

1.55 (a) Define correlation and write the properties of correlation.

(b) Calculate Correlation Coefficient of the following data and comment on it.

x	8.2	9.6	7.0	9.4	10.9	7.1
y	8.7	9.6	6.9	8.5	11.3	7.0

Karachi Board 1992

1.56 The following table shows the height and weight of 10 students.

CamScanner

Height (x) (in inches)	60	65	66	70	72	69	58	66	67	70
Weight (y) (in Kg)	50	55	58	60	62	65	59	68	70	72

(i) Find the Karl Pearson's Coefficient of Correlation.

(ii) Find the equation of Regression Line of weight on height.

(iii) If $u = \frac{x}{2}$ and $v = y - 30$, what would be the coefficient of correlation between u and v (calculation in not required).

Karachi Board 1999

1.57 **(a)** What is the difference between Regression and Correlation.

(b) Calculate the coefficient of correlation from the following data.

x	5	10	4	0	2	7
y	10	12	5	4	1	3

Karachi Board 1998

1.58 **(a)** Define Correlation. What are the properties of Correlation.

(b) Calculate the coefficient of correlation for the data given below:

Mileage (1000 Miles)	40	30	25	30	50	65
Selling Price (Rs. 1000)	50	40	35	35	20	22

Karachi Board 1991

1.59 **(a)** Why is Regression Line so important and so frequently used.

(b) Given the following results for heights and weights of 1000 students.

$\overline{y} = 68.00$ inches , $\overline{x} = 150.00$ lbs , $S_y = 2.50$ inches , $S_x = 20.00$ lbs , $r = 0.60$

Find the two regression lines. Given, Ahmed weighs 200 lbs. Kamil is 5 feet tall. Estimate the height of Ahmed from his weight and the weight of Kamil from his height.

Karachi Board 1996

1.60 **(a)** Define the coefficient of correlation. Interpret the meaning when $r = -1, 0$ and $+1$.

(b) Determine the line of Regression of y on x and then predict the value of y for $x = 7$

x	2	3	4	5	6
y	8	5	4	3	2

Karachi Board 1998

1.61 **(a)** Sketch the following scatter diagrams.

(i) When the coefficient of correlation is positive.

(ii) When the coefficient of correlation is negative.

(iii) When the coefficient of correlation is equal to +1

(iv) When the coefficient of correlation is equal to –1

(b) Calculate the coefficient of correlation between the heights of fathers and sons from the given data

Height of father	65	66	67	67	68	69	70	72
Height of son	67	68	65	68	72	72	69	71

Karachi Board 1991

1.62 **(a)** Find r_{xy} , b_{yx} , b_{xy} from the following data and show that $r_{xy} = \sqrt{b_{yx} \cdot b_{xy}}$

x	1	2	3	4	5
y	2	5	3	2	1

(b) Find a regression line by which we can find an estimate of y for a given value of $x = 3.5$

Karachi Board 1983

CS CamScanner

1.63 (a) Define Correlation. Write the properties of Correlation.

(b) The following data are given

x	100	150	200	300	300	450
y	30	35	50	60	70	85

(i) Find the coefficient of correlation of the data (ii) Estimate y if x = 500

Karachi Board 2002

1.64 (a) Describe correlation coefficient and draw the scatter diagram.

(b) The following results are obtained from the two variables x and y

$a = -78.23$, $b = 2.88$, $c = 22.20$, $d = 0.28$

(i) Determine the two lines of regression.

(ii) Estimate $\hat{y}$ and $\hat{x}$, when $x = 60$ and $y = 175$

(iii) Find the coefficient of Determination.

(iv) Find the coefficient of correlation.

Karachi Board 2002, 2008

1.65 (a) Define Regression and Correlation.

(b) Calculate the regression equation of y on x and the coefficient of correlation from the data given below:

Marks in Stats	40	38	35	42	36
Marks in Maths	30	35	40	36	29

Karachi Board 2002

1.66 (a) Define Coefficient of Correlation.

(b) Compute the Regression equation of y on x for the following data and estimate y for $x = 9$

x	5	6	8	10	12	13	15	16	17	14
y	16	19	23	28	36	41	44	45	50	48

Karachi Board 2003

1.67 (a) Define Correlation and Regression.

(b) For 25 pairs of values, the following computations are made:

$$\sum x = 125 \ , \ \sum y = 100 \ , \ \sum x^2 = 650 \ , \ \sum y^2 = 464 \ , \ \sum xy = 536$$

(i) Find the two lines of Regression (ii) Find the Coefficient of Determination.

Karachi Board 2003,1993

1.68 (a) Define Regression and coefficient of correlation.

(b) Calculate the Equation of Regression of y on x for the following data

x	5	7	8	10	12	15
y	40	35	25	28	18	20

Also calculate the coefficient of correlation.

Karachi Board 2003

1.69 (a) If $b_{yx} = 0.61$ and $b_{xy} = 1.36$ then find coefficient of correlation.

(b) If $b_{xy} = -0.61$ and $b_{yx} = -1.36$, calculate the coefficient of correlation.

Karachi Board 2010

1.70 For the following data show that $r = \sqrt{b_{yx} \times b_{xy}}$

x	0	1	2	3	4
y	3	5	6	6	7

1.71 (a) If the equations of the least squares regression lines are:

$Y = 0.648X + 2.64$ (Y on X) $X = 0.917Y - 1.91$ (X on Y)

Find Coefficient of Correlation.

Karachi Board 2011 (Supp.)

CamScanner

(b) The equations of two Regression lines obtained are as follows

$3x + 12y = 129$ 3x + 12y = 129

$3y + 9x = 46$ 3y + 9x = 46 H.W)

Obtain: (i) The value of Correlation Coefficient and

(ii) The mean of x (iii) The mean of y

(c) Comment on the following statement:

"The coefficeint of Regression of y on x is 1.2 and the coefficient of Regression of x on y is –1.2"

Karachi Board 2000

1.72 (a) Define coefficient of correlation.

(b) Prove that, the correlation is Independent of scale and origin.

(c) The following marks obtained by 5 students in Statistics (x) and Mathematics (y):

x	54	64	43	67	83
y	70	92	80	74	70

Calculate coefficient of correlation between x and y.

Karachi Board 2004

1.73 Same farmers used varying amount of a fertilizer in hundred-weight and obtained the yeild of corn in tons per acre as follows

Fertilizer	8	9	7	8	8	9	10	8	9	9
Yield	13	15	12	13	14	15	15	14	14	15

(a) Find the regression line of Y on X

(b) Predict the yield if the fertilizer is equal to 11 hundred-weight.

(c) Find the coefficient of correlation between the fertilizer and the yield.

Karachi Board 2004

1.74 (a) **Define Regression, Correlation and Rank Correlation.**

(b) Given the following data:

$V(X) = 9$, Regression equations: $4x - 5y + 33 = 0$

$20x - 9y - 107 = 0$

Find: (i) The mean values of x and y

(ii) The standard deviation of y

(iii) The coefficient of correlation r

Karachi Board 1988

1.75 (a) Data regarding income (x) and clothing expenditure (y) of 10 randomly selected families from an area were recorded and the following computations were made:

$\sum x = 200,\ \sum y = 42,\ \sum x^2 = 10000,\ \sum y^2 = 750,\ \sum xy = 1500$

(i) Find the least squares regression equation of y on x.

(ii) What amount is expected to increase in y if x is increased by 10 units.

(b) Prove that the coefficient of correlation is independent of the change of origin and scale.

Karachi Board 2001

1.76 (a) Define Coefficient of Correlation with the help of a scatter diagram. Also write any two properties of Correlation Coefficient.

Karachi Board 2001

(b) Calculate the equation of regression of y on x for the following data

$n = 6,\ \sum X = 225,\ \sum X^2 = 8875,\ \sum XY = 12905,\ \sum Y = 361,\ \sum Y^2 = 22641$

Estimate Y when $X = 37$

Karachi Board 2001

CS CamScanner

1.77 (a) Differentiate Correlation and Rank Correlation.

(b) Calculate the Correlation Coefficient of the following.

Income Rs. (000)	10	23	35	40	51	64
Expenditure Rs. (000)	8	21	25	35	35	56

Karachi Board 1991

1.78 (a) Define Rand Correlation Coefficient.

(b) The following table shows how 10 students arranged in alphabetical order, were ranked according to their achievements in both the laboratory and lecture portions of a Biology courses. Find the Coefficient of Rank Correlation.

Laborotory	8	3	9	2	7	10	4	6	1	5
Lecture	9	5	10	1	8	7	3	4	2	6

Karachi Board 1984

1.79 Two judges in a Music concert were asked to rank 8 candidates in their order preference, their rankings were as follows

Judge A	4	5	1	2	3	6	7	8
Judge B	8	6	2	3	1	4	5	7

Find Coefficient of rank correlation.

1.80 Find Spearman's Rank Correlation for the following data.

First Ranking	5	6	3	2	1	7	4
Second Ranking	3	6	5	4	1	7	2

1.81 Ten Competitors in a voice contest are ranked by two judges (A and B) in the following orders. Calculate Coefficient of Rank Correlation.

Rank of Judge A	1	6	5	10	3	2	4	9	7	8
Rank of Judge B	3	5	8	4	7	10	2	1	6	9

1.82 Six brands of vanilla ice cream are ranked by two taste experts to provide the following data.

Expert-1	1	2	3	4	5	6
Expert-2	3	2	1	4	5	6

Find Rank correlation coefficient (r_s).

1.83 Write short notes on the following

(i) Scatter Diagram.

(ii) Rank Correlation.

(iii) Correlation.

(iv) Regression Lines.

(v) Regression.

CamScanner

Chapter – 2

Important Tools of Probability

Before studying probability theory, it is necessary to discuss first the important tools and the basic concepts of Probability.

Counting Techniques

Counting the number of possible results of an experiment is an important part of probability theory.

Let us consider the following six particular problems which are based on different experiments and where the counting techniques may be useful to apply.

- In how many ways can a man go from Karachi to Quetta and come back by a different bus, if there are 10 buses running between Karachi and Quetta.
- In how many ways can a meal be ordered, if the menu lists 3 soups, 6 meat dishes, 5 desserts and 3 beverages.
- In how many ways can 4 people be seated in a row.
- How many three digits numbers can be formed from the digits 2, 4, 6, 8 if each digit can be used more than once.
- How many results are possible, if a coin is tossed 3 times.
- In how many ways can 7 boys be selected from 13 boys.

It is important to note that, all the above counting problems are related with the certain rules of counting.

These rules include the Fundamental Rule of Counting and the Rules of Permutations and Combinations.

CamScanner

Fundamental Rule of Counting

The fundamental rule is stated as:

If one thing can be done in n_1 ways and a second in n_2 ways and a third in n_3 ways and so forth, then the number of different ways in which they can be done, when taken altogether is $n_1 \times n_2 \times n_3$......... ways.

Example 2.1

There are 4 buses running between two cities. In how many ways can a man go from, one city to other and come back by a different bus.

Solution:

Since there are 4 buses available, therefore there are 4 ways of going from one city to the other, on his return there remain only 3 buses, because the man is to come back by a different bus. Hence there are three ways of return. Therefore total number of ways for round trip is $4 \times 3 = 12$

If B_1, B_2, B_3 and B_4 are the four buses. Then the 12 possible cases going and coming back are

$B_1\,B_2$	$B_1\,B_3$	$B_1\,B_4$
$B_2\,B_1$	$B_2\,B_3$	$B_2\,B_4$
$B_3\,B_1$	$B_3\,B_2$	$B_3\,B_4$
$B_4\,B_1$	$B_4\,B_2$	$B_4\,B_3$

Example 2.2

How many results are possible when a pair of dice is thrown.

Solution:

The first die can land in any of 1 to 6 ways. For each of these 6 ways the second die can also land in 6 ways. Therefore, the pair of dice can land in (6)(6) = 36 ways.

Example 2.3

How many lunches are possible consisting of soup and sandwich and a drink if one can select from 4 soups, 3 kinds of sandwiches and 2 soft drinks.

CamScanner

Solution:

The total of number of lunches would be (4)(3)(2) = 24

Example 2.4

Three travellers arrive at a town where there are four hotels. In how many ways can they take up their rooms, each at a different hotel.

Solution:

The first traveller has choice of four hotels, and when he has made his selection in any one way, the second traveller has a choice of three; therefore the first two can make their choice in 4 x 3 ways; and with any one such choice the third traveller can select his hotel in 2 ways; hence the required number of ways is 4 x 3 x 2 = 24

Example 2.5

Suppose we can go from city A to city B in three different ways (by Car, by Train, by Aeroplane) and from city B to our home H in two different ways (by Taxi, by Rickshaw), then in how many different ways can we go from city A to our home H via city B.

Solution:

The total number of ways of going from A to H via B is found to be 3 x 2 = 6

Example 2.6

How many three digit numbers can be formed from the digits 2, 4, 6 and 8 if repetitions are allowed.

Solution:

Hundred	Ten	Unit
⇓	⇓	⇓
4 choices	4 choices	4 choices

Therefore, we can form a total of (4)(4)(4) = 64, three digit numbers.

CS CamScanner

Example 2.7

In a certain town every telephone number consists of four digits, the first of which cannot be zero. How many different telephone numbers are possible.

Solution:

We think of writing a telephone number by writing the first digit, the second digit, then the third digit and finally, the fourth digit.

(task 1) We can write the first digit in 9 different ways since it can be any of the numbers 1, 2, 3,.........9.

(task 2) Then we can write the second digit in 10 different ways

(task 3) The third digit in 10 different ways and

(task 4) The fourth digit in 10 different ways.

Therefore we can do all four tasks in $9 \times 10 \times 10 \times 10 = 9000$ different ways. There are, therefore, 9000 possible four-digit telephone numbers.

Example 2.8

How may three digit numbers can be formed from the digits 0, 2, 4, 5 and 9 when each digit is used only once.

Solution:

The three positions are:

Hundred	Ten	Unit

There are 4 choices for the hundreds position. For each of these we have 4-choices of the tens position and 3-choices for the units position. Therefore we can form a total of $(4)(4)(3) = 48$, three digit numbers.

Example 2.9

How many even three digit numbers can be formed from the digits 3, 4, 5; 6, 9 if each digit can be used only once.

CamScanner

Solution:

The three positions are:

Hundred	Ten	Unit

Since the number must be even, we have only 2-choices for the units position. For each of these we have 4-choices for the hundreds position and 3-choices for the tens position.

Therefore we can form a total of $(2)(4)(3) = 24$, even three digit numbers

Example 2.10

How many new arrangements can be made from the letters of the word FAVOUR, so that vowel will occupy even place.

Solution:

There are 3 vowels and 3 consonants. Therefore the total 6 places with their number of ways where vowels occupy the even places are given as

Odd	Even	Odd	Even	Odd	Even
$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
3	3	2	2	1	1
ways	ways	ways	ways	way	way

Then total words $= 3 \times 3 \times 2 \times 2 \times 1 \times 1 = 36$

Permutations

A permutation is an ordered arrangement of objects. Suppose there are two letters A and B and we want to arrange them in a row. They can be arranged in two ways. like; AB and BA.

Similarly, if there are three letters A, B, C and we want to arrange them in a row. The six possible arrangements of A, B, C are:

ABC ACB BAC BCA CAB CBA

Each arrangement is a permutation. Hence there are six permutations.

CamScanner

When the number of letters increases it is not possible to write down all the arrangements. Therefore we must have a general formula. Let us consider again the same three letters A, B, C and we have to fill in three places i.e.

In other words we can say, there are 3 positions to be filled with the three letters A, B and C. Therefore we have 3 choices for the first position, and 2 for the second leaving only 1 choice for the last position. So that the 3 positions can be filled in $3 \times 2 \times 1 = 6$ ways. In other words we can say the three letters A, B and C can be arranged in 3! ways.

In general if we have 'n' different objects, we can arrange them in $n(n-1)(n-2)......(3)(2)(1)$ ways i.e. n! ways. We use the notation, ${}^{n}P_{n}$ which is the number of permutations or arrangements of 'n' objects taking them all at a time.

We illustrate the idea of permutations with the help of following examples.

Example 2.11

In how many ways the letters of the word CAT can be arranged .

Solution:

There are 3 letters in this word. the number of ways in which the letters can be arranged is

$${}^{3}p_{3} = 3! = 3 \times 2 \times 1 = 6$$

These arrangements can be listed as:

CAT CTA TAC TCA ATC ACT

Example 2.12

In how many ways can 5 people be lined up to get on a bus.

Solution:

$${}^{5}p_{5} = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120 \text{ ways.}$$

Example 2.13

How many numbers of different digits can be formed by using the digits 3, 4, 5.

CamScanner

Solution:

The number of permutations of three digits arranged three at a time is

$$^{3}p_{3}=3!=3\times2\times1=6$$

These six permutations are

345 354 453 435 534 543

Permutations of n different objects taken r at a time

The number of permutations of n different objects taken r at a time and when repetition is not allowed is given by

$$\boxed{^{n}p_{r}=\frac{n!}{(n-r)!}}$$

If n = r then $^{n}p_{n}=\frac{n!}{(n-n)!}=n!$ where $0!=1$

Example 2.14

Find the number of permutations of 4 objects A, B, C, D taken 2 at a time.

Solution:

The number of permutations of 4 objects A, B, C, D taken 2 at a time is

$$^{4}p_{2}=\frac{4!}{(4-2)!}=\frac{4!}{2!}=\frac{4\times3\times2\times1}{2\times1}=12$$

These permutations are:

AB AC AD BA BC BD CA CB CD DA DB DC

Example 2.15

In how many ways can 5 students be seated in a row having 3 seats.

CamScanner

Solution:

The 3 seats can be filled in ${}^{5}p_{3}$ ways.

Therefore $${}^{5}p_{3} = \frac{5!}{(5-3)!} = \frac{5!}{2!} = \frac{5\times4\times3\times2\times1}{2\times1} = 60$$

Example 2.16

Seven players of Pakistan's Hockey team can play in any of the five forward line positions. In how many ways can these positions be filled.

Solution:

The 5 positions can be filled in ${}^{7}p_{5}$ ways.

Therefore $${}^{7}p_{5} = \frac{7!}{(7-5)!} = \frac{7!}{2!} = \frac{7\times6\times5\times4\times3\times2\times1}{2\times1} = 2520$$

Permutations when Repetition is Allowed

The number of permutations when repetition is allowed, is given by

$$\boxed{{}^{n}p_{r} = n^{r}}$$

Example 2.17

In how many different ways can a 3-questions true-false examination be answered?

Solution:

Here $r = 3$ and $n = 2$, then the number of possible answers $(2)^3 = 8$

i.e. TTT TTF TFT FTT TFF FTF FFT FFF

Example 2.18

How many Licence plates of two letters followed by three digits can be made if the letters and digits can be repeated.

Solution:

There are 26 letters (i.e. from A to Z); and there are 10 digits (i.e. from 0 to 9.)

The two letters can be arranged in $(26)^2 = 676$ ways

and the three digits can be arranged in $(10)^3 = 1000$ ways

Therefore the number of license plates that can made in

$676 \times 1000 = 676000$ ways

Example 2.19

How many two digit numbers can be formed from the digits 1, 3, 7 and 9 when the digits are; (a) not repeated (b) repeated

Solution:

(a) Here $n = 4, r = 2$ and since the digits are not repeated, therefore

$$^4P_2 = \frac{4!}{(4-2)!} = \frac{4\times3\times2\times1}{2\times1} = 12 \text{ numbers}$$

These numbers are:

13, 17, 19, 31, 37, 39, 71, 73, 79, 91, 93, 97

(b) Since the digits are repeated, therefore $^4P_2 = 4^2 = 16$ numbers.

These numbers are:

11, 13, 17, 19, 31, 33, 37, 39, 71, 73, 77, 79, 91, 93, 97, 99

Example 2.20

How many four letters code words are possible using the letters in COIN,

If: (i) the letters may not be repeated

(ii) the letters may be repeated

Solution

(i) Here $n = 4$ and $r = 4$

CamScanner

The number of code words when repetition is not allowed, computed as

$$^{4}p_{4} = \frac{4!}{(4-4)!} = \frac{4!}{0!} = 4\times3\times2\times1 = 24 \qquad \text{where } 0! = 1$$

(ii) Here n = 4 and r = 4

The number of code words when repetition is allowed,

computed as $^{4}p_{4} = 4^4 = 256$

Permutations of "n" Objects when they are not all different

If there are n objects of which n_1 are alike, n_2 are alike, ...n_k are alike, such that

$$n_1 + n_2 + ... + n_k = n$$

Then the number of distinct permutations of these n objects is given by

$$\boxed{P = \frac{n!}{n_1!\cdot n_2!\cdot ....\cdot n_k!}}$$

Example 2.21

In how many ways can the word "BOOK" be arranged?

Solution:

Here $B = 1 = n_1$, $O = 2 = n_2$, $K = 1 = n_3$ and

$$n_1 + n_2 + n_3 = 1 + 2 + 1 = 4 = n$$

Then $P = \frac{4!}{1!\cdot2!\cdot1!} = \frac{4\times3\times2\times1}{1\times2\times1\times1} = 12$

These arrangements are listed as

BOOK	BOKO	BKOO	KBOO	KOBO	KOOB
OOBK	OOKB	OKOB	OBOK	OKBO	OBKO

Example 2.22

Find the number of permutations of 9995

Solution:

Here n = 4

$n_1 = 3$ (9-three times)

$n_2 = 1$ (5-one times)

then $P = \frac{4!}{3!\cdot 1!} = \frac{4\times3\times2\times1}{3\times2\times1\times1} = 4$

These arrangements are: 9995 9959 9599 5999

Example 2.23

In how many ways can the letters of the word STATISTICS be arranged.

Solution:

Here n = 4

$n_1 = 3$ (S-three times)

$n_2 = 3$ (T-three times)

$n_3 = 1$ (A-one time)

$n_4 = 2$ (I-two times)

$n_5 = 1$ (C-one time)

$$P = \frac{10!}{3!\cdot 3!\cdot 1!\cdot 2!\cdot 1!} = \frac{10\times9\times8\times7\times6\times5\times4\times3\times2\times1}{3\times2\times1\times3\times2\times1\times2\times1} = 50,400$$

Example 2.24

Find the number of permutations of 8 balls,taken 8 at a time when 3 are red, 4 are green and 1 is black and if one cannot distinguish between the same colour of balls.

Solution:

Here n = 8

$n_1 = 3$ (No. of red)

$n_2 = 4$ (No. of green)

$n_3 = 1$ (No. of black)

The number of permutations is calculated as

$$P = \frac{8!}{3!\cdot 4!\cdot 1!} = \frac{8\times 7\times 6\times 5\times 4\times 3\times 2\times 1}{3\times 2\times 1\times 4\times 3\times 2\times 1\times 1} = 280.$$

Combinations

Combination means selection consisting of r objects chosen from n different objects and when the order is not important. Thus, the possible selections or combinations of the letters A, B, C taking two at a time are : AB, AC and BC The number of combinations of n different objects taken r at a time is given by

$$\boxed{{}^{n}C_{r} = \frac{n!}{r!\,(n-r)!}}$$

Following are some of the important properties of combinations:

(i) ${}^{n}C_{r} = {}^{n}C_{n-r}$ (ii) ${}^{n}C_{0} = {}^{n}C_{n} = 1$ (iii) ${}^{n}C_{1} = {}^{n}C_{n-1} = n$

Example 2.25

In how many ways a committee of 3 students can be selected from 4 students (A, B, C and D)

Solution

Here $n = 4$ and $r = 3$

then, the number of ways to select 3 students out of 4 students is

$${}^{4}C_{3} = \frac{4!}{3!\,(4-3)!} = \frac{4!}{3!\;1!} = \frac{4\times 3\times 2\times 1}{3\times 2\times 1} = 4$$

The possible committees consisting of three students each are:

ABC ABD BCD ACD

Example 2.26

Evaluate the following combinations

(i) $^{20}C_4$ (ii) $^{18}C_{16}$ (iii) $^{9}C_0$ (iv) $^{52}C_2$ (v) $^{4}C_2 \cdot {}^{5}C_3$ (vi) $^{8}C_8$

Solution:

(i) $^{20}C_4 = \dfrac{20!}{4!(20-4)!} = \dfrac{20\times19\times18\times17\times16!}{24\times16!} = 4845$

(ii) $^{18}C_{16} = \dfrac{18!}{16!(18-16)!} = \dfrac{18\times17\times16!}{16!\times2!} = 153$

(iii) $^{9}C_0 = \dfrac{9!}{0!(9-0)!} = \dfrac{9!}{0!\times9!} = 1$ since $0! = 1$

(iv) $^{52}C_2 = \dfrac{52!}{2!(52-2)!} = \dfrac{52!}{2!\times50!} = \dfrac{52\times51\times50!}{2\times1!\times50!} = 1326$

(v) $^{4}C_2 = \dfrac{4!}{2!(4-2)!} = \dfrac{4\times3\times2!}{2\times1\times2!} = 6$ and $^{5}C_3 = \dfrac{5!}{3!(5-3)!} = \dfrac{5\times4\times3\times2!}{3\times2\times1\times2!} = 10$

Then $^{4}C_2 \cdot {}^{5}C_3 = 6\times10 = 60$

(vi) $^{8}C_8 = \dfrac{8!}{8!(8-8)!} = \dfrac{8!}{8!\times0!} = 1$

Example 2.27

From a group of 10 boys and 6 girls a committee of 3 boys and 2 girls is to be selected. In how many ways can this be done?

Solution:

3 boys can be selected from 10 boys in $^{10}C_3$ ways

i.e. $^{10}C_3 = \dfrac{10!}{3!(10-3)!} = \dfrac{10\times9\times8\times7!}{3\times2\times1\times7!} = 120$

2 girls can be selected from 6 girls in $^6C_2$ ways

i.e. $$^6C_2 = \frac{6!}{2!\,(6-2)!} = \frac{6\times5\times4!}{2\times1\times4!} = 15$$

The committee of 3 boys and 2 girls can be formed in

$$^{10}C_3 \times {}^6C_2 = 120\times15 = 1800 \text{ ways}$$

Example 2.28

In how many ways can the city football team of 11 players be selected from 16 players.

Solution:

Here $n = 16$ and $r = 11$

$$^{16}C_{11} = \frac{16!}{11!\,(16-11)!} = \frac{16!}{11!\times5!} = \frac{16\times15\times14\times13\times12\times11!}{11!\times5\times4\times3\times2\times1} = 4368 \text{ ways}$$

Example 2.29

In how many ways can a cricket eleven be chosen out of 14 players? How many of them will

(i) include a particular player (ii) exclude a particular player

Solution:

Total number of ways of selecting 11 players out of 14 players is $^{14}C_{11}$

i.e. $$^{14}C_{11} = \frac{14!}{11!\,(14-11)!} = \frac{14\times13\times12\times11!}{3\times2\times1\times11!} = 364 \text{ ways}$$

(i) Since a particular player is to be always included. Therefore, we have to select 10 players out of 13 players.

$$^{13}C_{10} = \frac{13!}{10!\,(13-10)!} = \frac{13\times12\times11\times10!}{3\times2\times1\times10!} = 286 \text{ ways}$$

CamScanner

(ii) Since a particular player is to be always excluded, then we have to select 11 players out of 13 players.

$$^{13}C_{11} = \frac{13!}{11!\,(13-11)!} = \frac{13\times12\times11!}{2\times1\times11!} = 78 \text{ ways}$$

Example 2.30

From a class containing 5 boys and 6 girls a group of 5 students is to be selected. There are how many combinations.

(i) 3 boys and 2 girls (ii) 2 boys and 3 girls (iii) 5 boys

Solution:

(i) The number of ways of selecting 3 boys is

$$^{5}C_{3} = \frac{5!}{3!\,(5-3)!} = \frac{5!}{3!\;4!} = 10$$

The number of ways of selecting 2 girls is

$$^{6}C_{2} = \frac{6!}{2!\,(6-2)!} = \frac{6!}{2!\;4!} = 15$$

Then number of ways of selecting 3 boys and 2 girls is

$$^{5}C_{3}\times{}^{6}C_{2} = 10\times15 = 150$$

(ii) The number of ways of selecting 2 boys is

$$^{5}C_{2} = \frac{5!}{2!\,(5-2)!} = \frac{5!}{2!\;3!} = 10$$

The number of ways of selecting 3 girls is

$$^{6}C_{3} = \frac{6!}{3!\,(6-3)!} = \frac{6!}{3!\;3!} = 20$$

The number of ways of selecting 2 boys and 3 girls is

$$^{5}C_{2}\times{}^{6}C_{3} = 10\times20 = 200$$

CamScanner

(iii) The number of ways of selecting 5 boys is

$${}^{5}C_{5} \times {}^{6}C_{0} = 1 \times 1 = 1$$

Example 2.31

Indicate in how many ways a hand of 4 cards can be selected from a deck of 52 cards.

Solution:

The number of ways of selecting 4 cards out of 52 cards is

$${}^{52}C_{4} = \frac{52!}{4!\,(52-4)!} = \frac{52!}{4!\ 48!} = 270725$$

What is a Set

Any well-defined list or collection of objects is called a set. For example, a set of natural numbers; a set of statistics books in a library, a set of dots ranging from 1 to 6 on the die, a set of students of Accounting, etc. The objects that are in a set, are called members or elements of the set.

Set are denoted by capital letters A, B, C, ..., X, Y, Z, whereas the members of a set are denoted by small letters a, b, c, ..., x, y, z. The members are enclosed by curly brackets { }, as shown in the following examples;

The set A consisting of numbers 4, 6, 8 and 10 may be written as

$$A = \{4, 6, 8, 10\}$$

or the set of B, of possible outcomes when a coin is tossed, may be written as

$$B = \{H, T\}$$

where H and T corresponds 'head' and 'tail' respectively.

The symbol $\in$ means "is an element of" or "belongs to" and $\notin$ means "is not an element of" or "does not belong to". If 'a' is an element of the set A and b is not, we write $a \in A$ and $b \notin A$. For example, if A = {4, 6, 8, 10} then $4 \in A$ and $9 \notin A$.

Equal Sets

Two sets are equal if they have exactly the same elements in them.

CS CamScanner

For example,

let $A = \{3, 5, 7\}$, $B = \{5, 3, 7\}$ and $C = \{3, 5, 7, 9\}$

then $A = B$, $A \neq C$ and $B \neq C$

Null Set

The null set or empty set is a set that contains no elements. We denote this set by the symbol ϕ. For example, let A be a set of people in Pakistan who are older than 300 years, then A must be the null set and it is written as

$$A = \{\ \} = \phi$$

Subsets

If every element of set B is also an element of set A, then B is called as a subset of A. Symbolically, it is written as $B \subseteq A$ which read as 'B' is a subset of A or B is contained in A.

For example,

if $A = \{3, 4, 5, 6, 7, 12\}$ and $B = \{3, 5, 7\}$

then $B \subseteq A$, i.e. B is a subset of A.

According to the above definition every set is a subset of itself. Any subset of a set that is not the set itself is called a proper subset of the set. Therefore, B is a proper subset of A if $B \subset A$ and $B \neq A$.

For example,

The set $B = \{2, 4\}$ is a proper subset of $A = \{1, 2, 3, 4, 5\}$. However the set $C = \{3, 2, 5, 1, 4\}$ is a subset of A but not a proper subset, since $A = C$.

Universal Set

It is to be noted that all sets are subsets of a particular set. This set is called the universal set and is usually denoted by U and sometimes by S in the name of sample space.

CamScanner

For example,

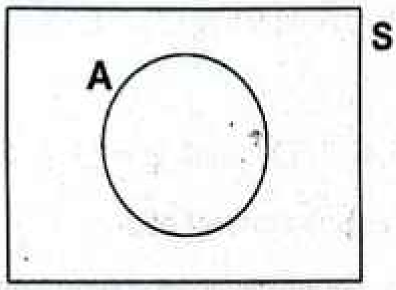
All the subsets of the universal set $S = \{1, 2, 3\}$ are $\{1, 2, 3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1\}, \{2\}, \{3\}$ and ϕ. Note that the number of subsets of a universal set containing three elements is $2^3 = 8$.

Venn Diagram

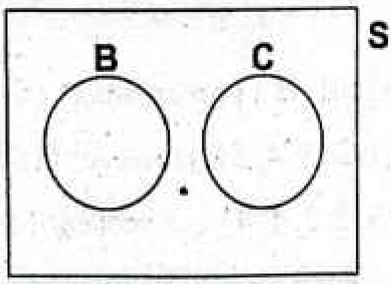
A graphic device to represent the relationship between the subsets and the universal set is known as Venn diagram. The universal set is represented by a rectangle and the subsets indicated by circles.

For example,

If the universal set is "all horses" and set A "all race horses", this relationship can be illustrated by the following Venn diagram:

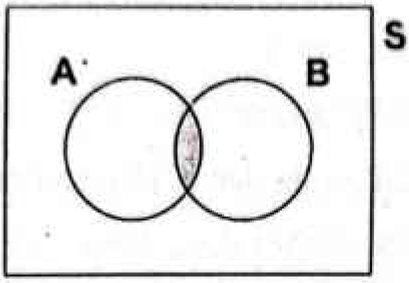


Suppose B denotes "all horses that are completely black" and C denotes "all horses that are completely white". This relationship is shown in the following figure



In this case no horse is a member of both sets and we say that B and C are disjoint sets. A and B are disjoint of no member of A is a member of B. Other examples of disjoint sets are the set of bicycles and the set of motor cars, the set of positive integers and the set of negative integers.

If we let A denote "all race horses" and B denote "all black horses", then Venn diagram for this relationship would be as shown in the following figure



Because some race horses are black, there are some horses that are members of both sets A and B (shaded area). In such examples, we say that set A meets B (or sets A and B are overlapping sets).

Operations on Sets

The basic operations are Union, intersection, difference and complementation.

Union of Sets

The union of the sets A and B is the set of elements that belonging to A or to B or to both A and B. This new set is denoted by $A \cup B$ (i.e. A union B).

Example 2.32

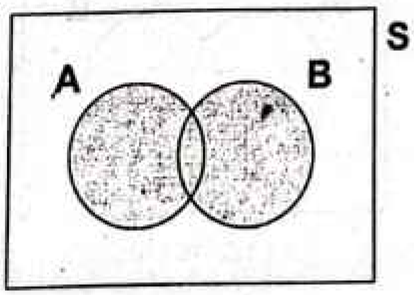
If S = {Bilal, Jamil, Javed, Pervaiz, Haider, Basit, Jehangir, Talha, Faiq}

A = {Jamil, Javed, Pervaiz}

B = {Bilal, Jamil Haider, Javed}

Then $A \cup B$ = {Bilal, Jamil Javed, Pervaiz, Haider}

Because these five boys are just those boys who belong either to A or to B or to both.



CamScanner

Example 2.33

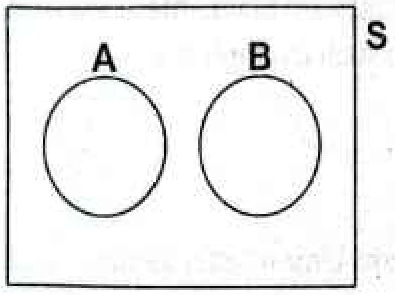
If S = {Bilal, Jamil, Javed, Pervaiz, Haider, Basit, Jehangir, Talha, Faiq}

A = {Javed, Pervaiz, Jamil}

B = {Haider, Basit, Jehangir}

Then $A \cup B$ = {Javed, Pervaiz, Jamil, Haider, Basit, Jehangir}

Because these six boys are just those boys who belong to A or to B. The Venn diagram will be;



observe that A and B are disjoint because no member of A is a member of B.

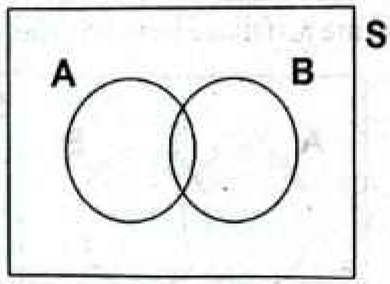
Intersection of Sets

The intersection of two sets A and B is the set of elements that are common to A and B. Symbolically, we write $A \cap B$ for the intersection of A and B. The elements in the set $A \cap B$ must be those and only those which belong to both A and B.

Example 2.34

Let A = {3, 4, 5, 6, 7} and B = {4, 6, 8, 10};

then $A \cap B$ = {4, 6}



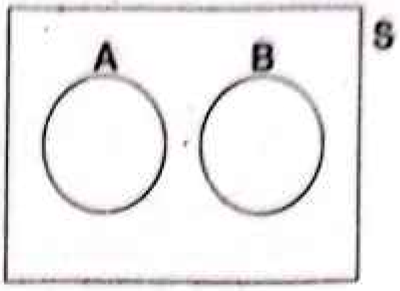
$A \cap B$ is shaded

Example 2.35

Let $A = \{a, e, i, o, u\}$ and $B = \{k, l, m\}$;

then $A \cap B = \{\ \} = \phi$

Since A and B have no elements in common, therefore the sets A and B are called disjoint sets.



Difference of Sets

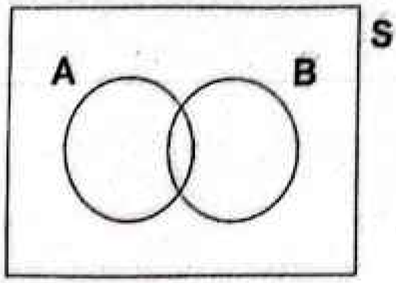
The difference of two sets A and B denoted by $A - B$ or $A \backslash B$ or $A - (A \cap B)$ read as A minus B.

Example 2.36

Let $A = \{0, 2, 4, 6, 8\}$ and $B = \{0, 4, 8\}$

Then $A - B = \{2, 6\}$

The shaded area of $A - B$ is shown in the following Venn diagram.



$A - B$ is shaded

Complement of Set

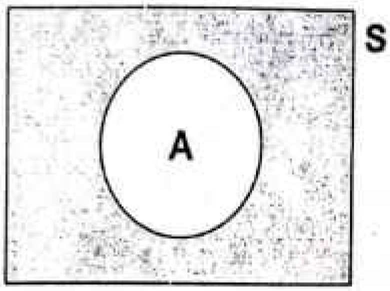
If A is a subset of the universal set S, then the difference of S and A i.e. $S - A$ is called complement of A and it is denoted by A'

Example 2.37

If S = {all college students}

A = {all male students}

Then $A' = S - A$ = {all female students}



A' is shaded

Laws of Algebra of Sets

1	$A \cup B = B \cup A$	9	$(A')' = A$
2	$A \cap B = B \cap A$	10	$S' = \phi$
3	$A \cup A = A$	11	$\phi' = S$
4	$A \cap A = A$	12	$A \cup S = S$
5	$A \cap \phi = \phi$	13	$A \cap S = A$
6	$A \cup \phi = A$	14	$(A \cup B)' = A' \cap B'$
7	$A \cap A' = \phi$	15	$(A \cap B)' = A' \cup B'$
8	$A \cup A' = S$	16	$A \cap B' = A - A \cap B$

Example 2.38

Let S = {1, 2, 3, 4, 5, 6}

A = {1, 2, 3}

B = {2, 3, 4}

Find (i) $A \cup B$ (ii) $A \cap B$ (iii) $B - A$ (iv) A'

(v) $(A \cup B)'$ (vi) $(A \cap B)'$ (vii) $(A')'$

Solution:

(i) A = {1, 2, 3}, B = {2, 3, 4} then $A \cup B = \{1, 2, 3, 4\}$

(ii) A = {1, 2, 3}, B = {2, 3, 4} then $A \cap B = \{2, 3\}$

(iii) A = {1, 2, 3}, B = {2, 3, 4} then $B - A = \{4\}$

(iv) $A' = S - A = \{4, 5, 6\}$

(v) $(A \cup B)' = S - A \cup B$

since S = {1, 2, 3, 4, 5, 6}

and $A \cup B = \{1, 2, 3, 4\}$

$(A \cup B)' = S - A \cup B = \{5, 6\}$

(vi) $(A \cap B)' = S - A \cap B$

since S = {1, 2, 3, 4, 5, 6}

and $A \cap B = \{2, 3\}$

$(A \cap B)' = S - A \cap B = \{1, 4, 5, 6\}$

(vii) $(A')' = S - A' = \{1, 2, 3, 4, 5, 6\} - \{4, 5, 6\}$

$= \{1, 2, 3\} = A$

Random Experiment

Any process or activity that generates observations (data) is called an experiment. If an experiment is conducted repeatedly under the homogenous conditions, and the result is not unique but may be any one of the various possible results (outcomes) is called a random experiment. The tossing of a coin, the rolling of a balanced die, drawing of a card from a well shuffled pack of 52 cards, etc. are some important examples of random experiments.

Outcome

A single result of an experiment is called an outcome.

CamScanner

Sample Space

A set of all possible outcomes of an experiment is called the sample space and is represented by S. The possible outcomes of sample space are the elements of S and are also called sample points.

Example 2.39

Consider the experiment of tossing a die. If we are interested in the number that shows on the top face, then the sample space would be

$$S = \{1, 2, 3, 4, 5, 6\}$$

where S consists of 6 possible outcomes.

Example 2.40

A cricket match may result in a win, a loss or a draw, so that there are three possible outcomes i.e. win, loss and draw in this experiment of match, then

$$S = \{\text{win, loss, draw}\}$$

Event

Any subset of a sample space is called an "Event" denoted by A, B, C, ...

There are two types of events:

(i) Simple Event (ii) Compound Event

If an event contains only one outcome, it is called simple event. A compound event is one that can be expressed as the union of simple events.

Note:

1. An empty set is also an event, sometimes called impossible event. Whereas the event S is called sure event.
2. Events are said to be equally likely if they have the same chance of occurrence.

The events may also be classified as complementary events, mutually exclusive events, independent and dependant events.

CamScanner

Example 2.41

If we toss a die two times then the sample space is represented as:

$$S = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

This sample space contains 36 outcomes. But if we are interested in the sum of 7 on both dice, therefore we will have a set {(3,4) (4,3) (2,5) (5,2) (1,6) (6,1)} and it is also the subset of S. This type of subset in statistical experiment is called an event and can be represented as $A = \{(3,4)\ (4,3)\ (2,5)\ (5,2)\ (1,6)\ (6,1)\}$. This event contains 6 outcomes.

Example 2.42

Suppose in an experiment a die and a coin be tossed once

a) List the elements of the sample space.

b) List the elements corresponding to event A that an even number occurs.

c) List the elements corresponding to event B that number occurs on the die is less than 4

d) List the elements corresponding to event C that 2 tails occurred.

Solution:

a) S = {(H,1) (H,2) (H,3) (H,4) (H,5) (H,6) (T,1) (T,2) (T,3) (T,4) (T,5) (T,6)}

b) A = {(H,2) (H,4) (H,6) (T,2) (T,4) (T,6)}

c) B = {(H,1) (H,2) (H,3) (T,1) (T,2) (T,3)}

d) $C = \{\ \} = \phi$

CamScanner

Example 2.43

An experiment consists of flipping a coin and then flipping it a second time if a head occurs. If a tail occurs on the first flip, then a die is tossed once.

a) List the elements of the sample space S

b) List the elements of S corresponding to the event A that a number less than 4 occurred on the die.

c) List the elements of S corresponding to event B that 2 tails occurred.

Solution:

a) S = {HH, HT, T1, T2, T3, T4, T5, T6}

b) A = {T1, T2, T3}

c) $B = \phi$

CS CamScanner

SAMPLE SPACE

Sample Space for the card randomly selected from a deck of 52 cards								
Denomination	SUIT							
	RED = 26				BLACK = 26			
	Hearts		Diamonds		Spades		Clubs	
King	K	♥	K	♦	K	♠	K	♣
Queen	Q	♥	Q	♦	Q	♠	Q	♣
Jack	J	♥	J	♦	J	♠	J	♣
10	10	♥	10	♦	10	♠	10	♣
9	9	♥	9	♦	9	♠	9	♣
8	8	♥	8	♦	8	♠	8	♣
7	7	♥	7	♦	7	♠	7	♣
6	6	♥	6	♦	6	♠	6	♣
5	5	♥	5	♦	5	♠	5	♣
4	4	♥	4	♦	4	♠	4	♣
3	3	♥	3	♦	3	♠	3	♣
2	2	♥	2	♦	2	♠	2	♣
Ace	A	♥	A	♦	A	♠	A	♣

CamScanner

EXERCISE – 2

2.1 A man has 4 suits, 3 ties and 2 pairs of shoes. In how many different ways can he dress up.

2.2 A restaurant menu lists 2 soups, 5 meat dishes, 3 beverages and 2 deserts. In how many ways can a meal be ordered.

2.3 How many motor cycle number plates can be made if each plate contains 2 different letters followed by 3 different digits.

2.4 There are 6 roads between the towns A and B and 4 other roads between the towns B and C.

(i) In how many ways can one drive from A to C via B.

(ii) In how many ways can one drive form A to C and back passing through B both ways.

Karachi Board 1994

2.5 A committee consists of ten people. It is decided to appoint a chairman, a vice-chairman and a secretary. In how many different ways can this be done.

Karachi Board 2001

2.6 How many different words can be formed from the letters of "RANDOM" ?

Karachi Board 1991

2.7 How many distinct arrangements of letters can be made from the word PRINCE.

2.8 How may three digit numbers can be formed from the digits 2, 4, 6, 8 when each digit is used only once.

2.9 In how many ways can 5 people be seated on a sofa if there are only 3 seats are available.

Karachi Board 1989

CamScanner

2.10 How many three digit numbers can be formed from the digits 1, 2, 4, 5 and 9 when each digit is used only once.

Karachi Board 1998

2.11 How many three digit numbers are possible from the digits 1, 3, 5, 6, 8 if the digits are not repeated.

Karachi Board 2003

2.12 (a) How many different words can be made using the letters of the word "COLUMNS".

(b) How many words of this permutation start with the letter N.

Karachi Board 1996

2.13 How many words can be formed out of the letters of the word "FRESH".

(i) How many of them will begin with S.

(ii) How many of them begin with R and end with H.

2.14 (a) There are 3 Men and 2 Women. In how many ways they can sit alternately.

(b) In how many ways can 10 people be selected on a bench if:

(i) only 4 seats are available. (ii) 10 seats are available

(iii) 8 seats are available.

Karachi Board 1988

2.15 How many four code words are possible using the letters in "CHIEF" if

(a) the letters may not be repeated **(b)** the letters may be repeated

2.16 How many three digit numbers are possible from the digits 1 to 5 if the digits are:

(a) repeated **(b)** not repeated.

2.17 How many three digit numbers are possible from the digits 0, 2, 4, 6, 8 if the digits are **(a)** repeated **(b)** not repeated.

2.18 (a) How many four digit numbers are possible from the digits 0, 1, 2, 3, 4. Numbers are repeated.

CamScanner

(b) How many three digit numbers can be formed from the digits 0, 3, 5, 6, 8 if

(i) each digit is to be used only once. (ii) the digits can be repeated.

Karachi Board 2004

2.19 There are 6 machines but there are only 3 spaces on the floor of the machine shop for the machines. In how many different ways can the 6 machines be arranged in three spaces.

Karachi Board 1991

2.20 How many two-digit numbers can be formed from the digits 3, 4, 5 if:

(i) Repetition of digits is allowed. **(ii)** Repetition of digits is not allowed.

Karachi Board 2001

2.21 Four married couples have bought 8 seats in a row for a concert. In how many different ways can they be seated.

(i) If each couple is to sit together

(ii) If all the men sit together to the right of all the women.

Karachi Board 2003

2.22 (a) Simplify the following permutations.

(i) ${}^{4}P_{2}$ (ii) ${}^{9}P_{1}$ (iii) ${}^{5}P_{0}$ (iv) ${}^{20}P_{4}$

(b) A student has 8 books to arrange on a shelf of a bookcase. How many ways is it possible to arrange the 8 books.

2.23 Find the number of permutations of 5 objects say A, B, C, D and E taken three at a time.

2.24 (a) Four persons enter in a bus in which 6 seats are vacant. In how many ways can they be seated.

(b) In how many ways can 5 people line up for a group photograph?

2.25 How many football teams can be formed with 5 men, who can play any front line positions and 5 men, who can play any back positions.

CS CamScanner

2.26 How many three digit numbers can be formed from the digits 0, 1, 3, 5, 7 if:

(i) Each digit is used only once.

(ii) Digits can be repeated.

Karachi Board 1998

2.27 (i) How many three digit numbers can be formed from the digits 0, 1, 2, 3, 4, 5 and 6 if each digit can be used only once.

(ii) How many of these are odd numbers?

(iii) How many of these are greater than 330?

Karachi Board 1999, 2002

2.28 (i) How many three digit numbers can be formed from the digits 0, 1, 2, 3, 4, 5, 6 & 7 if each digit is used only once?

(ii) How many are greater than 440 ?

Karachi Board 2000

2.29 How many 4 digit numbers can be formed with the digits 0, 1, 2, ..., 9 without repetition if the last digit is zero.

Karachi Board 1997

2.30 (a) In how many ways can 5 men and 4 women be seated in a row so that women occupy the even places.

Karachi Board 1997

(b) Three girls and two boys are invited to a party. In how many ways can they be seated if

(i) they can sit anywhere

(ii) the girls and boys sit alternately

(iii) two boys occupy the ends

Karachi Board 2004

CamScanner

2.31 Four different Mathematics books, six different physics books and two different chemistry books are to be arranged on a shelf. How many different arrangements are possible if all the books of each particular subject must be put together?

Karachi Board 1997

2.32 From 7 different consonants and 5 vowels how may words can be formed consisting of 4 different consonants and 3 different vowels.

Karachi Board 1993

2.33 How many 4 digit odd numbers can be made using the integers 1, 2, 3, 4, 5 and 6 without repetition.

Karachi Board 1992

2.34 How many three digit numbers can be formed using the integers 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 if the repetition of integers is not allowed in any number thus formed.

Karachi Board 1992

2.35 (a) How many three digit numbers can be formed from the digits 0, 1, 2, 3, 4, 5 and 6 if each digit is being used only once.

(b) How many of these are odd numbers.

(c) How many of these are even numbers.

Karachi Board 1996

2.36 (a) Define Combination and Permutation and give examples.

(b) How many five-digit odd numbers can be made using the integers 1, 2, 3, 4, 5, 6, 7 and 8 if the integers are not repeated.

Karachi Board 1995

2.37 In how many ways the letters in the word "TOMORROW" can be arranged.

2.38 Find the number of permutations that can be formed from the letters of the word "DISCUSS".

2.39 How many distinct arrangements of letters can be made from the word "PEOPLE".

CS CamScanner

2.40 How many distinguishable permutations can be made form the letters of "ARRANGEMENTS".

Karachi Board 1991

2.41 In how many different ways can the letters of the word "COMBINATION" be arranged.

Karachi Board 1993

2.42 In how many different ways can the letters of the word "STATISTICS" be arranged.

Karachi Board 2001, 2006, 2008

2.43 In how many ways can 5 red, 3 yellow and 4 green chips be arranged in row if the chips of the same colour are not distinguishable from each other.

Karachi Board 1992

2.44 In how many ways can 4 mango, 3 guava and 5 apple trees be arranged in a line if the trees of the same kind are not distinguishable.

Karachi Board 1997

2.45 There are 3 white, 2 red and 4 green flags. In how many ways these can be arranged in a row if the flags of same colour are not distinguishable.

2.46 In how many ways 4 Mathematics, 5 Statistics and 7 English books can be arranged on a shelf in a row if one can not distinguish between same types of books.

2.47 There are 2 Mangoes, 3 Apples and 2 Banana plants. In how many ways these can be planted in a line. If one cannot differentiate between plants of the same kind.

2.48 (a) In how many ways a committee of 3 students can be selected from 5 students.

(b) In how many ways can a committee of 5 people be chosen out of 9 people.

Karachi Board 1989

2.49 Evaluate the following combinations:

(i) $^{10}C_7$ (ii) $^{9}C_3$ (iii) $^{19}C_{10}$ (iv) $^{10}C_1$ (v) $^{20}C_2$

CamScanner

2.50 (a) A club consists of 15 members. In how many ways can a committee of 3 members be formed.

Karachi Board 2003

(b) There are 10 questions in an examination paper. In how many ways can a candidate select 5 questions.

2.51 (a) Differentiate between Combination and Permutation.

(b) In how many ways can a team of 6 boys be chosen out of 10 boys?

(c) In how many ways can 6 differently coloured chips be arranged in a row?

Karachi Board 2002

2.52 (a) A Committee of 5 members is to be selected from among 8 men and 5 women. In how many ways can a committee be chosen with 3 men and 2 women.

Karachi Board 1999

(b) In how many ways can a committee consisting of 4 men and 3 women be randomly selected from 6 men and 5 women.

Karachi Board 2002

2.53 A Committee of 3 members is to be selected from among 5 men and 4 women. In how many ways can the committee be chosen if no restrictions are imposed for the selection of men and women.

Karachi Board 1998

2.54 How many different Committees can be formed from 5 American, 4 Russian and 3 French delegates at U. N. O. if each Committee comprises 3 American, 2 Russian and 1 French delegates.

Karachi Board 1995

2.55 A bag contains 5 red and 4 black balls. In how many ways can 3 balls be drawn from the bag (without replacement) containing at least 2 red balls.

Karachi Board 1994

CamScanner

2.56 From a group of 5 men and 3 women, how many committees of size 3 are possible with 2 men and one women.

Karachi Board 1990

2.57 In how many ways can a committee of 3 persons be formed from 3 Chinese, 4 Pakistanis and 5 Americans to have:

(i) At least a Pakistani

(ii) All Chinese

(iii) At most two Americans

Karachi Board 1988

2.58 (a) Define Combination and Permutation and give examples.

(b) A Committee of 4 persons is to be formed from a group of 4 men and 6 women. How many committees consisting of at least 3 men are possible.

Karachi Board 1992

2.59 A bag contains 5 red and 4 white balls. Five balls are drawn at random from the bag without replacement. In how many ways can the balls be drawn if:

(i) no restriction is imposed.

(ii) less than 3 balls are white.

Karachi Board 1993

2.60 (a) Evaluate the following:

(i) ${}^{10}P_5$ (ii) ${}^{10}C_5$ (iii) $\binom{10}{2,\ 3,\ 5}$.

Karachi Board 1987

(b) Enumerate all possible combination and permutation of the three letters chosen from the four letters A, B, C, D

Karachi Board 2003

2.61 (a) Explain the difference between Permutation and Combination.

CS CamScanner

(b) A committee of 5 members has to be selected from 5 Asians, 3 Europeans and 2 Africans. Find the possible number of committees which contains:

(i) Exactly 3 Asians.

(ii) 2 Asians, 1 European and 2 Africans.

(iii) At least 2 Africans.

Karachi Board 1986

2.62 From a batch consisting of 3 boys and 5 girls a group of 3 students is to be selected. How many groups of at least 2 boys are possible?

Karachi Board 1992

2.63 In how many ways can 5 cards be selected from an ordinary deck of 52 cards if the selected cards consist of at least 2 kings.

Karachi Board 1997

2.64 Out of 5 Mathematicians and 7 Physicists a committee consisting of 2 mathematicians and 3 physicists is to be formed. In how many ways can this be done if:

(i) Any mathematician and any physicist can be included

(ii) One particular physicist must be in the committee.

Karachi Board 2003

2.65 A Committee of 4 members is to be selected from among 7 men and 3 women. In how many ways can a committee be chosen with at least 3 men.

Karachi Board 2000

2.66 (a) In how many ways a cricket eleven be chosen out of 14 players. How many of them will (i) include a particular player (ii) exclude a particular player.

(b) Out of 14 investment banking firms, 6 are classified as type A and 8 are classified as type B firms. In how many ways can five firms be selected if at least 4 firms are of type A.

Karachi Board 2001

2.67 In how many ways can 7 boys be selected from 13 boys, so as always to include 3 particular boys.

CamScanner

2.68 Out of 16 students, in how many ways a group of 7 students may be selected so that:

(i) particular 4 students will not come

(ii) particular 4 students will always come.

2.69 A party of 4 is to be formed from 8 boys and 5 girls so as to include 2 boys and 2 girls. In how many different ways can the party be formed if two particular girls refuse to join the same party.

2.70 Out of 5 men and 3 women, a committee of 6 is to be selected. In how many ways can this be done:

(i) when there are 4 men (ii) when there is a majority of men.

2.71 **(a)** Define Permutation and Combination.

(b) (i) How many three digit numbers can be formed from the digits 0, 1, 2, 3, 4, 5, 6 and 7 if each digit is to be used only once.

(ii) How many of these are greater than 400

(c) From a group of 5 men and 3 women, how many committees of size 3 can be drawn with no restrictions.

Karachi Board 2003

2.72 **(a)** Define the following terms:

(i) Outcome (ii) Sample Space (iii) Event

(b) An experiment consists of tossing a coin only once. List the elements of the sample space S.

2.73 Write the sample space in each of the following experiments

(i) Two dice are tossed simultaneously

(ii) Two coins are tossed simultaneously

(iii) A coin and a die are tossed simultaneously

(iv) Three coins are tossed simultaneously

2.74 An experiment consists of flipping a coin three times.

(i) List the elements of the sample space S

CamScanner

(ii) List the elements of S corresponding to event A that 2 tails occurred

(iii) List the elements of S corresponding to event B that at least 2 heads occurred.

Karachi Board 1988

2.75 An experiment involves tossing a pair of dice, 1 green and 1 red and recording the numbers that come up

(a) List the elements of the sample space S

(b) List the elements of S corresponding to event A that the sum is less than 5.

(c) List the elements of S corresponding to event B that a 4 occurs on either die

(d) List the elements of S corresponding to event C that a 3 comes up on the green die.

(e) List the elements of S corresponding to event D that the sum is greater than 12

2.76 An experiment consists of flipping a coin three times.

(a) List the elements of the sample space S.

(b) List the elements of S corresponding to event A that at least 2 tails occurred.

(c) List the elements of S corresponding to event B that one head occurred.

2.77 An experiment consists of flipping a coin four times

(i) List the elements of the sample space S.

(ii) List the elements of S corresponding to event A that at least 2 tails occurred.

(iii) List the elements of S corresponding to event B that 2 heads occurred.

Karachi Board 1998

2.78 Define sample space. Find sample space when:

(i) two coins are tossed once

(ii) three coins are tossed once

(iii) a die and a coin are tossed together once

Karachi Board 1992

Chapter – 3

Probability

Introduction

The idea of probability is familiar to every one. In our daily life we make statements such as; "Pakistan will probably win the Sharjah Cup", "I have 50:50 chance of getting an odd number when a die is tossed", "I have a fair chance of passing Intermediate examination", etc. In each case we are expressing an outcome of which we are not certain, but we have some degree of confidence in the validity of each of these statements because of past information. Therefore, probability is a measure of degree of belief in a particular statement or probability is a measure of the chance that an uncertain event will occur.

Definition of Probability

There are different definitions of probability. But here we shall discuss only the classical definition of probability.

Classical Definition of Probability

If A is any event in the sample space S, then probability of event A, is defined as

$$P(A) = \frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } S}$$

OR

$$P(A) = \frac{\text{Number of favourable cases}}{\text{Number of total possible cases}}$$

Symbolically: $P(A) = \frac{n(A)}{n(S)}$

Where all the cases (i.e. outcomes) are equally likely and disjoint.

CS CamScanner

Note: Two or more outcomes are said to be equally likely if the chance of their happening is equal.

Basic Properties of Probability

If A is any event in a sample space S, then

(i) $0 \le P(A) \le 1$

(ii) $P(S) = 1$ i.e., when the event is certain

(iii) $P(\phi) = 0$ i.e., when the event is impossible

(iv) $P(A') = P(\text{A does not occur}) = 1 - P(A)$

Example 3.1

A six-sided die is tossed only once. What is the probability of getting:

(i) an even number (ii) a number less than 3 (iii) a 4 or a higher number

(iv) a 7 (v) a number from 1 to 6

Solution:

The sample space is

(i) S = {1, 2, 3, 4, 5, 6}

Let A be the event of getting an even number, then A = { 2, 4, 6 }

Therefore $P(A) = \dfrac{n(A)}{n(S)} = \dfrac{3}{6} = \dfrac{1}{2}$

(ii) Let B be the event of getting a number less than 3, then B = { 1, 2 }

therefore $P(B) = \dfrac{n(B)}{n(S)} = \dfrac{2}{6} = \dfrac{1}{3}$

(iii) Let C be the event of getting a 4 or a higher number, then C = { 4, 5, 6}

therefore $P(C) = \dfrac{n(C)}{n(S)} = \dfrac{3}{6} = \dfrac{1}{2}$

(iv) Let D be the event of getting a 7, then D = { } = ϕ

therefore $P(D) = P(\phi) = 0 = P(\text{impossible event})$

CamScanner

(v) Let E be the event of getting a number from 1 to 6, then E={1,2,3,4,5,6}

therefore $$P(E) = \frac{n(E)}{n(S)} = \frac{6}{6} = 1 = P(\text{certain event})$$

Example 3.2

A fair coin is tossed three times. What is the probability of getting

(a) exactly two heads (b) at least two heads

Solution:

The sample space for this experiment is

S = { HHH, HHT, HTH, THH, HTT, THT, TTH, TTT} thus n(S) = 8

(a) Let A be the event that exactly two heads occur, then A = { HHT, HTH, THH }

thus n(A) = 3

$$P(A) = \frac{n(A)}{n(S)} = \frac{3}{8}$$

(b) Let B be the event that at least two heads occur

then B = { HHT, THH, HTH, HHH }, thus n(B) = 4

$$P(B) = \frac{n(B)}{n(S)} = \frac{4}{8} = \frac{1}{2}$$

Example 3.3

A pair of dice is rolled once. What is the probability of getting

(i) a total of 6 (ii) a total of 13

Solution:

The sample space is

$$S = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

thus $n(S) = 36$

(i) Let $A = D_1 + D_2 = 6$ then $A = \{ (2,4)\ (4,2)\ (1,5)\ (5,1)\ (3,3) \}$

Thus $n(A) = 5$. Hence $P(A) = \frac{n(A)}{n(S)} = \frac{5}{36}$

(ii) Let $B = D_1 + D_2 = 13$ then $B = \{\ \} = \phi$ Hence $P(B) = P(\phi) = 0$

Example 3.4

A card is drawn at random from an ordinary pack of 52 playing cards. Find the probability that the card (a) is a jack (b) is not a jack

Solution:

One card from 52 cards can be drawn in ${}^{52}C_1 = 52$ ways

Then, No. of sample points in S i.e. Total No. of cases = $n(S) = 52$

(a) Let A be the event that the card is a jack, then the number of ways of drawing a jack out of 4 cards is = ${}^{4}C_1 = 4$

The the number of outcomes in A = $n(A) = 4$

$$P(A) = \frac{n(A)}{n(S)} = \frac{4}{52} = \frac{1}{13}$$

Therefore, the probability that the card drawn is a jack is $\frac{1}{13}$

(b) Let A' be the event 'The card is not a jack'.

$$P(A') = 1 - P(A) = 1 - \frac{1}{13} = \frac{12}{13}$$

Therefore the probability that the card drawn is not a jack is $\frac{12}{13}$.

Example 3.5

A bag contains 3 Red and 4 Black balls. A ball is drawn at random from the bag. What is the probability that the ball is black.

Solution:

One ball from 7 balls can be drawn in ${}^{7}C_{1} = 7$ ways

Then No. of outcomes in S = n(S) = 7

Let A = ball drawn is black

The No. of ways of drawing one black ball is ${}^{4}C_{1} = 4$

Then No. of outcomes in A = n(A) = 4, therefore

$$P(A) = \frac{n(A)}{n(S)} = \frac{4}{7}$$

Example 3.6

A bag contains 3 white balls, 4 black balls, and 5 red balls. If 3 balls are drawn at random, determine the probability that

(i) all 3 are red (ii) all 3 are black (iii) 2 are red and one is black

(iv) one of each colour is drawn.

Solution:

(i) 3 balls can be selected out of 12 balls in ${}^{12}C_{3}$ ways.

i.e. $${}^{12}C_{3} = \frac{12!}{3!(12-3)!} = 220 \text{ ways.}$$

Then outcomes in S = n(S) = 220

Let A = all 3 balls are red

Then number of ways of drawing 3 red balls out of 5 red balls is

$$^5C_3 = \frac{5!}{3!(5-3)!} = 10$$

Then No. of outcomes in A = n(A) = 10

$$P(A) = \frac{n(A)}{n(S)} = \frac{10}{220} = \frac{1}{22}$$

(ii) Let B = all 3 balls are black

The number of ways of drawing 3 black balls out of 4 black balls is

$$^4C_3 = \frac{4!}{3!(4-3)!} = 4$$

Then No. of outcomes in B = n(B) = 4

$$P(B) = \frac{n(B)}{n(S)} = \frac{4}{220} = \frac{1}{55}$$

(iii) Let C = 2 balls are red and one is black

The number of ways of drawing 2 red and 1 black out of 5 red and 4 black balls is

$$^5C_2 \times {^4C_1} = 10 \times 4 = 40$$

Then No. of sample points in C = n(C) =40

$$P(C) = \frac{n(C)}{n(S)} = \frac{40}{220} = \frac{2}{11}$$

(iv) Let D = One ball of each colour is drawn

The number of ways of drawing one-white, one-black and one-red balls out of 3-white, 4-black and 5-red balls is

$$^3C_1 \times {^4C_1} \times {^5C_1} = 3 \times 4 \times 5 = 60$$

Then No. of outcomes in D = n(D) =60

CamScanner

$$P(D) = \frac{n(D)}{n(S)} = \frac{60}{220} = \frac{3}{11}$$

Example 3.7

Three card are drawn at random from a full pack of cards. Find the probability that the three cards drawn are a King, a Queen and a Jack.

Solution:

The number of ways of drawing 3 cards from 52 cards is ${}^{52}C_3 = 22100$

Then number of outcomes in S = n(S) = 22100

Let A = 3 cards consist of 1-King, 1-Queen and 1-Jack are drawn.

The number of ways of drawing one-King, one-Queen and one-Jack out of 4-King, 4-Queen and 4-Jack is

$${}^{4}C_1 \times {}^{4}C_1 \times {}^{4}C_1 = 4 \times 4 \times 4 = 64$$

Then number of outcomes in A = n(A) = 64

$$P(A) = \frac{n(A)}{n(S)} = \frac{64}{22100} = \frac{16}{5525}$$

Example 3.8

A room has 3 lamps. From a collection of 10 light bulbs of which 6 are not good, a person selects 3 at random and puts them in the sockets. What is the probability that he will have light from all the three lamps.

Solution:

There are 10 bulbs in all. Hence the number of ways of selecting

3 bulbs out of 10 = ${}^{10}C_3 = 120$

Then No. of outcomes in S = n(S) =120 which are equally likely.

Let A = person selects all three good bulbs = ${}^{4}C_3 = 4$

Then No. of outcomes in A = n(A) = 4, therefore

CamScanner

$$P(A) = \frac{n(A)}{n(S)} = \frac{4}{120} = \frac{1}{30}$$

Kinds of Events and Laws of Probability

Following are the four important kinds of events;

1. Mutually exclusive events
2. Not Mutually exclusive events
3. Independent events
4. Dependent events

Mutually Exclusive Events (Disjoint Events)

The two events A and B are said to be mutually exclusive if any one of them occurs at a time i.e. either A occurs or B occurs and they both cannot occur at the same time. Since events are disjoint, then $A \cap B = \phi$

Suppose a die is tossed. Let A be the event that an even number turns up and B be the event that an odd number shows. Then A = { 2, 4, 6 } and B = { 1, 3, 5 } and since $A \cap B = \phi$ therefore A and B are mutually exclusive.

The probability of these two mutually exclusive events is computed as

$$P(A \text{ or } B) = P(A) + P(B)$$

OR

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

Note: If $P(A \cup B) = P(A) + P(B) = P(S) = 1$ we say that events are exhaustive.

Example 3.9

Two dice are rolled once. What is the probability of getting a total of 7 or 11

Solution:

The sample space is

$$S=\begin{bmatrix}(1,1)&(1,2)&(1,3)&(1,4)&(1,5)&(1,6)\\(2,1)&(2,2)&(2,3)&(2,4)&(2,5)&(2,6)\\(3,1)&(3,2)&(3,3)&(3,4)&(3,5)&(3,6)\\(4,1)&(4,2)&(4,3)&(4,4)&(4,5)&(4,6)\\(5,1)&(5,2)&(5,3)&(5,4)&(5,5)&(5,6)\\(6,1)&(6,2)&(6,3)&(6,4)&(6,5)&(6,6)\end{bmatrix}$$

Let the event $A = D_1 + D_2 = 7$

then A = { (3,4) (4,3) (2,5) (5,2) (1,6) (6,1) } and the event

$B = D_1 + D_2 = 11$ then B = { (6,5) (5,6) }

Since $A \cap B = \phi$ or in other words a total of 7 and 11 both cannot occurs at the same time. Then A and B are mutually exclusive,

Therefore $P(A \cup B) = P(A) + P(B)$

where $P(A) = \frac{n(A)}{n(S)} = \frac{6}{36}$ and $P(B) = \frac{n(B)}{n(S)} = \frac{2}{36}$

Then $P(A \cup B) = \frac{6}{36} + \frac{2}{36} = \frac{8}{36} = \frac{2}{9}$.

Example 3.10

A card is drawn from a full pack. Find the probability that it a King or a Queen

Solution:

One card from 52 cards can be drawn is ${}^{52}C_1 = 52$ ways

Then number of outcomes is S = n(S) = 52

Let A = a King is drawn and B = a Queen is drawn

Then number of outcomes in A = n(A) = ${}^{4}C_1 = 4$

the number of outcomes in B = n(B) = ${}^{4}C_1 = 4$

Since any one of them (i.e. either a King or a Queen) occurs at a time, therefore events are mutually exclusive, and therefore

CamScanner

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

where $P(A) = \frac{n(A)}{n(S)} = \frac{4}{52} = \frac{1}{13}$ and $P(B) = \frac{n(B)}{n(S)} = \frac{4}{52} = \frac{1}{13}$

Then $P(A \cup B) = \frac{1}{13} + \frac{1}{13} = \frac{2}{13}$

Not Mutually Exclusive Events

The two events A and B are said to be not mutually exclusive if there is some thing common between A and B and therefore, $A \cap B \neq \phi$. In other words, if the event A occurs, the event B may also occurs and if the event B occurs, the event A may also occurs. Then, the events A and B are said to be not mutually exclusive.

Suppose following are the two events in an experiment

A = a King is drawn from a pack and B = a Red card is drawn from a pack

If we draw a King it may be a Red and if we draw a Red card it may be a King. Therefore theses two events are not mutually exclusive.

The probability of these two not mutually exclusive events (i.e. either A or B or both A and B occur) is computed as

$$\boxed{\begin{array}{c} P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \\ \text{OR} \\ P(A \cup B) = P(A) + P(B) - P(A \cap B) \end{array}}$$

Example 3.11

A class contains 16 boys and 10 girls of which half of the boys and half of the girls are fat. Find the probability that a student chosen at random is a boy or a fat student.

Solution;

One student can be chosen out of 26 student in ${}^{26}C_1 = 26$ ways

Then number of outcomes in S = n(S) = 26

CamScanner

Let A = 1 boy student can be chosen out of 16 boys students in ${}^{16}C_1 = 16$ ways.

Then number of outcomes in A = n(A) = 16

Let B = fat student is chosen

1 fat student can be chosen out of 13 fat students in ${}^{13}C_1 = 13$ ways.

Then number of outcomes in B = n(B) = 13

Since A and B are not mutually exclusive, then

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

where $P(A) = \frac{n(A)}{n(S)} = \frac{16}{26}$ and $P(B) = \frac{n(B)}{n(S)} = \frac{13}{26}$

$A \cap B$ shows that the student is a boy and fat.

Therefore, No. of students in $(A \cap B) = n(A \cap B) = {}^{8}C_1 = 8$

then $P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{8}{26}$

Thus $P(A \cup B) = \frac{16}{26} + \frac{13}{26} - \frac{8}{26} = \frac{21}{26}$

Example 3.12

A die is tossed only once. What is the probability that either a number less than 5 or an even number occurs on the die

Solution:

The sample space is S = { 1, 2, 3, 4, 5, 6 } then n(S) = 6

Let A = number is less than 5, then A = { 1, 2, 3, 4 } and then n(A) = 4

B = even number occurs then B = { 2, 4, 6 } and n(B) = 3

Since $A \cap B = \{ 2, 4 \} \neq \phi$

Then A and B are not mutually exclusive events.

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

CS CamScanner

$$P(A)=\frac{n(A)}{n(S)}=\frac{4}{6}=\frac{2}{3}\ ,\ P(B)=\frac{n(B)}{n(S)}=\frac{3}{6}=\frac{1}{2} \text{ and } P(A\cap B)=\frac{n(A\cap B)}{n(S)}=\frac{2}{6}=\frac{1}{3}$$

Then, $P(A\cup B)=\frac{2}{3}+\frac{1}{2}-\frac{1}{3}=\frac{5}{6}$

Example 3.13

A student applies for admission to college A and B. He estimates that the probability of being admitted to A is 0.7, the probability of being admitted to B is 0.5 and the probability of being admitted to both A and B is 0.4. What is the probability that he will be admitted to at least one of the college (i.e. either A or B or both A and B)

Solution:

Since P(A) = 0.7, P(B) = 0.5 and

$P(A\cap B)$ = P(Both A and B) = 0.4

Now to find $P(A\cup B)$ = P(either A or B or both A and B) = ?

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

$P(A\cup B)=0.7+0.5-0.4=0.8$

Independent Events

Suppose A and B are two events in an experiment. If A occurs firstly and B occurs secondly and if the probability of second event B has not changed due to occurrence of first event A, then the events A and B are called independent.

Suppose two cards are drawn one by one, from a pack of 52 cards, if the events are;

A = a king is drawn in first attempt, and

B = a king is card drawn in second attempt

whereas the first card is replaced before drawing the second card.

Therefore $P(A)=\frac{4}{52}$ and $P(B)=\frac{4}{52}$

Since the probability of second event B is not changed due to the occurrence of first event A, then events A and B are called independent

The probability of these two independent events A and B (i.e. both of them occur) is computed as

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

OR

$$P(A \cap B) = P(A) \cdot P(B)$$

Important Note:

If A and B are independent, then

(i) A' and B are independent

(ii) A and B' are independent

(iii) A' and B' are independent

Example 3.14

A bag contains 5 red and 7 black balls. A ball is drawn at random from the bag, the colour is noted and the ball is replaced. A second ball is then drawn. Find the probability that the first ball is red and the second is black.

Solution:

First ball can be drawn in ${}^{12}C_1 = 12$ ways, then n(S) =12

Let A = the first ball is red, then

the number of ways of drawing one red ball is ${}^{5}C_1 = 5$, then n(A) = 5

$$P(A) = \frac{n(A)}{n(S)} = \frac{5}{12}$$

Let B = the second ball is black.

As the first ball is replaced before the second draw is made, therefore, A and B are independent events

The number of ways of drawing one black ball is ${}^{7}C_1 = 7$, then n(B) = 7

CS CamScanner

$$P(B) = \frac{n(B)}{n(S)} = \frac{7}{12}$$

$$P(A\ and\ B) = P(A \cap B) = P(A) \cdot P(B) = \frac{5}{12} \times \frac{7}{12} = \frac{35}{144}$$

Example 3.15

A die is thrown two times. Find the probability of getting a 4 on the first throw and an odd number on the second throw.

The sample space is

$$S = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1;6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

Let A = 4 occurs on the first toss throw

then, A = { (4;1) (4,2) (4,3) (4,4) (4,5) (4,6) }

$$P(A) = \frac{n(A)}{n(S)} = \frac{6}{36} = \frac{1}{6}$$

Let B = an odd number is obtained on the second toss

$$B = \begin{bmatrix} (1,1) & (2,1) & (3,1) & (4,1) & (5,1) & (6,1) \\ (1,3) & (2,3) & (3,3) & (4,3) & (5,3) & (6,3) \\ (1,5) & (2,5) & (3,5) & (4,5) & (5,5) & (6,5) \end{bmatrix}$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{18}{36} = \frac{1}{2}$$

Since the result on the second toss is not affected in any way by the result on the first toss, therefore A and B are independent

Then $P(A \cap B) = P(A\ and\ B) = P(A) \cdot P(B) = \frac{1}{6} \times \frac{1}{2} = \frac{1}{12}$

Example 3.16

The probability that A will be alive in 30 years is 0.4 and the probability that B will be alive in 30 years is 08. What is the probability that

(i) both will be alive in 30 years

(ii) both of them will die

(iii) A will be alive and B dead

Solution:

P(A-alive) = 0.4 P(B-alive) = 0.8

P(A-dead) = 0.6 P(B-dead) = 0.2

Since events are independent, then

(i) P(both will alive) = P(A-alive) P(B-alive) = (0.4) (0.8) = 0.32

(ii) P (both will die) = P(A-dead). P(B-dead) = (0.6) (0.2) = 0.12

(iii) P(A will alive and B dead) = P(A-alive). P(B-dead) = (0.4) (0.2) = 0.08

Dependent Events

Suppose A and B are two events in an experiment, if A occurs firstly and B occurs secondly and if the probability of second event B changes due to the occurrence of first event A. Then these events are said to be dependent.

Suppose two cards are drawn one by one from a pack of 52 cards

if A = a card drawn in first attempt is red

B = a card drawn in second attempt is black

whereas the first card is not replaced, then

$$P(A) = \frac{^{26}C_1}{^{52}C_1} = \frac{26}{52} = \frac{1}{2}$$

Since first card is not replaced before drawing the second, therefore probability of 2nd event B has been changed due to the first event A i.e. $\dfrac{{}^{26}C_1}{{}^{51}C_1} = \dfrac{26}{51}$ and is denoted by $P(B|A)$

Then events A and B are called dependent events. The probability of these two dependent events (i.e. both of them occur) is computed as

$$\boxed{\begin{gathered} P(A \text{ and } B) = P(A) \cdot P(B|A) \\ \text{OR} \\ P(A \cap B) = P(A) \cdot P(B|A) \end{gathered}}$$

where $P(B|A)$ is called the **Conditional Probability** of B given A and is defined as

$$\boxed{P(B|A) = \frac{P(A \cap B)}{P(A)}}$$

Similarly conditional probability of A given B is

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

Note:

(i) If A and B are independent events then $P(A|B) = P(A)$ and $P(B|A) = P(B)$

(ii) $P(A'|B) = 1 - P(A|B)$

Example 3.17

Two cards are drawn in succession from a deck of 52 playing cards without replacement. What is the probability that both cards are spades.

Solution:

Two cards are drawn one by one

Let A = Card drawn in first attempt is spade, then $P(A) = \frac{^{13}C_1}{^{52}C_1} = \frac{13}{52} = \frac{1}{4}$

Let B = Card drawn in second attempt is spade

Since first card is not replaced therefore $P(B|A) = \frac{^{12}C_1}{^{51}C_1} = \frac{12}{51} = \frac{4}{17}$ and

since events are dependent, then $P(A \text{ and } B) = P(A) \cdot P(B|A)$

i.e. $P(A) \cdot P(B|A) = \frac{1}{4} \times \frac{4}{17} = \frac{1}{17}$

Hence, P(both Spades) $= \frac{1}{17}$

Example 3.18

A box contains 8 tickets bearing the numbers 1, 2, 3, 4, 5, 6, 8, 10. One ticket is drawn and kept aside. Then a second ticket is drawn. What is the probability that both the tickets show even numbers.

Solution:

Let A = First ticket drawn shows an even number

then $P(A) = \frac{^{5}C_1}{^{8}C_1} = \frac{5}{8}$

B = Second ticket drawn shows an even number

Since first ticket is not replaced then $P(B|A) = \frac{^{4}C_1}{^{7}C_1} = \frac{4}{7}$

Probability that both the tickets show even number is

$$P(A \cap B) = P(A) \cdot P(B|A) = \frac{5}{8} \times \frac{4}{7} = \frac{5}{14}$$

CS CamScanner

Example 3.19

If $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{3}$, $P(A\cap B)=\frac{1}{4}$

then find (i) $P(A|B)$ (ii) $P(B|A)$ (iii) $P(A'|B)$

Solution:

Since $P(A|B)$ and $P(B|A)$ are the conditional probabilities

Then (i) $P(A|B)=\dfrac{P(A\cap B)}{P(B)}=\dfrac{\frac{1}{4}}{\frac{1}{3}}=\dfrac{3}{4}$

(ii) $P(B|A)=\dfrac{P(A\cap B)}{P(A)}=\dfrac{\frac{1}{4}}{\frac{1}{2}}=\dfrac{1}{2}$

(iii) $P(A'|B)=1-P(A|B)=1-\dfrac{3}{4}=\dfrac{1}{4}$

MISCELLANEOUS SOLVED PROBLEMS

Problem 1

If one card is drawn at random from a well shuffled pack of 52 cards find the chance that the card is:

(i) Spade (ii) Black (iii) Not a Diamond (iv) An ace (v) A face card

Solution:

(i) One card can be drawn out of 52 in $^{52}C_1=52$ ways

then $n(S)=52$

Let A = card drawn is a Spade

The number of outcomes in A i.e. $n(A) = {}^{13}C_1 = 13$

then $P(A) = \dfrac{n(A)}{n(S)} = \dfrac{13}{52} = \dfrac{1}{4}$

(ii) Let B = card drawn is a Black

and $n(B) = {}^{26}C_1 = 26$ then $P(B) = \dfrac{n(B)}{n(S)} = \dfrac{26}{52} = \dfrac{1}{2}$

(iii) Let C = card drawn is a diamond

and C′ = card drawn is not a diamond

Then $n(C) = {}^{13}C_1 = 13$

and $P(C) = \dfrac{n(C)}{n(S)} = \dfrac{13}{52} = \dfrac{1}{4}$

$P(C') = P(\text{Not a diamond}) = 1 - P(C) = 1 - \dfrac{1}{4} = \dfrac{3}{4}$

(iv) Let D = card drawn is an ace, then $n(D) = {}^{4}C_1 = 4$

and $P(D) = \dfrac{n(D)}{n(S)} = \dfrac{4}{52} = \dfrac{1}{13}$

(v) Let E = card drawn is a face card, then $n(E) = {}^{12}C_1 = 12$

and $P(E) = \dfrac{n(E)}{n(S)} = \dfrac{12}{52} = \dfrac{3}{13}$

Problem 2

Two cards are drawn from a pack of cards at random. What is the probability that it will be

(i) A diamond and a heart (ii) A king and a queen (iii) Two kings

CamScanner

Solution:

(i) $n(S) = {}^{52}C_2 = 1326$

Let A = a diamond and a heart, then $n(A) = {}^{13}C_1 \times {}^{13}C_1 = 169$

$$P(A) = \frac{n(A)}{n(S)} = \frac{169}{1326} = \frac{13}{102}$$

(ii) Let B = a King and a Queen, then $n(B) = {}^{4}C_1 \times {}^{4}C_1 = 16$

$$P(B) = \frac{n(B)}{n(S)} = \frac{16}{1326} = \frac{8}{663}$$

(iii) Let C = 2 Kings, then $n(C) = {}^{4}C_2 = 6$

$$P(C) = \frac{n(C)}{n(S)} = \frac{6}{1326} = \frac{1}{221}$$

Problem 3

A bag contains 12 red marbles and 8 black marbles. If two marbles are drawn from the bag at random. What is the probability that

(i) Both are red (ii) Both are black (iii) One is red and one is black

Solution:

12-Red	8-Black

= 20

(i) 2 balls can be drawn out of 20 balls in ${}^{20}C_2 = 190$ ways, then $n(S) = 190$

Let A = Both are red then $n(A) = {}^{12}C_2 = 66$

$$P(A) = \frac{n(A)}{n(S)} = \frac{66}{190} = \frac{33}{95}$$

(ii) Let B = Both are Black, then $n(B) = {}^{8}C_2 = 28$

CamScanner

$$P(B) = \frac{n(B)}{n(S)} = \frac{28}{190} = \frac{14}{95}$$

(iii) Let C = One is red and one is black then $n(C) = {}^{12}C_1 \times {}^{8}C_1 = 96$

$$P(C) = \frac{n(C)}{n(S)} = \frac{96}{190} = \frac{48}{95}$$

Problem 4

A marble is drawn at random from a box containing 10 red, 30 white, 20 blue and 15 orange. Find the probability that it is

(i) white (ii) not blue (iii) orange or red (iv) red, white or blue

Solution:

10 Red	30 White	20 Blue	15 Orange

= 75 marbles

(i) One marble can be drawn out of 75 marbles in ${}^{75}C_1 = 75$ ways

then $n(S) = 75$

Let A = Marble is white, then $n(A) = {}^{30}C_1 = 30$

$$P(A) = \frac{n(A)}{n(S)} = \frac{30}{75} = \frac{2}{5}$$

(ii) Let B = Marble is blue, then $n(B) = {}^{20}C_1 = 20$

$$P(B) = \frac{n(B)}{n(S)} = \frac{20}{75} = \frac{4}{15}$$

$P(B') =$ P(Marble is not blue) $= 1 - P(B) = 1 - \frac{4}{15} = \frac{11}{15}$

(iii) Let A = Marble is orange and B = Marble is red

Since A and B are mutually exclusive, then

CS CamScanner

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

$$P(A) = \frac{n(A)}{n(S)} = \frac{^{15}C_1}{^{75}C_1} = \frac{15}{75} \quad \text{and} \quad P(B) = \frac{n(B)}{n(S)} = \frac{^{10}C_1}{^{75}C_1} = \frac{10}{75}$$

$$P(A \cup B) = \frac{15}{75} + \frac{10}{75} = \frac{25}{75} = \frac{1}{3}$$

(iv) Let A = Marble is red, B = Marble is white and C = Marble is blue

Since all these events are mutually exclusive, then

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{^{10}C_1}{^{75}C_1} = \frac{10}{75}$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{^{30}C_1}{^{75}C_1} = \frac{30}{75}$$

$$P(C) = \frac{n(C)}{n(S)} = \frac{^{20}C_1}{^{75}C_1} = \frac{20}{75}$$

$$P(A \cup B \cup C) = \frac{10}{75} + \frac{30}{75} + \frac{20}{75} = \frac{60}{75} = \frac{4}{5}$$

Problem 5

A bag contains 25 balls marked 1 to 25. One ball is drawn at random. What is the probability that it is marked with a number of multiple of 5 or 6

Solution:

The sample space is S = { 1, 2, 3, 4, 5, 6, 7, ..., 25 }

Let A = a number multiple of 5, then A = { 5, 10, 15, 20, 25 }

Let B = a number multiple of 6, then B = { 6, 12, 18, 24 }

Since $A \cap B = \phi$

therefore A and B are mutually exclusive

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

$$P(A)=\frac{n(A)}{n(S)}=\frac{5}{25} \quad \text{and} \quad P(B)=\frac{n(B)}{n(S)}=\frac{4}{25} \quad \text{and then}$$

$$P(A\cup B)=\frac{5}{25}+\frac{4}{25}=\frac{9}{25}$$

Problem 6

One ticket is drawn at random from a set of 20 tickets numbered form 1 to 20. What is the probability that the number of the ticket drawn is divisible by

(i) 2 or 3 (ii) 3 or 7.

Solution:

(i) The sample space is S = { 1, 2, 3, 4, 5, 6, ..., 20 }

Let A = Number is divisible by 2, then A = { 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 }

Let B = Number is divisible by 3, then B = { 3, 6, 9, 12, 15, 18 }

since $A\cap B=\{6, 12, 18\}\neq\phi$.

therefore A and B are not mutually exclusive, then

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

$$P(A)=\frac{n(A)}{n(S)}=\frac{10}{20}=\frac{1}{2} \quad , \quad P(B)=\frac{n(B)}{n(S)}=\frac{6}{20}=\frac{3}{10} \quad \text{and}$$

$$P(A\cap B)=\frac{n(A\cap B)}{n(S)}=\frac{3}{20}, \quad \text{then}$$

$$P(A\cup B)=\frac{1}{2}+\frac{3}{10}-\frac{3}{20}=\frac{13}{20}$$

(ii) Let A = Number is divisible 3, then A = { 3, 6, 9, 12, 15, 18 }

Let B = Number is divisible 7, then B = { 7, 14 }

and since $A\cap B=\phi$

therefore A and B are mutually exclusive

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

$P(A) = \dfrac{n(A)}{n(S)} = \dfrac{6}{20}$ and $P(B) = \dfrac{n(B)}{n(S)} = \dfrac{2}{20}$, then

$$P(A \cup B) = \frac{6}{20} + \frac{2}{20} = \frac{8}{20} = \frac{2}{5}$$

Problem 7

Out of 80 members of a club 30 drink tea, 20 drink coffee and 10 drink both. One member is selected at random. Find the probability that he drinks either tea or coffee.

Solution:

$$n(S) = {}^{80}C_1 = 80$$

Let A = Member drinks tea and B = Member drinks coffee

then $n(A) = {}^{30}C_1 = 30$, $n(B) = {}^{20}C_1 = 20$ and $n(A \cap B) = {}^{10}C_1 = 10$

Since A and B are not mutually exclusive events, then

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

$$P(A) = \frac{n(A)}{n(S)} = \frac{30}{80} = \frac{3}{8}$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{20}{80} = \frac{2}{8}$$

$$P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{10}{80} = \frac{1}{8}$$

$$P(A \cup B) = \frac{3}{8} + \frac{2}{8} - \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$$

Problem 8

The probability that a contractor gets an electric contract is 2/7 and that of his getting a plumbing contract is 3/5. If the probability of his getting at least one of the contracts is 5/7. What is the probability of his getting both the contracts.

CamScanner

Solution:

Let A = getting electric contract, then $P(A)=\frac{2}{4}$

B = getting plumbing contract, then $P(B)=\frac{3}{7}$ and

P(at least one of the contracts) = $P(A\cup B)=\frac{5}{7}$

Now we have to find $P(A\cap B)$ = ?

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

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

$P(A\cap B)=\frac{2}{7}+\frac{3}{5}-\frac{5}{7}$

$P(A\cap B)=\frac{6}{35}$ = P(both contracts)

Problem 9

The probability that Manzoor solves a problem is 3/7 and the probability that Rashid solves the same problem is 7/15. Find the probability that at least one of them solves the problem, assure that they attempt it independently.

Solution:

Let M = Manzoor solve the problem and R = Rashid solve the problem

At least one of them solves the problem means either (Manzoor or Rashid) or both solve the problem.

Since M and R are independent, they are not mutually exclusive.

$P(M \text{ or } R)=P(M)+P(R)-P(M \text{ and } R)$ or may be written as

$P(M\cup R)=P(M)+P(R)-P(M\cap R)$, then

$P(M\cup R)=\frac{3}{7}+\frac{7}{15}-\frac{3}{7}\times\frac{7}{15}=\frac{73}{105}$

Note: If two events are independents, they are not mutually exclusive.

CS CamScanner

Alternate Solution of Problem 9

P(at least one of them solves the problem)

= 1 – P(Neither Manzoor nor Rashid solves the problem)

$= 1 - P(M' \cap R')$

$= 1 - P(M') \cdot P(R')$

Since, $P(M) = \frac{3}{7}$ then $P(M') = \frac{4}{7}$

$P(R) = \frac{7}{15}$ then $P(R') = \frac{8}{15}$

Therefore:

P(at least one of them solves the problem) $= 1 - \frac{4}{7} \times \frac{8}{15}$

$= 1 - \frac{32}{105} = \frac{73}{105}$

Problem 10:

Let the population of adults in a small village be categorized according to sex and employment status as:

	Employed	*Unemployed*	*Total*
Male	460	40	500
Female	140	260	400
Total	600	300	900

If a person is chosen at random find the probability that it is:

(i) Male (ii) Employed person (iii) Male and Employed person

(iv) Male or Employed person

Solution:

$n(S) = {}^{900}C_1 = 900$

(i) Let M = person is a male

then $n(M) = {}^{500}C_1 = 500$ and $P(M) = \frac{n(M)}{n(S)} = \frac{500}{900} = \frac{5}{9}$

(ii) Let E = An employed person, then $n(E) = {}^{600}C_1 = 600$ and

$$P(E) = \frac{n(E)}{n(S)} = \frac{600}{900} = \frac{2}{3}$$

(iii) Since $P(M \cap E) = \frac{n(M \cap E)}{n(S)}$

Since $n(M \cap E) = {}^{460}C_1 = 460$ then $P(M \cap E) = \frac{460}{900} = \frac{46}{90} = \frac{23}{45}$

$$P(M \cup E) = P(M) + P(E) - P(M \cap E) = \frac{5}{9} + \frac{2}{3} - \frac{23}{45} = \frac{32}{45}$$

Problem 11

From a well shuffled pack of 52 playing cards, one card is drawn. Find the probability that it is

(i) a king (ii) a diamond (iii) the queen of heart

(iv) either the queen of heart or the jack of spades (v) either a two or a three

(vi) two or spade (vii) not the ace of spades (viii) not a club (ix) not a face card

(x) either a king, queen or a jack

Solution:

(i) $n(S) = {}^{52}C_1 = 52$

Let A = a King, then $n(A) = {}^{4}C_1 = 4$

$$P(A) = \frac{n(A)}{n(S)} = \frac{4}{52} = \frac{1}{13}$$

(ii) Let B = a diamond, then n(B) = $^{13}C_1 = 13$

$$P(B) = \frac{n(B)}{n(S)} = \frac{13}{52} = \frac{1}{4}$$

(iii) Let C = the queen of heart, then n(C) = $^{1}C_1 = 1$

$$P(C) = \frac{n(C)}{n(S)} = \frac{1}{52}$$

(iv) Let A = the queen of hearts, then n(A) = $^{1}C_1 = 1$

Let B = the jack of spades then n(B) = $^{1}C_1 = 1$

since A and B are mutually exclusive then $P(A \cup B) = P(A) + P(B) = \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)}$

$$P(A \cup B) = \frac{1}{52} + \frac{1}{52} = \frac{2}{52} = \frac{1}{26}$$

(v) Let A = a Two, then

n(A) = $^{4}C_1 = 4$ then $P(A) = \frac{n(A)}{n(S)} = \frac{4}{52}$

Let B = a Three, then

n(B) = $^{4}C_1 = 4$ then $P(B) = \frac{n(B)}{n(S)} = \frac{4}{52}$

Since A and B are mutually exclusive

then $P(A \cup B) = P(A) + P(B)$

$$P(A \cup B) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13}$$

(vi) Let A = a Two, then

n(A) = $^{4}C_1 = 4$ then $P(A) = \frac{n(A)}{n(S)} = \frac{4}{52} = \frac{1}{13}$

CS CamScanner

Let B = a Spade, then

$n(B) = {}^{13}C_1 = 13$ then $P(B) = \frac{n(B)}{n(S)} = \frac{13}{52} = \frac{1}{4}$

Since A and B are not mutually exclusive

$P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{1}{52}$

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

$P(A \cup B) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$

(vii) Let A = Ace of Spade, then $n(A) = {}^{1}C_1 = 1$

and $P(A) = \frac{n(A)}{n(S)} = \frac{1}{52}$

P(Not Ace of Space) $= P(A') = 1 - P(A) = 1 - \frac{1}{52} = \frac{51}{52}$

(viii) Let C = a Club, then $n(C) = {}^{13}C_1 = 13$

and $P(C) = \frac{n(C)}{n(S)} = \frac{13}{52} = \frac{1}{4}$

P(Not a Club) $= P(C') = 1 - P(C) = 1 - \frac{1}{4} = \frac{3}{4}$

(ix) F = Face card, then $n(F) = {}^{12}C_1 = 12$

and $P(F) = \frac{n(F)}{n(S)} = \frac{12}{52} = \frac{3}{13}$

P(Not a Face card) $= P(F') = 1 - P(F) = 1 - \frac{3}{13} = \frac{10}{13}$

(x) Let A = a King, then $n(A) = {}^{4}C_1 = 4$

Let B = a Queen, then $n(B) = {}^4C_1 = 4$

C = a Jack, then $n(C) = {}^4C_1 = 4$

and $P(A) = \frac{4}{52}$, $P(B) = \frac{4}{52}$, $P(C) = \frac{4}{52}$

and since A, B and C are mutually exclusive then

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$P(A \cup B \cup C) = \frac{4}{52} + \frac{4}{52} + \frac{4}{52} = \frac{12}{52} = \frac{3}{13}.$$

Problem 12

A bag contains 3 white and 5 red balls. One ball is drawn its colour noted and put back in the bag. Then a second ball is drawn. Find the probability that the two balls are of the same colour

Solution:

3 - White	5 - Red

= 8 balls

P(balls are of the same colour) = P(all white or all red) = P(all white) + P(all red)

$$= P(W_1 \cap W_2) + P(R_1 \cap R_2)$$

$$= P(W_1) \cdot P(W_2) + P(R_1) \cdot P(R_2)$$

$$= \frac{{}^3C_1}{{}^8C_1} \times \frac{{}^3C_1}{{}^8C_1} + \frac{{}^5C_1}{{}^8C_1} \times \frac{{}^5C_1}{{}^8C_1}$$

$$= \frac{9}{64} + \frac{25}{64} = \frac{34}{64} = \frac{17}{32}$$

Problem 13

A card is taken from a well-shuffled pack of cards. Then a second card is drawn. Find the probability that both are spades in case the first card

(i) is replaced before the second is drawn, and

(ii) is not replaced before the second is drawn

Solution:

(i) Let A = First Card is spade and B = Second card is spade

Since first card is replaced, therefore events are independent.

then $P(A\cap B)$ = P(both spade) $= P(A)\cdot P(B)$

$$=\frac{{}^{13}C_1}{{}^{52}C_1}\times\frac{{}^{13}C_1}{{}^{52}C_1}=\frac{13}{52}\times\frac{13}{52}=\frac{1}{4}\times\frac{1}{4}=\frac{1}{16}$$

(ii) Since first card is not replaced, therefore events are dependent.

$$P(A\cap B)=P(A)\cdot P(B|A)=\frac{{}^{13}C_1}{{}^{52}C_1}\times\frac{{}^{12}C_1}{{}^{51}C_1}=\frac{13}{52}\times\frac{12}{51}=\frac{1}{17}$$

Problem 14

Let A, B be the events in a sample space, such that

$P(A)=\frac{3}{4}$, $P(B)=\frac{1}{3}$ and $P(A\cup B)=\frac{5}{6}$

Find (i) $P(A\cap B)$ (ii) $P(A|B)$ (iii) $P(B|A)$

Solution:

(i) Since $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

then $P(A\cap B)=P(A)+P(B)-P(A\cup B)=\frac{3}{4}+\frac{1}{3}-\frac{5}{6}=\frac{3}{12}=\frac{1}{4}$

(ii) Since $P(A\cap B)=P(B)\cdot P(A|B)$

then $P(A|B)=\dfrac{P(A\cap B)}{P(B)}=\dfrac{\frac{1}{4}}{\frac{1}{3}}=\dfrac{3}{4}$

similarly $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$

Problem 15

Prove that:

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

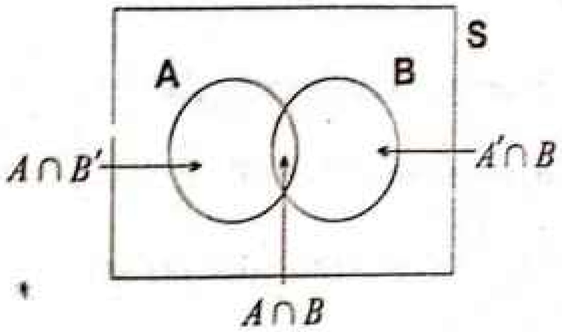
for two not mutually exclusive events A and B, and

(ii) $P(A \cup B) = P(A) + P(B)$

for two mutually exclusive events A and B.

Solution:

(i) Let A and B are two events as shown in the following Venn diagram



We have $A \cup B = A \cup [A' \cap B]$

Since A and $A' \cap B$ are disjoint (mutually exclusive)

then $P(A \cup B) = P(A) + P(A' \cap B)$. . . . . (1)

Now $B = (A \cap B) \cup (A' \cap B)$

where $A \cap B$ and $A' \cap B$ are mutually exclusive,

then $P(B) = P(A \cap B) + P(A' \cap B)$

and $P(A' \cap B) = P(B) - P(A \cap B)$ by putting in (1) we get

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

(ii) We know that, if two events A and B are mutually exclusive then $A \cap B = \phi$

and $P(A \cap B) = P(\phi) = 0$. Therefore, the formula becomes as

$$\boxed{P(A \cup B) = P(A) + P(B)}$$

Problem 16

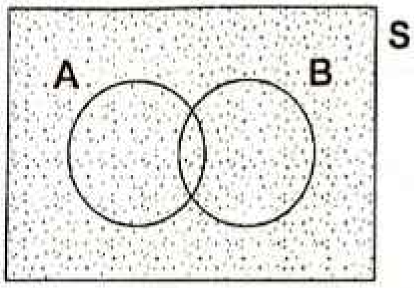
Prove that:

(i) $P(A \cap B) = P(A) \cdot P(B|A)$ for two dependent events A and B, and

(ii) $P(A \cap B) = P(A) \cdot P(B)$ for two independent events A and B.

Solution:

(i) Suppose a sample space S contains N equally likely elements (sample points) of which m_1 sample points belong to event A, m_2 sample points belong to event B and m_3 sample points belonging to both A and B (i.e. $A \cap B$) as shown in the following Venn diagram.



Then $P(A) = \dfrac{n(A)}{n(S)} = \dfrac{m_1}{N}$

$$P(B) = \frac{n(B)}{n(S)} = \frac{m_2}{N}$$

$$P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{m_3}{N}$$

and by the definition of conditional probability we know that

CamScanner

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{\frac{m_3}{N}}{\frac{m_1}{N}}=\frac{m_3}{m_1}$$

Now $P(A\cap B)=\frac{m_3}{N}\cdot\frac{m_1}{m_1}$ (multiplying and dividing by m_1)

$$=\frac{m_1}{N}\cdot\frac{m_3}{m_1}=P(A)\cdot P(B|A)$$

Hence $\boxed{P(A\cap B)=P(A)\cdot P(B|A)}$

(ii) If A and B are independent events, then

$P(A|B)=P(A)$ and $P(B|A)=P(B)$. Therefore, the formula for two independent events becomes as

$$\boxed{P(A\cap B)=P(A)\cdot P(B)}$$

EXERCISE – 3

3.1 If a six faced cubical die is thrown, what is the probability of an even number turning up, what is the probability of getting 4 or a higher number.

3.2 A coin is tossed two times, find the probability of getting

(i) One head **(ii)** Two heads **(iii)** At least one head

3.3 If a coin is tossed 2 times, what is the probability of getting the same face.

3.4 If three coins are tossed simultaneously, what is the probability of getting the same face.

3.5 Two dice are thrown simultaneously. What is the probability that the sum of the number thrown is

(i) less than 2 **(ii)** 9 **(iii)** Even

3.6 A pair of fair dice is rolled once. Find the probability of obtaining odd numbers on the two dice.

3.7 Determine the sample space, if three balanced coins are tossed together. Find the probability of at least one head in one throw.

3.8 If two dice are rolled, find the probability that:

(i) Two face shows the same number.

(ii) The sum of the two faces are greater than 10.

3.9 **(a)** Find the probability of appearing a total more than 7 when dice are thrown.

(b) A pair of fair dice is rolled once. What is the probability that two faces show the same number.

3.10 In throwing two dice, what is the probability that the sum of the dots on the top face of both the dice is:

CamScanner

(i) 9 (ii) not 9 (iii) greater than 2

Karachi Board 1992

3.11 **(a)** Define probability of an event.

(b) If a die is tossed twice, find the probability of getting a sum of seven dots.

Karachi Board 2002

3.12 Two fair dice are rolled at a time. What is the probability of getting 4 in one of the dice.

3.13 Determine the probability of getting a total of 8 in a single throw with two dice.

3.14 A pair of fair dice is tossed. What is the probability of getting a total of at most 4.

3.15 Two fair dice are thrown. Find the probability of

(a) Obtaining a pair of similar numbers

(b) Obtaining sum of Nos. on two dice as '10' or more.

(c) Obtaining sum of Nos. on two dice as less than '10'.

(d) Obtaining '4' in one of the dice

(e) Obtaining less than '12' as sum of No. on the two dice

3.16 Suppose two dice are thrown. What is the probability that the number on of the first die is even that of the second is an odd number.

3.17 Two dice are thrown. Find the probability that

(i) the first die shows 4

(ii) the total of the numbers on the dice is 9 or greater than 9.

(iii) the number on the first die is greater than the number of the second die.

3.18 **(a)** A fair die is tossed twice, calculate the probability of getting

(i) a total of 4 (ii) the sum of the numbers on the two faces is a prime number

Karachi Board 2002

(b) Two dice are rolled together. X is the sum of the dots on both the dice. Find the following probabilities.

CamScanner

(i) $P(X \geq 7)$ (ii) $P(X \leq 4)$ (iii) $P(X = 8)$

Karachi Board 2001

3.19 An Unbiased coin is thrown 3 times. Enlist the members of the sample space and hence find the probability of obtaining at the most one head.

Karachi Board 1994

3.20 A pair of fair dice is rolled once. What is the probability of obtaining

(i) The same numbers on the two dice.

(ii) a total of at least 5 on the two dice.

Karachi Board 1988

3.21 What is the probability of getting a total of seven points. If we throw two dice.

Karachi Board 1987

3.22 A pair of unbiased dice is rolled once. Find the probability of obtaining a total of seven on the two dice.

Karachi Board 1989

3.23 A bag contains 5 green and 10 white balls. If one ball is drawn from it, find the chance that the ball drawn is green.

3.24 A card is drawn from a pack of cards. Find the probability that it is

(i) a Jack (ii) a Spade

3.25 (a) 3 items are selected at random from 15 items out of which 6 are defective. Find the probability that one is defective.

(b) A class has 12 boys and 4 girls. If two students are selected at random from the class, find the probability that both are boys.

Karachi Board 2000 (Sup.)

3.26 In a group of 20 students, 13 students have passed in statistics. If a student is selected at random from this group what is the probability that the student has passed in statistics.

3.27 A box contains 5 blue and 8 green balls. If two balls are selected at random from this box, what is the probability that they are

CamScanner

(i) both blue (ii) both green (iii) blue and green.

3.28 There are three economists, four engineers, two statisticians and one doctor. A committee of four is to be formed from among them. Find the probability that the committee consists of one member of each group.

Karachi Board 2000 (Sup.)

3.29 What is the probability of getting 3 white balls in a draw of 3 balls from a box containing 6 white and 5 red balls.

3.30 Three cards are drawn from a full pack of cards. What is the probability that all the three are red.

3.31 A card is drawn at random from a full pack of cards. Find the probability that the card is not a face card.

3.32 A bag contains 3 red and 5 black balls. Two balls are drawn at random. Find the probability that:

(i) The balls are of different colour.

(ii) The balls are of the same colours.

Karachi Board 2001

3.33 Of 12 eggs in a refrigerator 2 are bad. From these eggs, four eggs are taken out at random to make a cake. What is the probability that exactly one egg is bad?

Karachi Board 1991

3.34 A bag contains 4 white and 3 black balls and two balls are drawn at random, find the probability that

(i) both are white

(ii) both are black

(iii) one white and one black

Karachi Board 1988, 2008

3.35 An urn contains 3 red, 4 white and 5 black balls. Three balls are drawn at random from the urn. Find the probability that

(i) all are black (ii) all are of different colours.

CamScanner

3.36 A bag contains 3 white balls, 4 black balls and 5 red balls. If 3 balls are drawn at random, determine the probability that

(i) all 3 are red (ii) all 3 are black (iii) 2 are red and one is black

(iv) one of each colour is drawn

3.37 Out of 20 students in a statistics class, 3 students are failing in the course. If 4 students from the class are picked up at random, what is the probability that one of the failing student will be among them.

3.38 There are 8 articles of which 2 are defective and 6 non-defective. If a sample of 5 is selected, find the probability of getting exactly 4 non defectives

3.39 Five cards are drawn from a pack. What is the probability of its being 4 kings and remaining card be other.

3.40 A committee of four members is to be selected at random from a group consisting of six men and four women. What is the probability that on this committee men and women will get equal representation?

3.41 In a bag there are 4 white and 10 yellow balls. Two balls are drawn at random. What is the probability that of these two balls, one is white and the other yellow.

3.42 A bag contains 7 red, 12 white and 4 green balls. What is the probability that

(i) 3 balls drawn are all white and

(ii) 3 balls drawn are one of each colour.

3.43 A bag contains 3 red, 4 white 3 green balls. If a sample of 3 balls is selected at random, what is the probability that:

(i) All are of different colour (ii) Two are of white colour

3.44 Find PROBABILITY, for each of the following events:

(i) The sum 8 appears, in a single toss of a pair of Two dice.

(ii) At least one head appears in Four tosses of a coin.

(iii) Exactly one king appears, if two playing cards are drawn from a pack of 52 cards.

3.45 Four persons are chosen at random from a group containing 3 men, 2 women and 4 children. Compute the probability that exactly two of them will be children.

CS CamScanner

3.46 A box contains 6 red 4 green and 5 white balls. Six balls are drawn simultaneously at random from the box. What is the probability that

(i) 2 red, 2 green and 2 white balls are drawn (ii) 4 are red.

3.47 A bag contain 5 green, 7 black and 3 red handkerchiefs. Three handkerchiefs are drawn simultaneously at random. What is the probability that one green, one black and one red handkerchief will be drawn.

3.48 Four cards are drawn from a full pack of cards. Find the probability that two are spades and two are hearts.

3.49 If two dice are thrown simultaneously, What is the probability of obtaining a sum of 5 or a sum of 9.

3.50 The following four events are defined to contain certain numbers

$A = \{5, 8, 9\}$, $B = \{1, 3, 5\}$, $C = \{1, 2, 3, 4\}$, $D = \{7, 9\}$

Which of the following are mutually exclusive

(i) A and B (ii) A and C (iii) A and D

(iv) B and C (v) B and D (vi) C and D

3.51 (a) Define Sample Space and Events, giving examples.

(b) If E and F are two mutually exclusive events, then prove that

$P(E \text{ or } F) = P(E) + P(F)$.

(c) A card is drawn at random from a pack of playing cards. What is the probability that it is either a 'Heart' or 'The Queen of Spades'?

Karachi Board 2002

3.52 (a) One card is drawn from a pack of 52 cards. What is the chance that the card is either a King or Queen.

(b) Define Mutually Exclusive Events, giving examples.

(c) A fair coin is tossed four times. Calculate the probability of getting three heads on them. Use the sample space.

Karachi Board 2003

3.53 If A and B are mutually exclusive events and

CamScanner

$P(A) = 0.4$ and $P(B) = 0.5$, find

(i) $P(A \cup B)$ (ii) $P(A \cap B)$ (iii) $P(A')$

3.54 A marble is drawn at random from a box containing 10 red, 30 white, 20 blue and 15 orange marbles. Find the probability that it is:

(i) Orange or red

(ii) not blue

(iii) white

(iv) red, white or blue

Karachi Board 1998

3.55 (a) Define Sample Space and Events, giving examples.

(b) If E and F are two mutually exclusive events, then prove that

P(E or F) = P(E) + P(F)

Karachi Board 2002

3.56 Three applicants are to be selected out of 6 boys and 4 girls. What is the probability of selecting

(i) all boys (ii) at least one boy.

Karachi Board 1992

3.57 A box contains 4 red, 4 white and 5 green balls. Three balls are drawn together at random from the box, find the probability that they may be

(i) all of different colours

(ii) all of the same colour

(iii) 2 white balls

Karachi Board 1996

3.58 Define Probability. State and prove the Addition Law of Probability.

Karachi Board 2002

3.59 One ball is drawn at random from a bag containing 5 red balls, 4 black balls and 3 green balls. Find the probability that the ball drawn is

CamScanner

(i) Black (ii) Not Red (iii) Red or Green

3.60 (a) There are four mangoes, two bananas and three oranges, if 3 fruits are picked at random. Find the probability that:

(i) All are different fruits (ii) At least two mangoes

(b) Find the probability for each of the following events:

(i) The sum 8 appears when two dice are thrown

(ii) Two heads appear when 4 fair coins are tossed

(iii) A king and a queen appear when two cards are drawn from a well shuffled deck of cards

(iv) Two balls drawn from a bag containing 5 white and 4 black balls are either white or black.

3.61 A room has three sockets for bulbs. From a collection of 10 light bulbs of which 6 are defective, a person selects 3 at random and puts them in the sockets. What is the probability that the room is lighted.

3.62 One ticket is drawn at random from a set of 20 tickets numbered 1 to 20. What is the probability that the number is a multiple of 3 or 7.

3.63 A card is selected at random from an ordinary well-shuffled pack of 52 cards. What is the probability of getting

(a) A King (b) A Spade (c) A King or an Ace (d) A Picture card

3.64 Two cards are drawn at random from a well shuffled pack of 52 cards. What is the probability that

(a) both are aces (b) both are red (c) at least one is an ace

3.65 A die is tossed once, what is the probability of getting either a number less than 5 or an even number occurs on the die.

3.66 A department of commerce enrolls 125 students out of which 50 are girls. If 15 girls and 25 boys of the department offer advance statistics as an additional subject, what is the probability that a student selected a random from the department is a girl or a student offering advance statistics.

CS CamScanner

3.67 (a) From the deck of 52 Cards, one is selected at random. What is the probability that it is a heart (H) or a Face card (F).

(b) A card is drawn at random from a deck of 52 cards. What is the probability that it is either a jack or a heart.

3.68 Prove the general addition theorem of Probability. Given $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{3}$ and $P(A+B)=\frac{3}{4}$ then obtain $P(A\cap B)$.

Karachi Board 2000

3.69 (a) State and prove Addition Law for mutually exclusive events.

(b) From a deck of 52 cards a card is drawn at random. What is the probability that it is a King or Club.

Karachi Board 1998

3.70 The probability that a student will succeed in Mathematics and Physics is 6/9 and 8/13 respectively. What is the probability that he will succeed in at least one subject.

Karachi Board 2002

3.71 (a) Define Probability. State and prove the Addition Law of Probability.

(b) From a well shuffled deck of 52 cards a card is drawn at random. What is the probability that it is either a Jack or a diamond.

Karachi Board 1999

3.72 Prove that $P(A\cup B)=P(A)+P(B)-P(A\cap B)$. Also criticize if $A\cap B=$Null Set.

Karachi Board 2000

3.73 (a) State and prove the Law of Addition of Probability for two Not-Mutually Exclusive Events.

(b) A card is selected at random from a deck of 52 cards. Find the probability that it is either a King or a red card.

Karachi Board 1995

CamScanner

3.74 (a) State and prove Addition Law of Probability for two non-mutually exclusive events.

(b) A can solve 75% and B can solve 70% of the problems in a given book. What is the probability that a problem chosen at random will be solved by at least one of them.

Karachi Board 1989

3.75 (a) State and prove the Addition theorem on probability for not mutually exclusive events.

(b) The probability of solving any problem from this paper by Mr. A is 4/5 and by Mr. B is 3/4. A problem is chosen at random from this paper find the probability that:

(i) Either Mr. A or Mr. B will solve the problem

(ii) Both of them will solve the problem

(iii) Neither of them will solve the problem.

Karachi Board 1986

3.76 (a) Define Probability. State and prove Addition Law of Probability.

(b) Find the probability of a 4 turning up at least once in two tosses of a fair die.

Karachi Board 1984

3.77 For two events A and B, P(A) = 0.5, P(B) = 0.6 and P(both A and B) = 0.4 . Find P(B′) and P(A or B).

3.78 (a) If A and B are two Not Mutually Exclusive Events, show that

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

Karachi Board 1993, 2002, 2003, 2007 (Supp.).

(b) A drawer contains 50 bolts and 100 nuts. Half of the bolts and half of the nuts are rusted. If an item from the drawer is chosen at random, what is the probability that:

(i) It is rusted and is a bolt

(ii) It is rusted or is a bolt

Karachi Board 1992

CS CamScanner

3.79 (a) State and prove Addition law of Probability for two non-mutually exclusive events.

(b) A class consists of 100 students out of which 60 are boys and 40 girls. 37 boys and 18 girls of the class take Statistics as their optiona' subject. A student is chosen at random from the class, what is the probability that the student chosen be either a girl or one taking Statistics.

Karachi Board 1999

3.80 (a) State Classical Probability, giving its obvious deficiencies.

(b) In a certain federal prison it is known that $\frac{2}{3}$ of the inmates are under 25 years of age. It is also known that $\frac{3}{5}$ of the inmates are male and that $\frac{5}{8}$ of the inmates are female or of 25 years of age or older. What is the probability that a prisoner selected at random from this prison is female and at least 25 years old or older.

Karachi Board 1996

3.81 (a) What do you mean by Probability? Give the definition of Probability.

(b) State and Prove Addition Law of Probability.

Karachi Board 1985

3.82 Following information are given about 100 applicants for the post of Accountant in a company.

	Married	Un-married
Graduate	10	40
Post Graduate	35	15

An Applicant was selected. Find out probability that selected candidate

(i) is a Graduate and Un-married

(ii) is a Graduate or Un-married

Karachi Board 2000

CS CamScanner

3.83 A person is known to hit the target in three out of four shots, whereas an other person is known to hit the target in two out of three shots. Find the Probability of the target being hit at all when they both try.

Karachi Board 2000

3.84 An auto engineer has found that a car will require a tune-up with probability 0.6, a brake-job with probability 0.1 and both with probability 0.02. What is the probability that a car requires either a tune-up or a brake-job or both?

3.85 (a) It is known that A can hit the target 4 times out of every 6 shots, B can hit 3 times out of every 4 shots and C can hit a target 5 times out of every 6 shots. They fire a shot each. Find the probability that the target is hit at all.

Karachi Board 1993

(b) A card is drawn from an ordinary pack of 52 playing cards. Find the probability that the card is (i) a club or a diamond (ii) a club or a king

Karachi Board 2001

3.86 Of 150 people, 60 men and 90 are women, 30 of the men and 60 of the women read newspaper daily. A person is selected a random from the lot. What is the chance that selected would be a man or read newspaper?

Karachi Board 2002

3.87 (a) If $P(A) = 0.4$, $P(B) = 0.3$ and $P(A \cap B) = 0.1$ Find (i) $P(A')$ (ii) $P(A \cup B)$

(b) Two balanced dice are rolled simultaneously. 'A' denotes the event that the sum appearing on them is 8 and 'B' denotes the event that the two dice have the same number; find $P(A)$, $P(B)$, $P(A \cap B)$ and $P(A \cup B)$

Karachi Board 2003

3.88 Twenty cards have been numbered from 1 to 20. The cards are shuffled and a card is drawn.

(i) What is the probability that the number so obtained is a multiple of 5 or 7.

(ii) What is the probability that the number so obtained is a multiple of 3 or 5.

3.89 Two playing cards are drawn from a pack of 52 cards. Find out the probability of getting

CamScanner

(i) either Aces or Kings (ii) either Aces or Black Cards

3.90 The probability that A can solve a problem is 2/3 and B can solve it is 4/5, if both try what is the probability that the problem is solved?

3.91 (a) Define the following terms:

(i) Mutually Exclusive Events

(ii) Independent Events

(iii) Equally Likely Events

Karachi Board 2003

(b) The probability that a student passes Mathematics is $\frac{2}{3}$ and the probability that he passes English is $\frac{4}{9}$. If the probability of passing at least one course is $\frac{4}{5}$, what is the probability that he will pass both courses.

Karachi Board 2007 (Supp.)

3.92 (a) If $P(A)=\frac{5}{12}$, $P(B)=\frac{1}{2}$, $P(A\cup B)=\frac{19}{24}$ and $P(A\cap B)=\frac{1}{8}$

State whether events A and B are

(i) mutually exclusive (ii) independent

(b) Tickets numbered from 1 to 100, are well shuffled and a ticket is drawn from them at random. What is the probability that the drawn ticket has:

(i) an odd number less than 20 (ii) the number 15 or the multiples of 15

(iii) the number 1 or the multiples of 1.

Karachi Board 1991

3.93 (a) Suppose the event E and F are independent and that

$P(E)=\frac{1}{3}$, $P(F)=\frac{1}{4}$

Determine: (i) P(E and F) (ii) P(E or F)

Karachi Board 2011 (Supp.)

(b) A coin and a die are tossed independently, compute the probability of observing a head on the coin and 2 or 4 on the die.

Karachi Board 2002

(c) A and B are two events associated with an experiment. Suppose that $P(A)=\frac{2}{5}$, $P(A\cup B)=\frac{7}{10}$ and $P(B)=p$

(i) What is the value of p if A and B are mutually exclusive.

(ii) What is the value of p if A and B are independent.

Karachi Board 2007

3.94 (a) Candidate "A" can solve 70% problems of this question paper while another candidate "B" can solve 60% problems. A problem is selected at random. Find the probability that it can be solved by: (i) A and B (ii) A or B

(b) The tickets numbered from 1 to 100, are well shuffled and a ticket is drawn from them at random. What is the probabilty that the drawn ticket has:

(i) an odd number less than 20 (ii) the number 25 or the multiples of 25

(iii) the number 1 or multiples of 1.

Karachi Board 2000 (Sup.)

3.95 A box contains 6 red 3 black balls. Two balls are drawn in succession. What is the probability that both the balls are of same colour. If the 1st ball drawn is

(i) replaced **(ii)** not replaced.

3.96 Two cards are drawn form a well shuffled pack of 52 cards. Find the chance that they are both Diamonds if the first card is **(i)** replaced **(ii)** not replaced

3.97 Two card are drawn from a well shuffled deck of 52 cards. Find the probability that they are both aces, if the first card is **(i)** replaced **(ii)** not replaced

CamScanner

3.98 A bag consists of 2 red, 3 white and 4 green balls. Two balls are drawn randomly without replacement. Find the probability that:

(i) Both are of the same colour **(ii)** Both are of different colours

3.99 (a) If A and B are two events such that $P(A)=\frac{2}{5}$, $P(A|B)=\frac{1}{2}$ and $P(B|A)=\frac{2}{3}$

Find (i) $P(A\cap B)$ (ii) $P(B)$ (iii) $P(A\cup B)$

Karachi Board 2010

(b) (i) Define Probability.

(ii) Differentiate between Independent and Dependent events and give examples.

(iii) From a pack of 52 well shuffled cards, two cards are drawn successively at random without replacement. Find the probability that first card is an ace and the second card is a King.

Karachi Board 1991

3.100 If $P(A|B)=\frac{2}{5}$, $P(A)=\frac{1}{3}$, $P(B)=\frac{1}{4}$

Find (i) $P(A\cap B)$ (ii) $P(B|A)$

3.101 (a) Given : P(A) = 0.60, P(B) = 0.40, $P(A\cap B)=0.24$ Find $P(A|B)$ and $P(A\cup B)$

(b) If A and B are independent events and P(A) = 0.3 and P(B) = 0.5; find $P(A\cup B)$

Karachi Board 2000 (Sup.)

3.102 If two events A and B are not independent, prove that

$$P(A\cap B)=P(A)\cdot P(B|A)=P(B)\cdot P(A|B)$$

Karachi Board 1994

3.103 (a) A and B are two dependent events and P(A) = 0.25, P(B) = 0.33 and $P(A\cap B)=0.15$ Find (i) $P(A\cup B)$ (ii) $P(B|A)$

(b) (i) Explain the Conditional Probability of Independent events with examples.

(ii) Prove the Addition Law of Probability for two non mutually exclusive events A and B.

Karachi Board 2004

CS CamScanner

3.104 A and B play 12 games of chess of which 6 are won by A, 4 are won by B and 2 end in a draw. They agree to play a tournament consisting 3 games. Find the probability that:

(i) A and B win alternately.

(ii) B wins at least one game.

Karachi Board 2003

3.105 (a) What do you mean by Probability.

(b) Two fair dice are thrown, find the probability of obtaining 4 in one of the dice.

(c) A bag contains 3 white and 5 black balls. One ball is drawn from the bag looked at and replaced. Then a second ball is drawn. What is the probability that both the balls are of the same colour

3.106 State and prove Multiplication law of Probability for two dependent events.

Karachi Board 2001

3.107 (a) Define Probability. State and prove the Multiplication Law of Probability.

(b) Three cards are drawn in succession without replacement from an ordinary deck of 52 cards. Find the probability that the first card is a red ace, the second a ten or jack and third is greater than 3 but less than 7.

Karachi Board 2003

3.108 (a) Define Conditional Probability.

(b) The probability that a missile 'A' hits the target is 0.8 and the probability that another missile 'B' hits the target is 0.85. If both the missiles are fired, find the probability that at least one will not hit the target.

Karachi Board 1997

3.109 There are 40 nuts and 20 bolts in a drawer. 15 nuts and 5 bolts are rusted. If an item is selected at random, what is the probability that:

(i) It is rusted or a bolt

(ii) It is a nut if the item selected is rusted.

Karachi Board 1997

CamScanner

Chapter – 4

Probability Distributions

Before studying probability distributions, it is necessary to introduce first, what is a random variable.

Random Variable

A random variable is a numerical quantity whose value depends on chance.

OR

If a variable (X) assumes different numerical values, each with a given probability, then it is called a random variable.

For example, an experiment of tossing a coin two times has four possible outcomes HH, HT, TH and TT and suppose we are interested in the number of heads. Let the variable X denote the number of heads in this experiment, then the possible values of X are 2, 1 and 0. This is a random variable because each value of the variable (X) is attached or related with probability.

Random variable may be discrete or continuous. If the random variable takes on the integer values (i.e. the values in whole numbers) such as 0, 1, 2, 3, ..., then it is called a discrete random variable. For example, the number of defective ball point pens in each carton of 12, the number of printing mistakes in each page of a book, the number of daily admissions to a general hospital, the number of accidents in a month at a place, etc,

If a random variable can assume any value within a given interval, it is called a continuous random variable, it has infinite possible values. For example, height of a person, weight of a baby, temperature at a place, etc.

CS CamScanner

Probability Distribution

For a discrete random variable, a table or a formula showing all possible values of a random variable, along with their corresponding probabilities is called a probability distribution of the random variable.

It is to be noted that in any probability distribution, the sum of all probabilities should be equal to unity (i.e. one).

Example 4.1

Suppose in an experiment a fair coin is tossed three times, then

S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

Let the random variable X showing the number of Heads, then X can take 0, 1, 2 and 3 possible values.

Now for each value of X we find probabilities as shown in the following table.

x	0	1	2	3
P(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Since the sum of all probabilities is unity, therefore we can say that we have a probability distribution of No. of Heads. Since the values of random variable are in whole numbers, therefore, it is called discrete probability distribution.

Example 4.2

Suppose in an experiment a die is rolled only once, then the sample space is

S = {1, 2, 3, 4, 5, 6}

If the random variable X represents the number of spots that appear on the die then X takes the values 1, 2, 3, 4, 5 and 6 with respective probabilities $\frac{1}{6}$ for each. Thus the values of X with their respective probabilities define the discrete probability distribution of X. Probability distribution is given below:

CamScanner

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

Example 4.3

A bag contains two white and three black balls. Two balls are selected at random. Find the probability distribution for the number of white balls.

Solution:

Let $X = $ No. of white balls, then the possible values of $X = 0, 1, 2$

Now find probabilities for $X = 0, 1, 2$

$$P(x=0) = P(\text{no white ball}) = \frac{\left({}^2C_0\right)\left({}^3C_2\right)}{{}^5C_2} = \frac{3}{10}$$

$$P(x=1) = P(\text{one white ball}) = \frac{\left({}^2C_1\right)\left({}^3C_1\right)}{{}^5C_2} = \frac{6}{10}$$

$$P(x=2) = P(\text{two white balls}) = \frac{\left({}^2C_2\right)\left({}^3C_0\right)}{{}^5C_2} = \frac{1}{10}$$

Therefore, the probability distribution for the number of white balls is:

x	0	1	2
P(x)	$\frac{3}{10}$	$\frac{6}{10}$	$\frac{1}{10}$

Note: A discrete probability distribution is also called "Probability Mass Function".

A Probability Mass Function can also be represented in equation form (i.e. in the form of formula) instead of tabular form.

CamScanner

Example 4.4

Suppose an unbiased coin is tossed 3 times, then find probability distribution of the random variable "No. of Heads" in the following forms:

(a) In Tabular Form (b) In Equation Form

Solution:

(a) In Tabular Form

Since S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

Let X = No. of Heads = 0, 1, 2 and 3

then, $P(x = 0) = P(\text{all Tails}) = \frac{1}{8}$

$P(x = 1) = P(1 - \text{Head and } 2 - \text{Tails}) = \frac{3}{8}$

$P(x = 2) = P(2 - \text{Heads and } 1 - \text{Tail}) = \frac{3}{8}$

$P(x = 3) = P(3 \text{ Heads}) = \frac{1}{8}$

Therefore, the probability distribution (i.e. probability mass function) in tabular form is

x	0	1	2	3
P(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

(b) In Equation Form

Since total No. of cases = 8 and the random variable X may take the values 0, 1, 2 and 3

Then, $P(X = 0) = \frac{{}^3C_0}{8}$; by definition of probability

$P(X = 1) = \frac{{}^3C_1}{8}$

$$P(X=2) = \frac{^3C_2}{8}$$

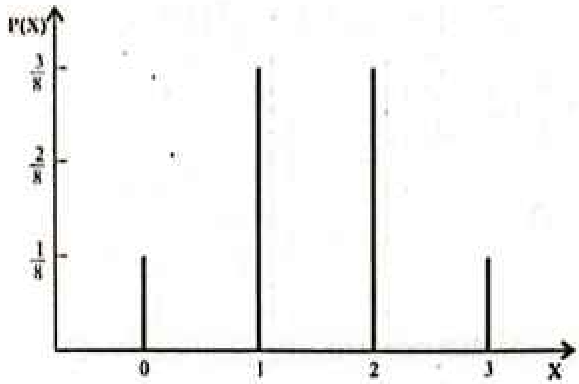
$$P(X=3) = \frac{^3C_3}{8}$$

In general $P(X=x) = \frac{^3C_x}{8}$, where $x = 0, 1, 2$ and 3

Therefore, the probability distribution (i. e. probability mass function) in the form of equation or in the form of formula is:

$$\boxed{P(X=x) = \frac{^3C_x}{8},\ x = 0, 1, 2 \text{ and } 3}$$

Note: The above probability distribution can also be represented in the form of graph as shown below:



Mathematical Expectation

Suppose we have following probability distribution of a discrete random variable X.

x	x_1	x_2	x_3	....	x_n
$P(x)$	$P(x_1)$	$P(x_2)$	$P(x_3)$	....	$P(x_n)$

The mathematical expectation or expected value of random variable x is defined as

$$E(X) = x_1P(x_1) + x_2P(x_2) + \ldots\ldots\ldots\ldots + x_nP(x_n)$$

$$E(X) = \sum_{i=1}^{n} x_i P(x_i) \quad \text{or simply may be written as}$$

$$\boxed{E(X) = \sum x P(x)}$$

Note: It should be noted that E(X) is the average value of the random variable X.

Example 4.5

Find the expected value of X in the following probability distribution.

x	0	1	2
$P(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Solution:

$$E(X) = \sum x P(x) = x_1P(x_1) + x_2P(x_2) + x_3P(x_3)$$

$$= 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4}$$

$$= 0 + \frac{1}{2} + \frac{1}{2} = 1$$

Example 4.6

For the following probability distribution

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

Find (i) E(X) (ii) $E(X^2)$

Solution:

(i) $E(X) = \sum x P(x) = x_1P(x_1) + x_2P(x_2) + x_3P(x_3)$

CamScanner

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

$$= \frac{3}{6} + \frac{4}{2} + \frac{5}{3} = \frac{25}{6} = 4.17$$

$$\boxed{E(X) = 4.17}$$

(ii) $E(X^2) = \sum x^2 P(x) = x_1^2P(x_1) + x_2^2P(x_2) + x_3^2P(x_3)$

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

$$E(X^2) = \frac{9}{6} + \frac{16}{2} + \frac{25}{3} = 17.8$$

$$\boxed{E(X^2) = 17.8}$$

Mean and Variance of a Random Variable

The mean of a discrete random variable X is defined as

$$\boxed{\text{Mean of } x = E(X) = \sum xP(x)}$$

The variance of discrete random variable X is defined as

$$\boxed{\text{Variance of } X = V(X) = E[X - E(X)]^2 = E(X^2) - [E(X)]^2}$$

where $E(X^2) = \sum x^2 P(x) = x_1^2P(x_1) + x_2^2P(x_2) + \dots\dots\dots\dots + x_n^2P(x_n)$

and $[E(X)]^2 = \left[\sum xP(x)\right]^2 = [x_1P(x_1) + x_2P(x_2) + \dots\dots\dots\dots + x_nP(x_n)]^2$

Note: Standard Deviation of $x = \sqrt{V(x)}$

Example 4.7

The random variable X has the following probability distribution

x	1	2	3
$P(x)$	0.25	0.5	0.25

Calculate Mean and Variance of X

Solution:

Mean of X = $E(X) = (1)(0.25) + (2)(0.5) + (3)(0.25)$

Mean of X = $0.25 + 1.0 + 0.75 = 2$

Since $V(X) = E(X^2) - [E(X)]^2$

where $E(X) = 2$ and $[E(X)]^2 = (2)^2 = 4$

Now compute $E(X^2) = \sum x^2 P(x)$

$$E(X^2) = (1)^2(0.25) + (2)^2(0.5) + (3)^2(0.25)$$

$$E(X^2) = 0.25 + 2.0 + 2.25 = 4.5$$

Then $V(X) = 4.5 - 4 = 0.5$

Example 4.8

The probability distribution of the random variable X is given below

x	0	1	2	3	4
$P(x)$	0.05	0.3	0.4	0.2	0.05

Calculate Mean and Standard Deviation.

CamScanner

Solution:

x	$P(x)$	$xP(x)$	$x^2P(x)$
0	0.05	0	0
1	0.30	0.30	0.3
2	0.40	0.85	1.6
3	0.20	0.60	1.8
4	0.05	0.20	0.8
Total	1	1.90	4.5

Mean of X $= E(X) = \sum xP(x) = 1.9$

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

where $E(X^2) = \sum x^2 P(x) = 4.5$ and $E(x) = 1.9$

Now $V(X) = 4.5 - (1.9)^2 = 0.89$

Standard Deviation of $X = \sqrt{V(X)}$

Hence S. D. of $X = \sqrt{0.89} = 0.94$

Example 4.9

The probability distribution of a random variable X is given below

$$P(x) = \frac{{}^3C_x}{8}, \quad x = 0, 1, 2, 3$$

Find mean and variance of the random variable X

Solution:

Mean of X $= E(X) = \sum xP(x) = \sum_{x=0}^{3} x \frac{{}^3C_x}{8}$

$$= 0 \times \frac{{}^3C_0}{8} + 1 \times \frac{{}^3C_1}{8} + 2 \times \frac{{}^3C_2}{8} + 3 \times \frac{{}^3C_3}{8}$$

CamScanner

Mean of X $= 0 + \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = \frac{12}{8} = 1.5$

To find V(X), we find first $E(X^2)$

$$E(X^2) = \sum x^2 P(x) = \sum_{x=0}^{3} x^2 \frac{{}^3C_x}{8}$$

$$= (0)^2 \times \frac{{}^3C_0}{8} + (1)^2 \times \frac{{}^3C_1}{8} + (2)^2 \times \frac{{}^3C_2}{8} + (3)^2 \times \frac{{}^3C_3}{8}$$

$$= 0 + \frac{3}{8} + \frac{12}{8} + \frac{9}{8} = \frac{24}{8} = 3$$

Then $V(X) = E(X^2) - [E(X)]^2$

$$= 3 - (1.5)^2 = 0.75$$

Properties of Mathematical Expectation

(i) $E(c) = c$, where c is any constant

(ii) $E(aX) = aE(X)$, where a is any constant

(iii) $E(X \pm b) = E(X) \pm b$, where b is any constant

(iv) $E(aX \pm b) = aE(X) \pm b$, where a and b are some constants

(v) $E[X - E(X)] = 0$,

(vi) $E(X \pm Y) = E(X) \pm E(Y)$,

(vii) $E(XY) = E(X) \cdot E(Y)$, if X and Y are independent

CamScanner

MISCELLANEOUS SOLVED PROBLEMS ON EXPECTATION

Problem 1

For the following probability distribution of the random variable X

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

Find: (i) $E(X)$ (ii) $E(X^2)$ (iii) $E(X-2)$ (iv) $E(2X)$

(v) $E\left(\frac{X}{2}\right)$ (vi) $E(X+3)^2$

Solution:

(i) $E(X)=\sum xP(x)=3\times\frac{1}{4}+4\times\frac{1}{2}+5\times\frac{1}{4}$

$E(X)=\frac{3}{4}+\frac{4}{2}+\frac{5}{4}=4$

(ii) $E(X^2)=\sum x^2P(x)=3^2\times\frac{1}{4}+4^2\times\frac{1}{2}+5^2\times\frac{1}{4}$

$E(X^2)=\frac{9}{4}+\frac{16}{2}+\frac{25}{4}=16.5$

(iii) $E(X-2)=E(X)-2=4-2=2$

(iv) $E(2X)=2E(X)=2\times 4=8$

(v) $E\left(\frac{X}{2}\right)=\frac{1}{2}\cdot E(X)=\frac{4}{2}=2$

(vi) $E(X+3)^2=E(X^2+6X+9)$

CamScanner

$E(X+3)^2 = E(X^2) + 6E(X) + 9$

$E(X+3)^2 = 16.5 + 6 \times 4 + 9 = 49.5$

Problem 2

Prove that E(c) = c, where c is any constant.

Solution:

Suppose we have following probability distribution for X = c

where $x_1 = c, x_2 = c, \ldots\ldots, x_n = c$

X = c	$x_1 = c$	$x_2 = c$	....	$x_n = c$
$P(x)$	$P(x_1)$	$P(x_2)$	....	$P(x_n)$

Since $E(X) = \sum x P(x) = x_1P(x_1) + x_2P(x_2) + \ldots\ldots + x_nP(x_n)$

then $E(C) = (c)P(x_1) + (c)P(x_2) + \ldots\ldots + (c)P(x_n)$

$E(C) = cP(x_1) + cP(x_2) + \ldots\ldots + cP(x_n)$

$E(C) = c[P(x_1) + P(x_2) + \ldots\ldots + P(x_n)]$

$E(C) = c[1] = c$ Since $\sum P(x) = 1$

Hence $E(C) = c$

Problem 3

Prove that $E(aX) = aE(X)$ where a is any constant.

Solution:

$$E(aX) = \sum_{i=1}^{n} ax_iP(x_i) = a\sum_{i=1}^{n} x_iP(x_i)$$

$E(aX) = aE(X)$

Problem 4

Prove that $E(X + b) = E(X) + b$ where b is any constant.

CS CamScanner

Solution:

$$E(X+b)=\sum_{i=1}^{n}(x_i+b)P(x_i)$$

$$E(X+b)=\sum_{i=1}^{n}x_iP(x_i)+\sum_{i=1}^{n}bP(x_i)$$

$$E(X+b)=\sum_{i=1}^{n}x_iP(x_i)+b\sum_{i=1}^{n}P(x_i)$$

$$E(X+b)=E(X)+b \quad \text{where} \sum_{i=1}^{n}P(x_i)=1$$

Problem 5

Prove that $E(aX-b)=aE(X)-b$ where a and b are some constants.

Solution:

$$E(aX-b)=\sum_{i=1}^{n}(ax_i-b)P(x_i)$$

$$E(aX-b)=\sum_{i=1}^{n}ax_iP(x_i)-\sum_{i=1}^{n}bP(x_i)$$

$$E(aX-b)=a\sum_{i=1}^{n}x_iP(x_i)-b\sum_{i=1}^{n}P(x_i)$$

$$E(aX-b)=a\,E(X)-b \qquad \text{where} \sum_{i=1}^{n}P(x_i)=1$$

Problem 6

Prove that $E(X+Y)=E(X)+E(Y)$

Solution:

Let X and Y are two discrete random variable having following probability distribution.

CS CamScanner

$x \backslash y$	y_1	y_2	....	y_j	....	y_n	**Total**
x_1	$P(x_1 y_1)$	$P(x_1 y_2)$	....	$P(x_1 y_j)$	....	$P(x_1 y_n)$	$P(x_1)$
x_2	$P(x_2 y_1)$	$P(x_2 y_2)$	....	$P(x_2 y_j)$	....	$P(x_2 y_n)$	$P(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	$P(x_i y_1)$	$P(x_i y_2)$	....	$P(x_i y_j)$	....	$P(x_i y_n)$	$P(x_i)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_m	$P(x_m y_1)$	$P(x_m y_2)$	....	$P(x_m y_j)$	....	$P(x_m y_n)$	$P(x_m)$
Total	$P(y_1)$	$P(y_2)$	....	$P(y_j)$	....	$P(y_n)$	1

Therefore $P(x_i) = \sum_{j=1}^{n} P(x_i y_j)$ and $P(y_j) = \sum_{i=1}^{m} P(x_i y_j)$

Now $E(X+Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_i + y_j)P(x_i y_j)$

$$E(X+Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i P(x_i y_j) + \sum_{i=1}^{m}\sum_{j=1}^{n} y_j P(x_i y_j)$$

$$E(X+Y) = \sum_{i=1}^{m} x_i \sum_{j=1}^{n} P(x_i y_j) + \sum_{j=1}^{n} y_j \sum_{i=1}^{m} P(x_i y_j)$$

$$E(X+Y) = \sum_{i=1}^{m} x_i P(x_i) + \sum_{j=1}^{n} y_j P(y_j)$$

$$E(X+Y) = E(X) + E(Y)$$

Binomial Probability Distribution

Binomial Probability Distribution is one of the special type of discrete probability distributions.

Binomial Probability Distribution is a mathematical formula used to calculate the probabilities for the possible values of the random variable X. The possible values of X are 0, 1, 2, ..., n. Hence, the formula of Binomial Probability Distribution is given by

$$P(x) = {}^{n}C_{x} \cdot p^{x} \cdot q^{n-x}\ ,\quad x = 0, 1, 2, \ldots\ldots\ldots, n$$

where: n = No. of independent trials

x = No. of successes in n trials

p = Probability of a success in a single trial remains constant from trial to trial

$q = 1 - p$ = Probability of a failure in a single trial

The above formula is based on an experiment, this experiment is called a Binomial Experiment, if it possess the following properties

(i) The No. of trials must be a fixed number n.

(ii) Each trial must contain two possible outcomes (success and failure).

(iii) the probability p of success remains constant from trial to trial.

(iv) Every trial must be independent of the others.

Note: When an experiment with only two possible outcomes is repeated in such a way that the probabilities with only two possible outcomes remains constant from trial to trial, the trials are called Bernoulli's Trials.

Example 4.10

A fair coin is tossed 4 times. What is the probability of obtaining

(i) 3 heads (ii) no head

Solution:

$n = 4$ = No. of trials

$x = 3$ = No. of Heads / No. of successes

$p = \frac{1}{2}$ = P(a Head) = P(Success)

$q = 1 - \frac{1}{2}$ = P(a Tail) = P(Failure)

CamScanner

This is the problem of binomial experiment, therefore we apply the formula of Binomial Distribution to find the probability of (i) 3 heads (ii) no heads.

(i) $P(x=3) = {}^4C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{4-3} = 4\times\frac{1}{8}\times\frac{1}{2}=\frac{1}{4}$

(ii) $P(\text{no head}) = P(x=0) = {}^4C_0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{4-0} = \frac{1}{16}$

Example 4.11

The probability that a patient recovers from a delicate heart operation is 0.85. What is the probability that at least 3 of the next 4 patients having this operation survive.

Solution:

P(Survive) = P(recovers) = $p = 0.85$

P(not survive) = $q = 1-p = 1-0.85 = 0.15$

$n = 4$ and $x = 3$ or 4

$P(x\geq 3) = P(x=3)+P(x=4) = {}^4C_3(0.85)^3(0.15)^{4-3}+{}^4C_4(0.85)^4(0.15)^0 = 0.89$

Properties of Binomial Distribution

(i) It is the discrete type of probability distribution.

(ii) It has two parameters (i.e. constants) "n" and "p"

(iii) $\sum_{x=0}^{n} {}^nC_x \cdot p^x \cdot q^{n-x} = (q+p)^n = 1$ since $q+p=1$

(iv) Mean = np and Variance = npq

(v) Mean > Variance

(vi) If $p=\frac{1}{2}$ then distribution becomes a symmetrical distribution.

CS CamScanner

Binomial Frequency Distribution

If the binomial probability distribution is multiplied by N, the No. of experiments. The resulting distribution is known as the binomial frequency distribution.

Symbolically: $f_e = N\binom{n}{x} \cdot p^x \cdot q^{n-x}$

Binomial frequency distribution is used to find expected number of successes (i.e. expected frequencies), when a Binomial Experiment containing 'n' independent trials is repeated 'N' times.

Note: It is to be noted that n independent trials constitute one experiment (or one set).

Example 4.12

Out of 400 families with 3 children each, how many would you expect to have 2 boys. Assume equal probability for boys and girls.

Solution:

$$P(\text{boy}) = P(\text{girl}) = \frac{1}{2}$$

$$P(2\text{ boys}) = {}^3C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{3-2} = \frac{3}{8}$$

$$\text{Expected No. of families with 2 boys} = 400 \times \frac{3}{8} = 150$$

Example 4.13

Take 100 sets of 10 tosses of an unbiased coin. In how many cases do you expect to get 7 heads.

CamScanner

Solution:

The probability of getting a head when an unbiased coin is thrown is $P(H)=\frac{1}{2}$

Probability of getting a tail is $P(T)=\frac{1}{2}$ Therefore, probability of getting 7 heads in 10 tosses of a coin is ${}^{10}C_7\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^{10-7}$

Therefore, No. of cases of 7 heads in 100 sets is $100\cdot{}^{10}C_7\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^3=12$ nearly.

Fitting a Binomial Distribution

Fitting a Binomial Distribution means to calculate the expected frequencies against every value of the random variable X by assuming that the given data really follows to Binomial Distribution.

The procedure for fitting a Binomial Distribution is illustrated by the following example.

Example 4.14

Fit a binomial distribution to the following data:

No. of Heads (x)	0	1	2	3	4	Total
Frequency (f)	10	40	50	20	5	125

Solution:

First we compute $\overline{X}=\frac{\sum fx}{\sum f}$ and then equate to np to find p where $n=4$

$$\overline{X}=\frac{\sum fx}{\sum f}=\frac{10\times0+40\times1+50\times2+20\times3+5\times4}{125}$$

$$\overline{X}=\frac{220}{125}=1.76$$

Now equating $\overline{X}$ and np

i.e. $\bar{X} = np$

$$1.76 = np$$

$$p = \frac{1.76}{n} = \frac{1.76}{4} = 0.44$$

$$q = 1 - p = 1 - 0.44 = 0.56$$

Fitted Probability Distribution is given by

$$P(x) = {}^4C_x \cdot (0.44)^x (0.56)^{4-x}, \; x = 0, 1, 2, 3, 4$$

The fitted frequency distribution of Binomial is given by

$$f_e = N\binom{n}{x} \cdot p^x \cdot q^{n-x} = 125\binom{4}{x}(0.44)^x (0.56)^{4-x}, \; x = 0, 1, 2, 3, 4, \ldots\ldots$$

The probabilities and the expected frequencies for $x = 0, 1, 2, 3$ and 4 are shown in the following table.

No. of Heads (x)	Probability $P(x)$	Expected frequency $f_e = N\,P(x) = 125P(x)$
0	${}^4C_0(0.44)^0(0.56)^{4-0} = 0.0983$	$125 \times 0.0983 = 12.3$
1	${}^4C_1(0.44)^1(0.56)^{4-1} = 0.3091$	$125 \times 0.3091 = 38.6$
2	${}^4C_2(0.44)^2(0.56)^{4-2} = 0.3643$	$125 \times 0.3643 = 45.5$
3	${}^4C_3(0.44)^3(0.56)^{4-3} = 0.1908$	$125 \times 0.1908 = 23.9$
4	${}^4C_4(0.44)^4(0.56)^0 = 0.0375$	$125 \times 0.0375 = 4.7$
Total	1	125

Example 4.15

Three coins are thrown 64 times and the distribution of the number of heads is observed to be

CS CamScanner

No. of Heads (x)	0	1	2	3	Total
Observed freq. (f)	21	31	12	0	64

Fit a Binomial Distribution

Solution:

$$\bar{X} = \frac{\sum fx}{\sum f} = \frac{20\times 0 + 31\times 1 + 12\times 2 + 0\times 3}{64}$$

$$\bar{X} = \frac{0+31+24+0}{64} = 0.8594$$

Now equating $\bar{X}$ and np

$$\bar{X} = np \qquad \text{where } n = 3$$

then $p = \frac{\bar{X}}{n} = \frac{0.8594}{3} = 0.29$

$$q = 1 - p = 1 - 0.29 = 0.71$$

The fitted probability distribution of binomial is

$$P(x) = {}^3C_x (0.29)^x (0.71)^{3-x}, \; x = 0, 1, 2, 3$$

The fitted frequency distribution of binomial is given by

$$f_e = NP(x) = 64\binom{n}{x}(0.29)^x(0.71)^{3-x}, \; x = 0, 1, 2, 3$$

The probabilities and the expected frequencies for $x = 0, 1, 2$ and 3 are shown in the following table.

No. of Heads (x)	Probability $P(x)$	Expected frequency $f_e = N\,P(x) = 64P(x)$
0	${}^3C_0(0.29)^0(0.71)^{3-0} = 0.3579$	22.9
1	${}^3C_1(0.29)^1(0.71)^{3-1} = 0.4386$	28.1
2	${}^3C_2(0.29)^2(0.71)^{3-2} = 0.1791$	11.5
3	${}^3C_3(0.29)^3(0.71)^0 = 0.0244$	1.6
Total	1	≈ 64

MISCELLANEOUS SOLVED PROBLEMS ON BINOMIAL DISTRIBUTION

Problem 1

Find Mean and Variance of Binomial Probability Distribution.

Solution:

$$\text{Mean} = E(X) = \sum_{x=0}^{n} xP(x) = \sum_{x=0}^{n} x \cdot {}^nC_x \cdot p^x \cdot q^{n-x}$$

$$\text{Mean} = 0 \cdot {}^nC_0 \cdot p^0 \cdot q^{n-0} + 1 \cdot {}^nC_1 \cdot p^1 \cdot q^{n-1} + 2 \cdot {}^nC_2 \cdot p^2 \cdot q^{n-2} + \dots\dots\dots + n \cdot {}^nC_n \cdot p^n \cdot q^0$$

$$\text{Mean} = 0 + npq^{n-1} + n(n-1)p^2q^{n-2} + \dots\dots\dots + np^n$$

$$\text{Mean} = np\left[q^{n-1} + (n-1)pq^{n-2} + \dots\dots\dots + p^{n-1}\right]$$

$$\text{Mean} = np\,[q+p]^{n-1} = np\,(1)^{n-1} = np$$

Hence $\text{Mean} = np$

$$\text{Variance} = E(X^2) - [E(X)]^2$$

Now find $E(X^2)$ and it may be written as

$$E(X^2) = E[X(X-1)+X] = EX(X-1)+E(X)$$

$$E(X^2) = EX(X-1)+np \quad \cdots\cdots\cdots\cdots\cdots (1)$$

$$EX(X-1) = \sum_{x=0}^{n} x(x-1)\cdot {}^nC_x \cdot p^x \cdot q^{n-x}$$

$$EX(X-1) = 0+0+2(2-1)\cdot {}^nC_2 \cdot p^2 \cdot q^{n-2} + 3(3-1)\cdot {}^nC_3 \cdot p^3 \cdot q^{n-3} + \ldots\ldots$$
$$\ldots\ldots + n(n-1)\cdot {}^nC_n \cdot p^n \cdot q^0$$

$$EX(X-1) = n(n-1)p^2q^{n-2} + n(n-1)(n-2)p^3q^{n-3} + \ldots\ldots + n(n-1)p^n$$

$$EX(X-1) = n(n-1)p^2\left[q^{n-2} + (n-2)pq^{n-3} + \ldots\ldots + p^{n-2}\right]$$

$$EX(X-1) = n(n-1)p^2[q+p]^{n-2} = n(n-1)p^2(1)^{n-2} = n(n-1)p^2$$

By putting in (1) we get

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

Hence, Variance $= V(x) = n(n-1)p^2 + np - (np)^2$

$$V(x) = np[(n-1)p+1-np]$$

$$V(x) = np[np-p+1-np]$$

$$\boxed{\text{Variance} = np(1-p) = npq}$$

Problem 2

Compute the mean and variance of Binomial distribution when $n = 10$ and $p = 0.6$

Solution:

Mean $= np = 10 \times 0.6 = 6$

Variance $= npq = 10 \times 0.6 \times 0.4$ Since $p + q = 1$ and $q = 1 - p = 1 - 0.6 = 0.4$

Variance $= 2.4$

Problem 3

Mean of a binomial distribution is 20 and the variance is 16. Calculate n, p and q.

Solution:

$$\text{Mean} = np = 20 \cdots\cdots\cdots (1)$$

and $$\text{Variance} = npq = 16 \cdots\cdots (2)$$

By putting (1) in (2), we get

$$20q = 16 \quad \text{and} \quad q = \frac{16}{20} = \frac{4}{5} = 0.8$$

$$p = 1 - q = 1 - 0.8 = 0.2$$

As $np = 20$ and $p = 0.2$, then

$$0.2 \times n = 20 \quad \text{and} \quad n = \frac{20}{0.2} = 100$$

Hence $n = 100$, $p = 0.2$ and $q = 0.8$

Problem 4

Show that $\sum_{x=0}^{n} {}^{n}C_{x} \cdot p^{x} \cdot q^{n-x} = (p+q)^{n} = 1$

Solution:

$$\sum_{x=0}^{n} {}^{n}C_{x} \cdot p^{x} \cdot q^{n-x} = {}^{n}C_{0} \cdot p^{0} \cdot q^{n-0} + {}^{n}C_{1} \cdot p^{1} \cdot q^{n-1} + {}^{n}C_{2} \cdot p^{2} \cdot q^{n-2} + \dots\dots\dots + {}^{n}C_{n} \cdot p^{n} \cdot q^{0}$$

Since $${}^{n}C_{0} = \frac{n!}{0!(n-0)!} = 1 \quad \text{and} \quad {}^{n}C_{n} = \frac{n!}{n!(n-n)!} = 1$$

also $p^{0} = 1$ and $q^{0} = 1$

then $$\sum_{x=0}^{n} {}^{n}C_{x} \cdot p^{x} \cdot q^{n-x} = q^{n} + {}^{n}C_{1} \cdot p^{1} \cdot q^{n-1} + {}^{n}C_{2} \cdot p^{2} \cdot q^{n-2} + \dots\dots\dots + p^{n}$$

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

Proved

CS CamScanner

Problem 5

In a binomial distribution $q = \frac{3}{5}$ and $n = 5$, write down the full expansion of $(q+p)^5$ and find the mean and standard deviation of the distribution.

Solution:

$$(q+p)^5 = q^5 + {}^5C_1 q^{5-1} p + {}^5C_2 q^{5-2} p^2 + {}^5C_3 q^{5-3} p^3 + {}^5C_4 q^{5-4} p^4 + p^5$$

Since $q = \frac{3}{5}$ and $p = 1 - \frac{3}{5} = \frac{2}{5}$

then $$\left(\frac{3}{5}+\frac{2}{5}\right)^5 = \left(\frac{3}{5}\right)^5 + 5\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right) + 10\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^2 + 10\left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right)^3 + 5\left(\frac{3}{5}\right)^1\left(\frac{2}{5}\right)^4 + \left(\frac{2}{5}\right)^5$$

$$\left(\frac{3}{5}+\frac{2}{5}\right)^5 = \frac{243}{3125} + \frac{810}{3125} + \frac{1080}{3125} + \frac{720}{3125} + \frac{240}{3125} + \frac{32}{3125}$$

Now Mean $= np = 5 \times \frac{2}{5} = 2$

and S. D. $= \sqrt{npq} = \sqrt{5 \times \frac{2}{5} \times \frac{3}{5}} = 1.095$

Problem 6

Derive the formula of Binomial Probability Distribution.

$$P(x) = {}^nC_x \cdot p^x \cdot q^{n-x}, \ x = 0, 1, 2, \ldots\ldots\ldots, n$$

Solution:

Suppose a coin is tossed n – times. If head occurs, we say it out success and if tail occurs, we say it our failure.

If P(Head) $= p$ and P(Tail) $= q$

Such that $p + q = 1$

CamScanner

Let X = No. of Heads = 0, 1, 2, ..., n

Now $P(X = 0) = P(\text{all Tails}) = P(TTT\ldots\ldots T)$

$$= P(T)\,P(T)\,P(T)\ldots\ldots\ldots\ldots P(T)$$

$$= q\cdot q\cdot q\cdots\ldots\ldots\ldots\cdot q = q^n$$

$P(X = 1) = P(1 - \text{Head and } n - 1 \text{ Tails})$

$$= P(HTT\ldots\ldots T) + P(THT\ldots\ldots T) + \ldots\ldots + P(TTT\ldots\ldots H)$$

$$= pq^{n-1} + pq^{n-1} + \ldots\ldots + pq^{n-1}$$

Since pq^{n-1} is repeated ${}^nC_1$ times, the above result may be written as

$$P(X = 1) = {}^nC_1 p\cdot q^{n-1}$$

Similarly

$$P(X = 2) = {}^nC_2 p^2\cdot q^{n-2}$$

$$P(X = 3) = {}^nC_3 p^3\cdot q^{n-3}$$

In general

$$P(X = x) = {}^nC_x\cdot p^x\cdot q^{n-x}$$

Hence $\boxed{P(x) = {}^nC_x\cdot p^x\cdot q^{n-x}, \qquad x = 0, 1, 2, \ldots\ldots\ldots, n}$

Continuous Probability Distribution

So far we have been dealing with discrete probability distributions i.e. distributions in which the random variable X takes on integer values such as 0, 1, 2, 3, ... But when we deal with random variables like heights and weights, we find that such random variables can take any value in the given interval $a \le x \le b$. Such random variables are called continuous random variables and their probability distributions are accordingly known as continuous probability distributions.

The most important distribution defined on a continuous random variable is Normal Probability Distribution.

Normal Probability Distribution

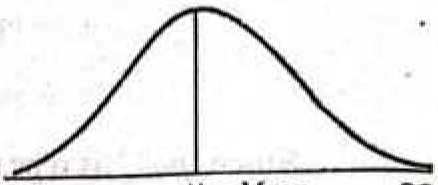
The equation of the normal probability distribution is

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty \le x \le \infty$$

where μ = Mean , σ = Standard Deviation

π (pi) = 3.14159 and $e = 2.71828$

Graph (or shape) of the normal distribution is displayed by the curve; where the total area under the curve is unity.



Properties of Normal Distribution

Following are the important properties of Normal Distribution:

(a) It is the continuous type of probability distribution.

(b) It has two parameters (i.e. constants) 'μ' and 'σ'

(c) Mean = μ and Variance = σ^2

(d) Mean = Median = Mode = μ

(e) The curve of Normal distribution is Bell-Shaped.

(f) The whole area under the curve is unity.

(g) (i) Area of Normal curve under the limits $\mu \pm \sigma$ (i.e. between $\mu - \sigma$ and $\mu + \sigma$) is 68.26%

(ii) Area of Normal curve under the limits $\mu \pm 2\sigma$ (i.e. between $\mu - 2\sigma$ and $\mu + 2\sigma$) is 95.44%

(iii) Area of Normal curve under the limits $\mu \pm 3\sigma$ (i.e. between $\mu - 3\sigma$ and $\mu + 3\sigma$) is 99.73%

(h) If 'x' is a normal variate (or normal random variable) with mean = μ and variance = σ^2 then $z = \frac{x - \mu}{\sigma}$ is called the Standard Normal Variate with mean = 0 and variance = 1. The p. d. f. (or the equation) of standard normal distribution is:

CamScanner

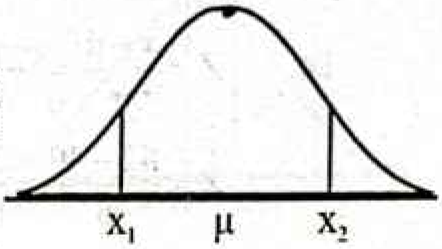
$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}, \quad -\infty \le z \le \infty$$

(i) Coeff. of Skewness $=\beta_1=0$ and Coeff. of Kurtosis $=\beta_2=3$

Area Under the Normal Curve

The curve of any continuous probability distribution is constructed so that the area under the curve bounded by two ordinates $X=x_1$ and $X=x_2$ equals the probability that the random variable X takes a value between $X=x_1$ and $X=x_2$

Thus, for the normal curve in the given figure, the $Pr(x_1<X<x_2)$ is represented by the area of the shaded region.



i.e. $Pr(x_1<X<x_2)=$ Area of the shaded region

To determine the area or probability of an interval of a normal distribution with mean = μ and standard deviation = σ, the following steps are taken:

(i) First we convert x values into z values by the help of the following formula

$$z=\frac{x-\mu}{\sigma}$$

where z is called standard normal variable with Mean = 0 and Standard deviation = 1

Hence we have $Pr(x_1<X<x_2)=Pr(z_1<Z<z_2)$

(ii) Determine the probability or area for each value of z (i.e. $z=z_1$ and $z=z_2$) from the "Normal Area Table" enclosed at the end of the book.

(iii) Subtract the area below the value of $z=z_1$ from the area below the value $z=z_2$

Therefore $Pr(x_1<X<x_2)=Pr(z_1<Z<z_2)=P(Z<z_2)-P(Z<z_1)$

Note: Normal area table enclosed at the end of the book gives the area under the standard normal curve corresponding $P(Z<z)$ i.e. P(less than the given value of Z = z or $P(-\infty<Z<z)$ or area under the curve to the left of the Z-value.

CamScanner

Example 4.16

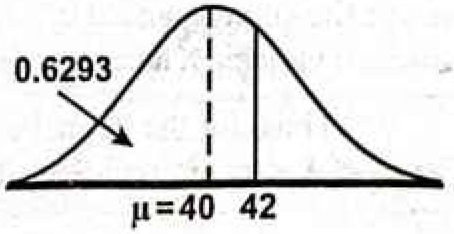
Given a normal distribution with $\mu = 40$ and $\sigma = 6$, find the probability (or area) that X assumes a value (i) Below 42 (ii) Above 27 (iii) Between 42 and 51

Solution:

Here $\mu = 40$ and $\sigma = 6$, then

(i) P(below 42) = P(x < 42)

Now transform the x-value into the value of z

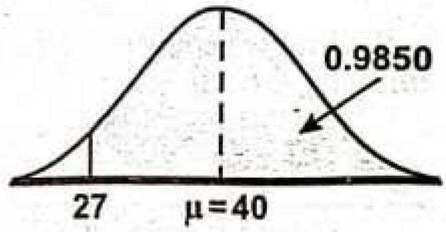


i.e. $Z = \frac{x-\mu}{\sigma} = \frac{x-40}{6}$ and for x = 42,

$$Z = \frac{42-40}{6} = \frac{2}{6} = 0.33$$

then P(x < 42) = P(z < 0.33) = 0.6293 (from Normal area table)

(ii) P(Above 27) = P(x > 27) = 1 − P(x < 27)



Since $Z = \frac{x-\mu}{\sigma} = \frac{x-40}{6}$ and for x = 27,

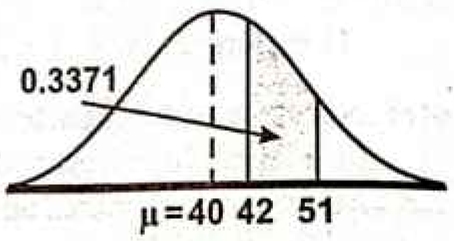
$$Z = \frac{27-40}{6} = -2.17$$

therefore P(x > 27) = 1 − P(x < 27) = 1 − P(z < −2.17)

$$= 1 - 0.0151 = 0.9850$$

(iii) P(between 42 and 51) = P(42 < x < 51) = P(x < 51) − P(x < 42)

since $Z = \frac{x-\mu}{\sigma} = \frac{x-40}{6}$ and for x = 51,



$$Z = \frac{51-40}{6} = 1.83$$

and for x = 42, $Z = \frac{42-40}{6} = 0.33$

therefore P(x < 51) = P(z < 1.83) = 0.9664

$$P(x < 42) = P(z < 0.33) = 0.6293$$

Hence P(Between 42 and 51) = 0.9664 – 0.6293 = 0.3371

Example 4.17

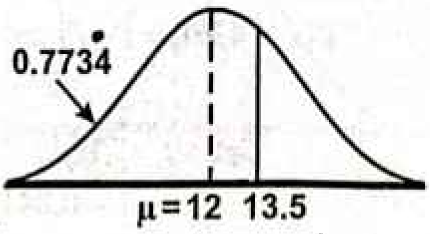
Given a normal distribution with $\mu = 12$ and $\sigma = 2$ find

(a) The area below 13.5 (b) The area above 10 (c) The area between 6 and 18

Solution:

(a) $P(x < 13.5) = P\left(z < \frac{13.5-12}{2}\right)$

$= P(z < 0.75) = 0.7734$

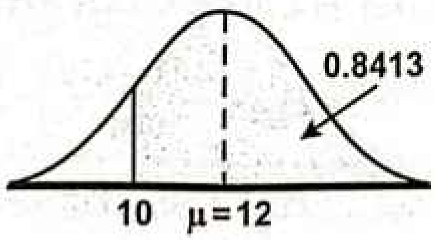


(b) $P(x > 10) = 1 - P(x < 10)$

$= 1 - P\left(z < \frac{10-12}{2}\right)$

$= 1 - P(z < -1.00)$

$= 1 - 0.1587 = 0.8413$

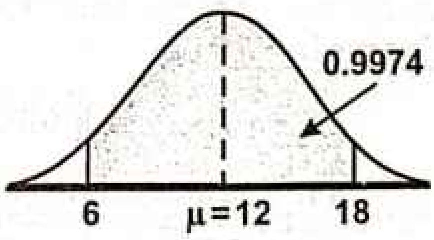


(c) $P(6 < x < 18) = P(x < 18) - P(x < 6)$

$= P\left(z < \frac{18-12}{2}\right) - P\left(z < \frac{6-12}{2}\right)$

$= P(z < 3.00) - P(z < -3.00)$

$= 0.9987 - 0.0013 = 0.9974$



Example 4.18

The burning time of an experimental rocket is a random variable having the normal distribution with $\mu = 4.76$ seconds and $\sigma = 0.04$ seconds. What is the probability that this kind of rocket will burn.

(a) less than 4.66 seconds (b) more than 4.80 seconds

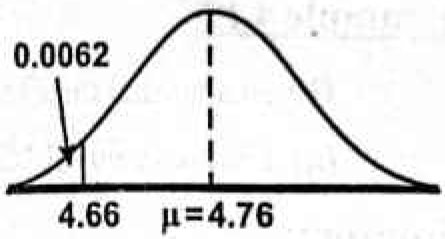
(c) anywhere from 4.70 to 4.82 seconds

Solution:

Since $\mu = 4.76$ and $\sigma = 0.04$, then

(a) $P(x < 4.66) = P\left(z < \dfrac{4.66 - 4.76}{0.04}\right)$

$= P(z < -2.50) = 0.0062$

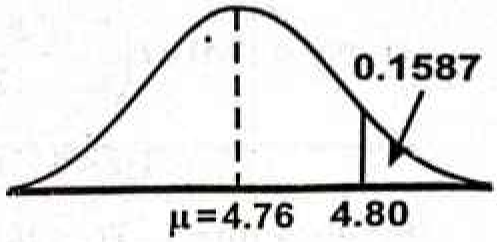


(b) $P(x > 4.80) = 1 - P(x < 4.80)$

$P(x > 4.80) = 1 - P\left(z < \dfrac{4.80 - 4.76}{0.04}\right)$

$= 1 - P(z < 1.00)$

$= 1 - 0.8413 = 0.1587$

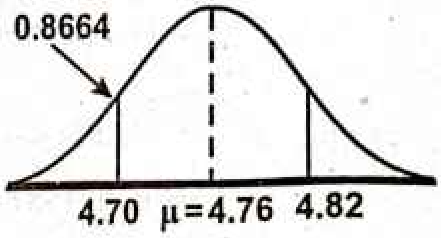


(c) $P(4.70 < x < 4.82) = P(x < 4.82) - P(x < 4.70)$

$= P\left(z < \dfrac{4.82 - 4.76}{0.04}\right) - P\left(z < \dfrac{4.70 - 4.76}{0.04}\right)$

$= P(z < 1.50) - P(z < -1.50)$

$= 0.9332 - 0.0668 = 0.8664$



Example 4.19

Given a normal distribution with $\mu = 100$ and $\sigma = 5$, find

(a) The area below 107 (b) The area above 89.5

(c) The area between 94 and 103

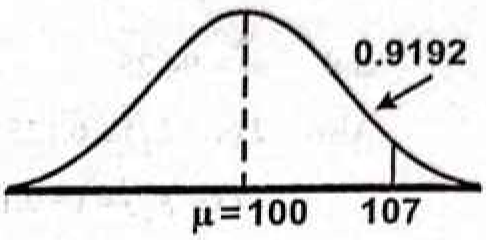
(d) The point that has 80% of the area below it

(e) The two points containing the middle 75% of the area

Solution:

(a) $P(x < 107) = P\left(z < \dfrac{107-100}{5}\right)$

$= P(z < 1.40) = 0.9192$

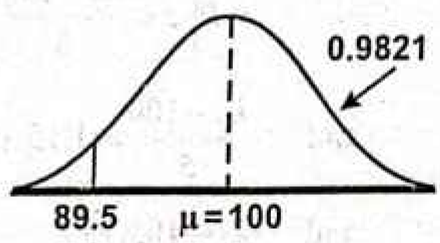


(b) $P(x > 89.5) = 1 - P(x < 89.5)$

$= 1 - P\left(z < \dfrac{89.5-100}{5}\right)$

$= 1 - P(z < -2.10)$

$= 1 - 0.0179 = 0.9821$

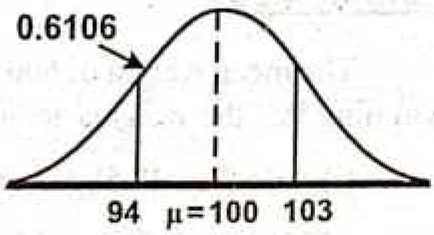


(c) $P(94 < x < 103) = P(x < 103) - P(x < 94)$

$= P\left(z < \dfrac{103-100}{5}\right) - P\left(z < \dfrac{94-100}{5}\right)$

$= P(z < 0.60) - P(z < -1.20)$

$= 0.7257 - 0.1151 = 0.6106$



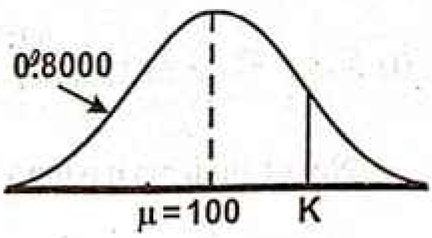
(d) Let 'k' be the point that has 80% of the area below it as shown in the given figure.

In other words $P(x < k) = 0.80$

then $P\left(z < \dfrac{k-100}{5}\right) = 0.80$

and $\dfrac{k-100}{5} = 0.84$ (from normal table)

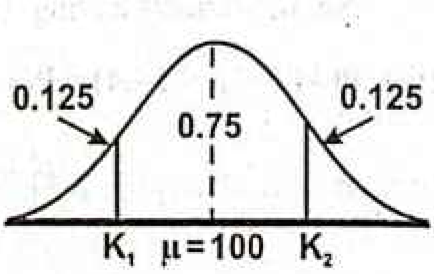
$k = 104.2$



(e) Let k_1 and k_2 be the two points containing the middle 75% of the area as shown in the given figure

then $P(x < k_1) = 0.125$

$P\left(z < \dfrac{k_1-100}{5}\right) = 0.125$



CS CamScanner

and $\frac{k_1 - 100}{5} = -1.15$ (from normal table)

and $k_1 = 94.25$

Also $P(x > k_2) = 0.125$

$$P(x < k_2) = 1 - 0.125 = 0.875$$

$$= P\left(z < \frac{k_2 - 100}{5}\right) = 0.875$$

and $\frac{k_2 - 100}{5} = 1.15$ (from normal table)

and $k_2 = 105.75$

Example 4.20

The mean weight of 600 students in a school is 50kg and standard deviation is 6kg. Assuming that the weights are normally distributed, find how many students weight

(i) less than 42.5kg (ii) greater than 56.5kg

(iii) between 44.5 and 54.4 kg.

Solution:

(i) $P(x < 42.5) = P\left(z < \frac{42.5 - 50}{6}\right) = P(z < -1.25) = 0.1056$

No. of students having weight less than 42.5kg $= 600 \times 0.1056 = 63$

(ii) $P(x > 56.5) = 1 - P(x < 56.5) = 1 - P\left(z < \frac{56.5 - 50}{6}\right) = 1 - P(z < 1.08)$

$$= 1 - 0.8599 = 0.1401$$

No. of students having weight greater than 56.5 $= 600 \times 0.1401 = 84$

(iii) $P(44.5 < x < 54.4) = P(x < 54.4) - P(x < 44.5)$

$$= P\left(z < \frac{54.4 - 50}{6}\right) - P\left(z < \frac{44.5 - 50}{6}\right)$$

CamScanner

$$= P(z < 0.73) - P(z < -0.92) = 0.7673 - 0.1788 = 0.5885$$

No. of students having weight between 44.5kg and 54.4kg is $= 600 - 0.5885 = 353$

Normal Approximation to Binomial Distribution

When n is large and p is not very close to 0 or to 1, then binomial distribution approximates to normal distribution having $\mu = np$ and $\sigma = \sqrt{npq}$ where $z = \dfrac{x - np}{\sqrt{npq}}$ is the standard normal variable.

Note: It is to be noted that when we use Normal distribution (i.e. continuous distribution) to approximate a discrete random variable. So we make a continuity correction by adding or subtracting half the unit of the measurement to each value of the discrete variable.

For example:

If x is binomial (i.e. a discrete) variable, then x can take on any of the values 0, 1, 2, 3, 4. Therefore, the values of discrete variables are transformed as

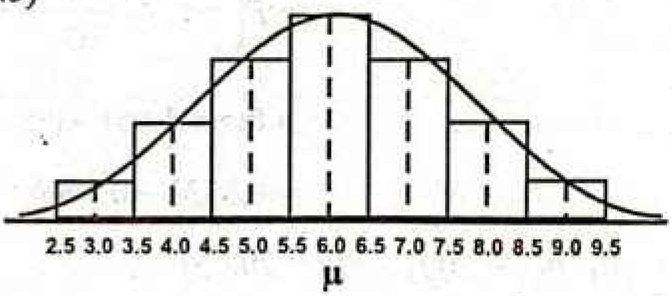
$P(x = 4)$ transformed to $P(3.5 < x < 4.5)$

$P(x < 4)$ transformed to $P(x < 3.5)$

$P(x \leq 4)$ transformed to $P(x < 4.5)$

$P(x \geq 4)$ transformed to $P(x > 3.5)$

$P(x > 4)$ transformed to $P(x > 4.5)$



$P(5 \leq x \leq 8)$ transformed to $P(4.5 < x < 8.5)$

$P(5 < x \leq 8)$ transformed to $P(5.5 < x < 8.5)$

$P(5 \leq x < 8)$ transformed to $P(4.5 < x < 7.5)$

$P(5 < x < 8)$ transformed to $P(5.5 < x < 7.5)$

As shown in the above probability Histogram.

Example 4.21

A Coin is tossed 400 times. Use the normal curve approximation to find the probability of obtaining

(a) Between 185 and 210 heads inclusive (b) Less than 209 heads

(c) At least 200 heads (d) Exactly 205 heads

Solution:

$n = 400$ and $p = \frac{1}{2}$ then $\mu = np = (400)\frac{1}{2} = 200$

and $\sigma = \sqrt{npq} = \sqrt{400 \times \frac{1}{2} \times \frac{1}{2}} = 10$

where $z = \frac{x - np}{\sqrt{npq}} = \frac{x - 200}{10}$

Since the data (No. of heads) are discrete and measured in integers we need the continuity correction.

(a) $P(185 \le x \le 210) = P(184.5 < x < 210.5)$

$$= P(x < 210.5) - P(x < 184.5)$$

$$= P\left(z < \frac{210.5 - 200}{10}\right) - P\left(z < \frac{184.5 - 200}{10}\right)$$

$$= P(z < 1.05) - P(z < -1.55)$$

$$= 0.8531 - 0.0606 = 0.7925$$

(b) $P(x < 209) = P(x < 208.5)$

$$= P\left(z < \frac{208.5 - 200}{10}\right)$$

$$= P(z < 0.85) = 0.8023$$

(c) $P(x \ge 200) = 1 - P(x < 200)$

$$= 1 - P(x < 199.5)$$

$$= 1 - P\left(z < \frac{199.5 - 200}{10}\right)$$

$$= 1 - P(z < -0.05) = 1 - 0.4801 = 0.5199$$

(d) $P(x = 205) = P(204.5 < x < 205.5)$

$$= P(x < 205.5) - P(x < 204.5)$$

$$= P\left(z < \frac{205.5 - 200}{10}\right) - P\left(z < \frac{204.5 - 200}{10}\right)$$

$$= P(z < 0.55) - P(z < 0.45)$$

$$= 0.7088 - 0.6736 = 0.0352$$

Example 4.22

A hockey player hits on 60% successful shots in penalty corners. What is the probability that he makes less than 50 successful shots, out of 100 opportunities to be given to him in various matches to be played next month.

Solution:

$n = 100,\ p = 0.6,\ q = 0.4$

Then $\mu = np = (100)(0.6) = 60$

and $\sigma = \sqrt{npq} = \sqrt{100 \times 0.6 \times 0.4} = 49$

$$P(x < 50) = P(x < 49.5) = P\left(z < \frac{495 - 60}{49}\right)$$

$$= P(z < -2.14) = 0.0162$$

EXERCISE – 4

4.1 Define Random Variable with examples.

Karachi Board 2003

4.2 (a) A coin is tossed three times. Find the probability distribution of the random variable 'number of heads'.

Karachi Board 1998

(b) A coin is tossed 3 times. Find the probability distribution of the random variable 'number of heads' and calculate the mathematical expectation of the random variable.

Karachi Board 2003

4.3 (a) In a departmental store there are 3 women and 5 men workers, 3 workers are selected at random. If X is a random variable indicating the number of women workers selected, find the probability distribution of X.

Karachi Board 1987

(b) A coin is tossed 5 times, find the probability of $x = 0, 1, 2, 3, 4, 5$ heads. Also find E(X)

Karachi Board 2002

4.4 (a) Define Mathematical Expectation of a random variable.

(b) From the following probability distribution

x	0	1	2	3	4	5
P(x)	0.02	0.19	0.28	0.24	0.23	0.04

Calculate: (i) Mean of X (ii) Variance of X

Karachi Board 1988

4.5 (a) Define Random Variable. Explain its types.

(b) A coin is tossed 3 times. If X denotes the number of tails appearing, find the probability distribution of the random variable X. Also find the E(X) and V(X) of the random variable X.

Karachi Board 1991

4.6 (a) Define Random Variable and Probability Distribution.

(b) Three balls are drawn at random from a bag containing 4 white and 4 black balls. If X denotes the number of white balls drawn from the bag, find the probability distribution of X. Also find the expected number of white balls.

Karachi Board 1995

4.7 The possible values of a random variable X are the integers from 1 to 8. If the values are equally likely, find

(i) Probability distribution of X.

(ii) E(X) and V(X)

Karachi Board 2001

4.8 (a) If two playing cards are drawn from a pack of 52 cards. Obtain the probability distribution for X = No. of Kings drawn. Also find E(X), $E(X-\overline{X})^2$

Karachi Board 2000

(b) Define Mathematical Expectation of random variable.

(c) From the following probability distribution

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

Calculate: (i) Mean of X (ii) Variance of X

Karachi Board 1998

4.9 Find the mean and the variance of the following probability distribution.

CamScanner

x	0	1	2	3
P(x)	0.1	0.5	0.3	0.1

Karachi Board 1998

4.10 3 balls are drawn from a box containing 3 white and 4 black balls. If X denotes the number of black balls drawn, obtain the probability distribution of X. Find the mean and variance of this distribution.

Karachi Board 1999

4.11 (a) Define Random Variable and Mathematical Expectation.

(b) For the following probability distribution $f(x)=\frac{1}{6}$, $x = 1, 2, 3, 4, 5, 6.$ Calculate $E[X-E(X)]$

Karachi Board 1996

4.12 (a) Prove that $E(aX-b) = aE(X)-b$, where X is a discrete random variable and 'a' and 'b' are constants.

(b) If X is a discrete random variable which can assume the values 0, 1, 2, 3 and 4 with corresponding associate probabilities 0.70, 0.10, 0.08, 0.07 and 0.05 respectively. Calculate:

(i) $E(X)$ (ii) $E(3X-2)$ (iii) Also verify that $E(3X-2) = 3E(X)-2$

Karachi Board 1993

4.13 If X represents the dots shown on a fair die, find (i) $E(X)$ and (ii) $E(X^2)$

Karachi Board 1989

4.14 (a) If 'a' and 'b' are constants, prove that $E(aX+b) = aE(X)+b$

(b) If $f(x)=\frac{6-|7-x|}{36}$; $x=2,3,4,\ldots\ldots,12$ then find the mean and variance of the random variable X.

Karachi Board 1990

4.15 (a) Differentiate between variable and random variable. Define mathematical expectation of a random variable.

CamScanner

(b) Find mean and variance of the following probability distribution.

x	8	12	16	20	24
P(x)	1/8	1/6	3/8	1/4	1/12

Karachi Board 1985

4.16 (a) Define Random Variable.

(b) Define Mathematical Expectation of a random variable.

(c) Prove that the expected value of a constant is constant itself.

Karachi Board 1984

4.17 (a) State the properties of Mathematical Expectation.

Karachi Board 1992

(b) Define Random Variable and Mathematical Expectation.

(c) A random variable X takes the values $-3, -2, 2, 3$ and 4 with probabilities P(x) equal to 1/5, 1/10, 1/10, 1/5 and 2/5 respectively, compute E(X) and show that $E(5X + 10) = 5E(X) + 10$

Karachi Board 1988

4.18 Let the probability function of X be $P(X=x) = \frac{x}{10}$; $x = 1, 2, 3, 4$ find the probability distribution of X and mean of X. Also find $E(2X + 3)$.

Karachi Board 2003, 2008 (Supp.)

4.19 (a) Show that $E(AX + B) = AE(X) + B$, where A and B are constants and X is a discrete random variable.

Karachi Board 2007 (Supp.)

(b) The number of automobile accidents that occur in a certain city is 0, 1, 2, 3 and 4 with the corresponding probabilities 0.80, 0.11, 0.06, 0.02 and 0.01 respectively.

(i) Calculate the expected number of accidents.

(ii) Make the probability distribution of $y = 3x + 2$ and also calculate $E(3X + 2)$

CamScanner

(iii) Verify that $E(3X + 2) = 3E(X) + 2$

Karachi Board 1992

4.20 Prove $E(aX + b) = aE(X) + b$, where X is a random variable and a and b are constants.

Karachi Board 2003, 2002, 2000

4.21 (a) Define (i) Random Variable (ii) Mathematical Expectation

(b) A random variable 'X' takes the values $-1, 0, 1$ and 2 with probabilities 0.25, 0.15, 0.30 and 0.30 respectively, find E(X).

Karachi Board 2010

4.22 (a) Differentiate between a variable and a Random Variable and give an example.

(b) Prove that $E[X - E(X)] = 0$

(c) A, B and C toss a coin in succession for a prize of Rs 126/- for obtaining the head first. Find their mathematical expectation.

Karachi Board 1994

4.23 (a) Prove that $E(X + Y) = E(X) + E(Y)$.

(b) Let X have the following probability distribution

x	1	2	3	4	5
P(x)	0.2	0.3	0.25	0.15	0.1

Find $E(X^2 + 2X - 3)$

Karachi Board 1987

4.24 (a) Let X and Y be distributed independently. Show that $E(X + Y) = E(X) + E(Y)$.

(b) If X is distributed with mean 4 and variance 16, calculate mean and variance of $Y = -3X + 12$

Karachi Board 1988

4.25 (a) Define Mathematical Expectation. Justify that the mean of the random variable X is the same as the expected value of X.

CamScanner

(b) If you buy a ticket in a lottery in which there is a prize of \$1000 and ten prizes of \$50 each and if 1000 tickets are sold find the expected value of your ticket. When is the game said to be fair?

Karachi Board 1996

4.26 (a) Use the properties of expected value to show the following:

(i) $V(c) = 0$, where c is any constant.

(ii) $V(cx) = c^2V(x)$

(b) Compute mean and variance of the table given below:

x	1	2	3	4
$P(X = x)$	3/8	1/8	1/4	1/4

Karachi Board 2000

4.27 (a) Show that $E(X + Y) = E(X) + E(Y)$, where 'X' and 'Y' are the discrete random variables.

Karachi Board 1997, 2011 (Supp.)

(b) The random variable X has the probabilities as shown in the following table

x	−2	−1	0	1	k
P(x)	0.1	0.1	0.3	0.4	0.1

Find the value of k if $E(X) = 0.3$

Karachi Board 2008

(c) X is the random variable 'the score on a die' Y is the random variable number of heads obtained when two coins are tossed. Find

(i) V(X) (ii) V(Y) (iii) V(X + Y)

Karachi Board 2001

4.28 (a) Define Mathematical Expectation and discuss its properties.

Karachi Board 1998

CamScanner

(b) Show that $V(aX) = a^2 V(X)$

(c) The random variable 'X' representing the number of chocolate chips in a cookie has the following probability distribution

x	2	3	4	5	6
P(x)	0.01	0.25	0.40	0.30	0.04

Find V(X).

Karachi Board 2003

(d) Suppose x and y are two random variables with the following joint probability density.

y \ x	2	4
1	0.10	0.15
3	0.20	0.30
5	0.10	0.15

Find: E(X), E(Y), E(XY) and E(X + Y)

Karachi Board 2010 (Supp.)

4.29 (a) Define Bernoulli Trials.

(b) Show that $\sum_{x=0}^{n} {}^{n}C_{x} \cdot p^{x} \cdot q^{n-x} = 1$

Karachi Board 1997

(c) Define and derive Binomial Probability Distribution.

(d) A and B play a game in which A's probability of winning is 2/3. In a series of 8 games, what is the probability that A wins exactly 4 games.

Karachi Board 2003

CamScanner

(c) What is the probability of obtaining less than two heads in six throws when a coin is tossed.

Karachi Board 2002

4.30 (a) Define binomial probability distribution and write down some important properties of this distribution.

(b) In a binomial distribution $q = \frac{2}{5}$ and n = 10, write down the full expansion of $(q + p)^n$ and find the mean and standard deviation of the distribution.

Karachi Board 1988

4.31 What is the probability of obtaining exactly 3 heads in 8 throws with a single coin.

Karachi Board 2001

4.32 If A and B play a game in which the probability that A wins is $\frac{2}{3}$. In a series of 8 games what is the probability that:

(i) First three games are won by A.

(ii) B wins at least two games.

Karachi Board 1997

4.33 (a) Write down the properties of a binomial distribution.

(b) For a binomial distribution given below:

$$P(x) = {}^{n}C_{x} \cdot p^{x} \cdot q^{n-x}, \; x = 0, 1, 2, \dots\dots\dots, n$$

Prove that E(X) = np

Karachi Board 1998, 2007

4.34 (a) Find Mean and Variance of the following probability distribution:

$$P(X = x) = {}^{4}C_{x} \cdot (0.6)^{x} (0.4)^{4-x}, \text{ where } x = 0, 1, 2, 3, 4$$

Karachi Board 1997

CamScanner

(b) The probability distribution of a discrete random variable X is:

$$f(x) = {}^3C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, \quad x = 0, 1, 2, 3;$$ find the mean and the variance.

Karachi Board 2002

4.35 (a) Calculate Mean and Variance of the binomial distribution if

(i) n = 60 p = 0.5 (ii) n = 100 p = 0.1

Karachi Board 1998

(b) For the following probability distribution

$$P(x) = \binom{9}{x}\left(\frac{1}{2}\right)^{9-x}\left(\frac{1}{2}\right)^x, \; x = 0, 1, 2, \ldots\ldots, 9$$

fill in the blanks:

(i) The name of the distribution is __________

(ii) The mean of the distribution is __________

(iii) The variance of the distribution is __________

(iv) The number of trials is ________

(v) The nature of the variable in the distribution is ________

Karachi Board 1988 (Sup.)

4.36 (a) Derive Binomial Probability Distribution.

(b) If on the average the rain falls on twelve days in every thirty days, find the probability that

(i) The first three days of a given week will be fine and remaining wet.

(ii) The rain falls just three days in a given week.

Karachi Board 1996

4.37 (a) Write the properties of Binomial experiment.

(b) Define Binomial Probability Distribution.

CamScanner

(c) Show that the mean and variance of the binomial distribution $b(x\ ;\ n, p)$ are 'np' and 'npq' respectively.

Karachi Board 1995, 2004

4.38 (a) Briefly narrate the necessary conditions when an experiment is called a binomial experiment.

(b) An estate agent claims to provide suitable accommodation to 60% of his clients. If on a particular occasion 8 clients approach him independently, calculate the probability that at the most 6 clients will be accommodated satisfactorily.

Karachi Board 1993

4.39 (a) Define Binomial Distribution and find its mean.

(b) The probability of a defective item produced by a machine is 0.20. Find the probability that at least 3 of the next 5 items produced by the machine will be defective.

Karachi Board 1997

4.40 Assuming that the sexes are equally likely, find the distribution of the number of boys in a family of 5 children. Also calculate the average number of boys expected in this family.

Karachi Board 1994

4.41 (a) Find the mean and variance of Binomial Distribution.

(b) The probability that a patient recovers from a delicate heart operation is 0.8. What is the probability that exactly 5 of the next 9 patients having this operation will survive.

Karachi Board 1991

4.42 (a) Derive the binomial probability distribution.

(b) In a binomial distribution the mean and the standard deviation were found to be 36 and 4.8 respectively; find 'p' and 'n'.

CamScanner

(c) Is it possible to have a binomial distribution with mean 5 and standard deviation 3? Why?

Karachi Board 1990

4.43 (a) Derive the expressions for mean and variance of a Binomial Probability Distribution.

(b) In a Binomial Distribution, n = 5 and p = 0.5, find $P(x \geq 2)$

Karachi Board 1989

4.44 (a) Derive an expression for the probability of x successes in an experiment of n trials having a constant probability of success in each trial.

(b) Find the mean of the following distribution.

$$P(x) = \binom{3}{x}\left(\frac{1}{2}\right)^3, \qquad x = 0, 1, 2, 3$$

Karachi Board 1986

4.45 (a) In a binomial distribution, n = 8, p = 0.5; find the following:

(i) Mean (ii) Variance (iii) Standard Deviation (iv) Coefficient of Variation

Karachi Board 2009 (Supp.)

(b) Calculate mean and variance from the following probability distribution

x	5	10	15	20	25
P(x)	0.05	0.25	0.40	0.25	0.05

Karachi Board 2004

4.46 (a) Define binomial probability distribution.

(b) Derive Mean and Variance of binomial distribution.

Karachi Board 1985

CS CamScanner

(c) Determine mean and variance of the binomial distribution

(i) $n = 100$ and $p = 0.1$ (ii) $n = 60$ and $q = 0.5$

Karachi Board 2003

4.47 **(i)** What do you mean by Binomial Frequency Distribution.

(ii) 5 dice are shown 243 times. How many times at least are 4 dice expected to show a 3 or 5.

Karachi Board 1994

4.48 **(a)** The probability distribution of a discrete random variable 'X' is

$$f(x) = \binom{4}{x}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{4-x}, \qquad x = 0, 1, 2, 3, 4$$

Find E(X) and V(X).

(b) Derive Binomial Distribution.

Karachi Board 1992, 2006 (Supp.)

(c) Form a complete Binomial Probability Distribution of a number of heads if a coin is tossed 4 times.

(d) Find E(X) and V(X) for the above Binomial Probability Distribution.

Karachi Board 2002

4.49 **(a)** State and explain Binomial Distribution. Also give three examples of this distribution.

(b) Find μ and σ for each of the following binomials.

(i) $\left(\frac{1}{2}+\frac{1}{2}\right)^{7}$ (ii) $\left(\frac{1}{6}+\frac{5}{6}\right)^{4}$

(c) Show that $\left(\frac{1}{3}+\frac{2}{3}\right)^{18}$ may be written as $\sum_{x=0}^{18} f(x)$, where $f(x) = \binom{18}{x}\frac{2^{x}}{3}18$

Karachi Board 1996

CS CamScanner

4.50 (a) Derive the expression for binomial probability distribution.

(b) Find the probability that in five tosses of a fair die 3 appears:

(i) twice on them (ii) at most once

Karachi Board 2003

(c) Briefly narrate the necessary conditions when an experiment is called a binomial experiment.

Karachi Board 1993

4.51 A coin is tossed 7 times; find the probability of getting at least 5 heads.

Karachi Board 1998

4.52 (a) Define Binomial Probability Distribution. Obtain mean and variance of the binomial distribution.

(b) What is the probability of getting exactly 4 heads in 6 tosses of a fair coin?

Karachi Board 2002

4.53 (a) An unbalanced coin is tossed 5 times. If X is a random variate showing the number of heads, then:

(i) Construct the binomial distribution for X if the probability of head in a single toss is 2/3.

(ii) Find the Mean and Variance of X.

Karachi Board 2001

(b) Define Binomial Distribution with an example.

(c) A bent coin has a probability of falling heads equal to 0.4. This coin is tossed 5 times. What is the probability of getting at least 3 heads.

Karachi Board 2003

4.54 (a) Write the properties of binomial distribution.

(b) When binomial distribution becomes a symmetrical distribution.

CS CamScanner

(c) If mean of binomial distribution is 50 and variance is 25. Find parameters of binomial distribution.

Karachi Board 2000

(d) In a Binomial Distribution, the mean and standard deviation were found to be 10 and 2 respectively; find p and n of the Binomial Distribution.

Karachi Board 2003

4.55 (a) Derive the expressions for the Mean and Variance of a Binomial Probability Distribution.

Karachi Board 2001, 2002, 2003

(b) Find the Mean and Variance of a Binomial Distribution.

(c) In a Binomial Distribution the Mean and Standard Deviation were found to be 36 and 4.8 respectively; find n and p.

Karachi Board 2000

4.56 (a) Define Binomial Probability Distribution and give its derivation.

(b) Show that the mean and variance of a Binomial Distribution are np and npq.

Karachi Board 1999

(c) A national centre for health statistics report based on 1985 data states that 30% of American adults smoke. Consider a sample of 15 adults selected at random. Find the probability that the number of smokers in the sample would be:

(i) Three (ii) Less than five

(d) Calculate the coefficient of variation for the binomial distribution with $n = 10$ and $p = 0.5$

Karachi Board 2001

4.57 (a) Define a Normal Distribution.

(b) The mean and variance of a normal distribution are 50 and 25. Find the following probabilities:

(i) $P(X \leq 60)$ (ii) $P(X > 40)$ (iii) $P(39 \leq X \leq 65)$

CS CamScanner

4.58 (a) Let X be a normally distributed random variable with mean 200 and variance 100; calculate the following:

(i) $P(X < 179)$ (ii) $P(X < 214)$

Karachi Board 1998

(b) The burning time of an experimental rocket is a random variable having the normal distribution with mean = 4.76 seconds and standard deviation = 0.04 seconds. What is the probability that this kind of rocket will burn:

(i) In less than 4.66 seconds (ii) In more than 4.80 seconds

(iii) Between 4.70 to 4.82 seconds

Karachi Board 2003

4.59 (a) Write down the properties of Normal distribution.

(b) Let X be a normally distributed random variable with mean 50 and standard deviation 10. Calculate the following:

(i) $P(45 \leq X \leq 62)$ (ii) $P(X \geq 64)$ (iii) $P(X \leq 40)$

Karachi Board 1990

4.60 (a) Define Normal Distribution. What are its important properties?

(b) A normal random variable X has mean 62 and variance 64; determine the probability that X lies between 46.32 and 77.68.

Karachi Board 2002

4.61 (a) Write down the properties of Normal Distribution.

(b) If X is a normally distributed variate with $\mu = 50$ and $\sigma^2 = 25$, find the probability that:

(i) $0 \leq X \leq 40$ (ii) $55 \leq X \leq 100$ (iii) $X > 54$ (iv) $X < 57$

Karachi Board 2001

CamScanner

4.62 (a) Write down the properties of Normal distribution.

(b) Let X be a Normally Distributed Random Variable with mean 70 and standard deviation 10, calculate the following probabilities:

(i) $P(62 \leq X \leq 77)$ (ii) $P(56 \leq X \leq 63)$ (iii) $P(X \geq 81)$

Karachi Board 1989

4.63 (a) Write the properties of Normal distribution.

(b) Let X be a random variable normally distributed with Mean 250 and Standard Deviation 25. What is the probability that:

(i) X lies between 250 and 260 (ii) X is greater than 260

(iii) X is smaller than 230.

Karachi Board 1993

4.64 (a) Define Normal Distribution and Standard Normal Distribution.

(b) In a Normal Distribution the mean is equal to 20 and the standard deviation is 5, find the area:

(i) between 14 and 22 (ii) above 15 (iii) below 16

Karachi Board 1988

4.65 (a) State the properties of Normal Probability Distribution.

(b) A random variable X is normally distributed with mean 60 and variance 36; find the probability that

(i) It will fall between 54 and 66, (ii) It will be larger than 69

(iii) It will be smaller than 57.

Karachi Board 1995

4.66 (a) Briefly state the properties of a Normal Distribution.

(b) Given $\mu = 500$, $\sigma = 80$, find the following probabilities:

(i) $P(X > 600)$ (ii) $P(X < 600)$ (iii) $P(400 \leq X \leq 550)$

Karachi Board 1994

CS CamScanner

4.67 (a) Define Normal Distribution.

(b) The mean height of a number of soldiers is 68.12 inches with a variance of 10.8 $(\text{inches})^2$. Assuming the distribution of height is normal, how many soldiers in a regiment of 1000 soldiers would you expect to be:

(i) over 72 inches (ii) below 68 inches

(iii) between 66 inches and 73 inches.

Karachi Board 1997

4.68 (a) Write the properties of Normal Distribution.

(b) The mean length of 500 laurel leaves from a certain bush is 151mm and the standard deviation is 15mm. Assuming that the length of the leaves is normally distributed find how many leaves measure the following

(i) Between 120mm and 150mm

(ii) More than 185mm

(iii) Smaller than 110mm

Karachi Board 1997

(c) The weights of certain population of young adult females are approximately normally distributed with mean 132 pounds and standard deviation 15 pounds. Find the probability that a female selected at random from this population will weight:

(i) More than 155 pounds (ii) 100 pounds or less

(iii) Between 105 and 143 pounds

Karachi Board 2001

(d) Define Normal distribution and give its properties.

(e) 1000 students appeared at a certain examination. The average marks obtained by them were 40% and standard deviation was 10%. How many students do you expect to get

(i) More than 50% marks (ii) More than 35% marks

Karachi Board 1988

CS CamScanner

4.69 (a) Discuss the relationship between binomial distribution and normal distribution.

(b) Write down the properties of normal distribution.

(c) If a normal distribution has mean = 10 and a variance = 1 then find $P(6 \le x \le 14)$ without using statistical table.

Karachi Board 2000

4.70 (a) Write down the properties of Normal Distribution.

(b) If $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-8)^2}$, where $-\infty \le x \le \infty$.

Find its mean and variance.

Karachi Board 1988

4.71 (a) Write down the properties of the normal distribution.

(b) The mean and variance of a normal random variable X are 100 and 36. Find the probabilities:

(i) $P(X \le 115)$ (ii) $P(X \ge 85)$ (iii) $P(85 \le X \le 115)$

(iv) $P(91.84 \le X \le 102.58)$

Karachi Board 2004

(c) Given a normal distribution with mean 200 and variance 100 find.

(i) The area below 214 (ii) The area above 179

(iii) The area between 188 and 206

Karachi Board 2003

(d) The Mean and Variance of a normal random variable are 50 and 25 respectively. Find the probabilities of the following

(i) $P(X \le 62.5)$ (ii) $P(X \ge 37.5)$ (iii) $P(40.2 \le X \le 59.8)$

Karachi Board 2004

CS CamScanner

(e) Given a normal distribution with mean 255 and variance 625, find the probability that 'X' assumes a value

(i) between 235 and 260 (ii) more than 265 (iii) less than 230

Karachi Board 2003

4.72 (a) Define Normal Distribution and Standard Normal Distribution.

(b) Let X be a normally distributed random variable with mean 70 and standard deviation 10, calculate the following

(i) $P(62 \le X \le 77)$ (ii) $P(X > 37)$

Karachi Board 1998

(c) An electrical firm manufactures light bulbs that have a length of life that is normally distributed with mean equal to 800 hours and standard deviation of 40 hours; find the probability that a bulb lights:

(i) between 778 and 834 hours (ii) more than 840 hours

Karachi Board 2002, 2006

(d) An institute gives aptitude tests to B. S. programmers. The scores of the tests are normally distributed with mean 255 and variance 625. What is the probability that a score selected at random will be

(i) between 235 to 260 (ii) less than 230 (iii) more than 265

Karachi Board 2002

4.73 (a) Under what conditions does a binomial distribution approach a normal distribution.

(b) Give the equation of normal curve with mean = zero and variance = 1

(c) N. E. D. University arranged a test for admission. The score obtained by the students were normally distributed with mean 250 and standard deviation 10. What is the probability that a score selected at random will be

(i) between 235 and 256

(ii) more than 265

CamScanner

(iii) less than 240

Karachi Board 1999

4.74 **(a)** Explain fully as you can, the relationship between Binomial and Normal Distribution.

(b) (i) Define standard Normal Variate (Z).

(ii) Prove that:

$E(Z) = 0, \quad V(Z) = 1$

Hence comment on the role of Z in Statistics.

Karachi Board 2000, 2002

(c) Describe five important properties of a Normal Distribution. (Proof is not required).

(d) If X is a Normal variate with Mean 10 and Variance 16, find:

(i) The value of Standard Normal Variate z when $X = 5$

(ii) The value of Normal Variate X when $z = 2.4$

Karachi Board 2000

(e) If X is a normal variate with mean 10 and variance 16, find the value of a standard normal variate z when $x = 5$

Karachi Board 1992

4.75 **(a)** Write down five important properties of Normal Probability Distribution.

(b) A production supervisor found that employees on the average, complete a certain task in 10 minutes. The times required to complete the tasks are approximately normally distributed with a variance of 9 square minutes. Find the proportion of employees completing the task in more than 10 minutes.

Karachi Board 2001

CS CamScanner

4.76 (a) Explain Normal Distribution with its properties.

(b) Under what conditions does a Binomial Distribution tend to be a Normal Distribution. Explain.

Karachi Board 1999

4.77 For the following Continuous probability distribution.

$$f(x) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-10}{5}\right)^2}, \quad -\infty \le x \le \infty$$

Fill in the blanks:

(i) The mean of the distribution is __________.

(ii) The name of the distribution is __________.

(iii) The variance of the distribution is __________.

(iv) The distribution is symmetrical about __________.

(v) All odd order moments of the distribution are __________.

(vi) The distribution is a limiting case of __________.

(vii) The maximum ordinate of the distribution is __________.

(viii) The maximum ordinate at the point x = is __________.

(ix) The nature of the variable in the distribution is __________.

(x) The graph of the distribution is __________ shaped.

Karachi Board 1986

4.78 (a) State and Prove Binomial Probability Distribution.

(b) Show that $\sum_{x=0}^{n} {}^nC_x \cdot p^x \cdot q^{n-x} = 1$

(c) State under what conditions the binomial distribution tends to normal distribution.

Karachi Board 1984

4.79 (a) The mean and standard deviation of a normal random variable are 40 and 5 respectively, find the following areas

(i) to the left of 49

(ii) to the right of 55

(b) Sketch Normal Curves to show the above areas in (a).

Karachi Board 2002

4.80 (a) Define Normal Distribution.

(b) Using the Normal Distribution:

$$f(x) = \frac{1}{100\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-500}{100}\right)^2} \quad \text{where } -\infty < x < \infty$$

Calculate the following probabilities

(i) P(x > 650) (ii) P(x < 250) (iii) P(325 < x < 675)

Karachi Board 1992

4.81 If X is normally distributed with mean μ and standard deviation σ, find the probability:

(a) $P(\mu - \sigma \le X \le \mu + \sigma)$ (b) $P(\mu - 2\sigma \le X \le \mu + 2\sigma)$

Karachi Board 2001

4.82 (a) Under what conditions does a binomial distribution tend to normal distribution? What are the true values of its measures of skewness and kurtosis for a normal distribution.

(b) The total time an ambulance takes to reach its destination in emergency calls is a normal variable with mean 8.9 minutes and standard deviation 1.8 minutes. What is the probability for any call that will take

(i) less than 7 minutes (ii) between 6 minutes and 7 minutes.

Karachi Board 1996

CamScanner

(c) For the given p.d.f. $f(x)=\frac{1}{\sqrt{2\pi}15}e^{-\frac{1}{2}\left(\frac{x-100}{15}\right)^2}$, find the probabilities of the following:

(i) $P(x \geq 124)$ (ii) $P\left[|x-100|<\sqrt{50}\right]$ (iii) $P(91<x<127)$

Karachi Board 2006 (Supp.)

4.83 (a) Define Normal Distribution.

(b) Under what conditions does binomial distribution tend to be normal distribution.

(c) A fair coin is tossed 80 times. Find the probability for the following outcomes:

(i) The number of heads is between 40 and 50 including both.

(ii) The number of heads is less than 55

(iii) The number of heads and tails is equal.

Karachi Board 1992

4.84 A pair of dice is rolled 180 times. What is the probability that:

(i) the total of 7 occurs at least 25 times.

(ii) the total is between 33 and 41 times inclusive.

(iii) the total is exactly 30 times.

Karachi Board 2007 (Supp.)

4.85 Write short notes on the following

(i) Normal Distribution (ii) Mathematical Expectation

(iii) Binomial Distribution (iv) Discrete and Continuous Random Variables

(v) Random Variable.

CamScanner

Chapter – 5

Sampling

Introduction

The process of selecting a small part from a large collection, such that the selected part will show all the characteristics of the large collection is called sampling. The small part which is taken out from a large collection is called "Sample" and the large collection is called "Population".

Sampling is quite often used in our day-to-day practical life. For example, in a shop we assess the quality of rice, sugar and any other commodity by taking a handful of it from the large and then decide to purchase it or not. A house wife normally tests the cooked products to find if they are properly cooked and contain the proper quantity of salt.

There are so many techniques to draw a sample from a population. Here we shall discuss only the techniques of Simple Random Sampling, Stratified Random Sampling and the Systematic Sampling.

Before introducing these sampling techniques some of the important terms and the basic concepts are defined first.

Population

A complete collection of individuals, objects or measurements under consideration in a statistical study will be called a population.

OR

A population may be defined as the large collection of similar units.

For example, a population may be of heights of all students of a college, a population may be of all books in a library, a population may be of all persons using a particular medicine, a population may be of prices of a commodity over a period of time, etc.

CamScanner

The number of units in the population (i.e. size of a population) is always denoted by N.

A population may be finite or infinite depending upon whether it contains a countable number of units or an uncountable number of units.

Sample

A small part of a population is called sample. The number of units in the sample (i.e. size of a sample) is always denoted by n.

Parameter

Any quantity calculated from a population is called a parameter. For example, population mean, population variance, etc. are therefore parameters. The population mean is denoted by μ and the population variance is denoted by σ^2.

Statistic

Any quantity calculated from a sample is called a statistic. For example, sample mean, sample variance, etc. are therefore statistics. The sample mean is denoted by $\overline{X}$ and the sample variance is denoted by s^2.

Census and Sample Survey

Census or Complete Enumeration means to get the information about each and every unit in the population.

For example, if we wish to know the average income of all doctors in Karachi, then the data are to be obtained about the annual income of all doctors in Karachi from all hospitals and clinics in Karachi. This technique of collecting data regarding all individuals in the population is called Census.

A complete census is an expensive operation as it involves large number of enumerators, supervisory and administrative staff. In view of the large volume of work and other administrative and fiscal problems, a census can only be conducted after a long period of time.

A Sample Survey is a technique of getting information about the characteristics of the population by studying only a part (i.e. by studying a sample) of the population.

CS CamScanner

Here we collect the data about only a few selected units from the population. The set of selected units is the sample and method of collecting data about them is called Sample Survey.

For example, if we want to find the height of all 10,000 students studying in the University of Karachi, we may select a sample of 100 students from them and on the basis of mean height of these 100 students we can estimate the mean height of all 10,000 students. Here the survey in which the data regarding heights of 100 students are obtained is a sample survey.

Advantages of Sampling

The important advantages of sampling are listed below:

(i) Sampling method is cheaper to collect information as compared to census (i.e complete enumeration).

(ii) The data may be collected, classified and analysed much more quickly with a sample than with a census enquiry.

(iii) A sample is often used as a check to verify the accuracy of complete count.

(iv) It provides greater accuracy because the volume of work is reduced in the sample survey.

Sampling with and without replacement

Sampling is said to be with replacement when we draw a unit from a finite population and return it to the population before the next unit is drawn. In this case each unit can be drawn more than once and the probability of drawing of each unit remains constant throughout the sampling procedure. It is important to note that, in case of sampling with replacement the population is considered infinite.

Sampling is said to be without replacement, if we do not return the selected unit to the population and draw the next unit. In this case each unit cannot be drawn more than once and the probability of drawing of each unit changes throughout the sampling procedure.

Sampling Techniques

Now we shall discuss the following three important sampling techniques one by one.

(i) Simple Random Sampling

CamScanner

(ii) Stratified Random Sampling

(iii) Systematic Sampling

Simple Random Sampling

This is the simplest and the easiest method of drawing a sample from a population. According to this method each and every unit in the population has an equal chance of being included in the sample and also each possible sample of the same size has an equal probability of being chosen.

Suppose, there is a population of 'N' units and we want to draw a sample of 'n' units, then the possible number of samples in case of sampling without replacement will be ${}^{N}C_{n} = \frac{N!}{n!\,(N-n)}$ and in case of sampling with replacement the possible number of samples will be N^{n}.

As an illustration, suppose there is a population of 4 persons, identified as A, B, C and D and we want to select a random sample of 2 persons.

Therefore, $N = 4, n = 2$ and there will ${}^{N}C_{n} = {}^{4}C_{2} = \frac{4!}{2!\,(4-2)} = 6$ possible samples in case of sampling without replacement.

The possible samples are: AB, AC, AD, BC, BD, CD

When the sampling is done with replacement, the possible samples will be $N^{n} = 4^{2} = 16$ which are listed below:

AA	BA	CA	DA
AB	BB	CB	DB
AC	BC	CC	DC
AD	BD	CD	DD

In practice a simple random sample is drawn unit by unit by means of "Lottery System" or by the use of "Random Numbers Table" or by the use of "Computer Programs" that provide Random Numbers. The detail is not needed to be discussed here as it is supposed to be beyond of the scope of the book.

Note: Simple Random Sampling can be referred to as Random Sampling and the sample obtained by this procedure is called Random Sample.

CamScanner

Selection of a Simple Random Sample

Suppose there is a population of 2000 students in a college and we want to draw a simple random sample of 100 students from this college. According to this technique first of all we allot serial numbers from 1 to 1000 to all students i.e. 0001, 0002, 0003, ..., 2000 and then prepare 2000 chits of these numbers. After shuffling the chits we draw any 100 chits one by one. The students bearing the numbers on the slips will be included in our required sample. This procedure of making chits and then drawing is called lottery method. This method is not convenient when we have a large population, so we mostly apply computer programs.

Advantages and Disadvantages of Simple Random Sampling

(i) It provides unbiased estimates of population mean, population total and population variance, etc.

(ii) When sample size increases the sample results approach to the population results.

(iii) This method is expensive and time consuming when population is large.

(iv) If the data are not homogenous, a simple random sample might not be representative of the population.

Sampling Distributions

Consider all possible samples of size n which can be drawn from a given population (either with or without replacement). For each sample we can compute a statistic, such as the mean, variance, etc. Which will vary from sample to sample. In this way we obtain a distribution of the statistic which is called its sampling distribution. Therefore, the sampling distributions may be of mean, variance, etc.

Sampling Distribution of the mean ($\bar{x}$)

If we draw all possible samples each of size n from a finite population of N units with mean μ and variance σ^2. Then we compute mean of every sample i.e. $\overline{X}$. Therefore, the statistic ($\overline{X}$) is now a random variable and form a probability distribution. This distribution is called sampling distribution of mean.

The sampling distribution of mean has the following properties in case of sampling without replacement and in case of sampling with replacement.

Properties in case of Sampling without replacement

1. Mean of all sample means is equal to the population mean i.e. $E(\overline{X}) = \mu$

OR

$\overline{X}$ is said to be an Unbiased Estimator of μ.

2. $V(\overline{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$ where σ^2 is the population variance.

Properties in case of Sampling with replacement

1. Mean of all sample means is equal to the population mean i.e. $E(\overline{X}) = \mu$

OR

$\overline{X}$ is said to be an Unbiased Estimator of μ.

2. $V(\overline{X}) = \frac{\sigma^2}{n}$ where σ^2 is the population variance.

Standard Error

The standard deviation of sampling distribution of $\overline{X}$ is called standard error of $\overline{X}$.

Thus S. E. of $(\overline{X}) = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ in case of without replacement

and S. E. of $(\overline{X}) = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ in case of with replacement

Example 5.1

A population consists of five numbers 0, 2, 4, 6, 8

(a) List all possible samples of size 2 that can be drawn from this population without replacement.

(b) Find mean of each sample.

(c) Construct sampling distribution of $\overline{X}$.

(d) Verify that mean of all sample means is equal to population mean.

CamScanner

Solution:

Since $N = 5$ and $n = 2$ and the sampling is done without replacement, then all possible samples $= {}^5C_2 = \frac{5!}{2!\,(5-2)} = 10$ as shown below;

Sample No.	Ail Possible Samples	Sample Mean $\bar{x}$
1	(0, 2)	1
2	(0, 4)	2
3	(0, 6)	3
4	(0, 8)	4
5	(2, 4)	3
6	(2, 6)	4
7	(2, 8)	5
8	(4, 6)	5
9	(4, 8)	6
10	(6, 8)	7
Total	–	40

Sampling Distribution of $\bar{x}$ is

$\bar{x}$	1	2	3	4	5	6	7
f	1	1	2	2	2	1	1

Now find mean of all means i.e. mean of sampling distribution of $\bar{x}$ is computed as

$\bar{x}$	f	$f\bar{x}$
1	1	1
2	1	2
3	2	6
4	2	8
5	2	10
6	1	6
7	1	7
Total	10	40

The Mean of $\overline{X} = E(\overline{X}) = \dfrac{\sum f\bar{x}}{\sum f} = \dfrac{40}{10} = 4$

and since, population mean $\mu = \dfrac{0+2+4+6+8}{5} = 4$

Therefore, it is verified that, mean of all sample means is equal to population mean.

Example 5.2

A population consists of five numbers 0, 3, 6, 9, 12

(a) List all possible samples of size 3 that can be drawn form this population without replacement.

(b) Verify that, Mean of $\overline{X} = E(\overline{X}) = \mu$ and $V(\overline{X}) = \dfrac{\sigma^2}{n} \cdot \dfrac{N-n}{N-1}$

Solution:

Since $N = 5$ and $n = 3$ and sampling is done without replacement. Then all possible samples $= {}^5C_3 = \dfrac{5!}{3!\,(5-3)} = 10$

Population mean and variance are computed as

$$\mu = \frac{\sum x}{N} = \frac{30}{5} = 6$$

$$\sigma^2 = \frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2$$

$$\sigma^2 = \frac{270}{2} - \left(\frac{30}{5}\right)^2 = 18$$

	x	x^2
	0	0
	3	9
	6	36
	9	81
	12	144
Total	30	270

The mean and variance of $\overline{X}$ are computed as

All Possible Samples	Sample Mean $\bar{x}$	$\bar{x}^2$
(0, 3, 6)	3	9
(0, 3, 9)	4	16
(0, 3, 12)	5	25
(0, 6, 9)	5	25
(0, 6, 12)	6	36
(0, 9, 12)	7	49
(3, 6, 9)	6	36
(3, 6, 12)	7	49
(6, 9, 12)	9	81
(3, 9, 12)	8	64
Total	60	390

Mean of $\overline{X} = E(\overline{X}) = \frac{\sum \bar{x}}{m} = \frac{60}{10} = 6 = \mu$, where m = No. of samples

Therefore, it is verified that $E(\overline{X}) = \mu$

Now, $V(\overline{X}) = \frac{\sum \bar{x}^2}{m} - \left(\frac{\sum \bar{x}}{m}\right)^2$ where m = number of samples

$$V(\overline{X}) = \frac{390}{10} - \left(\frac{60}{10}\right)^2 = 3$$

Also $V(\overline{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1} = \frac{18}{3} \cdot \frac{5-3}{5-1} = 3$

Hence it is verified that $V(\overline{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$

Example 5.3

Draw all possible samples each of size 2 from the population 2, 4, 6 and 8 using sampling with replacement. Find mean of each sample and verify that

(i) Mean of $\overline{X} = \mu$ and (ii) $V(\overline{X}) = \frac{\sigma^2}{2}$

Solution:

Firstly, we compute Population mean and variance

$$\mu = \frac{\sum x}{N} = \frac{20}{4} = 5$$

$$\sigma^2 = \frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2$$

$$\sigma^2 = \frac{120}{4} - \left(\frac{20}{4}\right)^2 = 5$$

	x	x^2
	2	4
	4	16
	6	36
	8	64
Total	20	120

Since, $N = 4$ and $n = 2$ then, possible samples $= N^n = 4^2 = 16$

The possible samples and their means are computed as

Possible Samples	Sample Mean $\bar{x}$	$\bar{x}^2$
(2, 2)	2	4
(2, 4)	3	9
(2, 6)	4	16
(2, 8)	5	25
(4, 2)	3	9
(4, 4)	4	16
(4, 6)	5	25
(4, 8)	6	36
(6, 2)	4	16
(6, 4)	5	25
(6, 6)	6	36
(6, 8)	7	49
(8, 2)	5	25
(8, 4)	6	36
(8, 6)	7	49
(8, 8)	8	64
Total	80	440

Mean of $\overline{X} = \dfrac{\sum \bar{x}}{m} = \dfrac{80}{16} = 5 = \mu$, Therefore, Mean of $(\overline{X}) = \mu$

Now, $V(\overline{X}) = \dfrac{\sum \bar{x}^2}{m} - \left(\dfrac{\sum \bar{x}}{m}\right)^2$ where m = number of samples

$$V(\overline{X}) = \frac{440}{16} - \left(\frac{80}{16}\right)^2 = 2.5$$

Also $V(\overline{X}) = \dfrac{\sigma^2}{n} = \dfrac{5}{2} = 2.5$

Hence it is also verified that $V(\overline{X}) = \dfrac{\sigma^2}{n}$

Example 5.4

Draw all possible samples each of size $n = 3$ from the population 2, 11 by using sampling with replacement. Find standard error of sampling distribution of $\bar{X}$.

Solution:

Since $N = 2$ and $n = 3$ then possible samples $= N^n = 2^3 = 8$

Possible Samples	Sample Mean $\bar{x}$
(2, 2, 2)	2
(2, 2, 11)	5
(2, 11, 2)	5
(11, 2, 2)	5
(2, 11, 11)	8
(11, 2, 11)	8
(11, 11, 2)	8
(11, 11, 11)	11

Sampling distribution of $\bar{x}$ is

$\bar{x}$	2	5	8	11
f	1	3	3	1

Now compute Standard Error of $\bar{x}$

$\bar{x}$	f	$f\bar{x}$	$f\bar{x}^2$
2	1	2	4
5	3	15	75
8	3	24	192
11	1	11	121
Total	8	52	392

CS CamScanner

$$\text{S. E. of } (\bar{X}) = \sigma_{\bar{x}} = \sqrt{\frac{\sum f\bar{x}^2}{\sum f} - \left(\frac{\sum f\bar{x}}{\sum f}\right)^2} = \sqrt{\frac{392}{8} - \left(\frac{52}{8}\right)^2} = 2.6$$

Standard Error of $\bar{x}$ may also be calculated as:

$$\text{S. E. of } (\bar{X}) = \frac{\sigma}{\sqrt{n}}$$ where σ is the Population Standard Deviation

Firstly, we compute population standard deviation σ

$$\sigma = \sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2}$$

$$\sigma = \sqrt{\frac{125}{2} - \left(\frac{13}{2}\right)^2}$$

	x	x^2
	2	4
	11	121
Total	13	125

Since $\text{S. E. of } (\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{4.5}{\sqrt{3}} = 2.6$

Example 5.5

If the size of a sample is 64 and standard error of mean is 1.5. What should be the sample if standard error reduced to 0.6

Solution:

Since $n = 64$

Since $\text{S. E. of } (\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{4.5}{\sqrt{64}} = 1.5$

then $\sigma = 12$

if $\text{S. E. of } (\bar{X}) = \frac{\sigma}{\sqrt{n}} = 0.6$

and $\frac{12}{\sqrt{n}} = 0.6$ and $\sqrt{n} = \frac{12}{0.6} = 20$

$n = 400$

CamScanner

Stratified Random Sampling

A simple random sample gives satisfactory results if the population is homogeneous. When the population is non homogeneous (or heterogeneous) but can be divided into a number of homogeneous groups, the method of Stratified Random Sampling is used.

According to this technique first of all we sub-divide the whole of the non homogeneous population into different groups (called Strata), in such a way that each group (called Stratum) has homogeneous units (or members) and select units from each group (or Stratum) by simple random sampling.

Let us understand this technique by the following practical example; suppose we want to estimate the per capita income of a city. For this purpose, we shall divide first the whole population of the city into different income groups e.g. low income group, middle income group and higher income group because the population of incomes is non homogeneous. Then we will select people from each of these income groups by using the technique of simple random sampling and the per capita income of the city will be estimated by the selected stratified random sample in a better form.

Advantages and Disadvantages of Stratified Sampling

(i) Accuracy of the estimate is increased.

(ii) By Stratification (or grouping) a large population is divided in small parts. Therefore, a large population may be studied in a better form.

(iii) To divide the population into homogeneous groups (or strata), it requires more money and time.

(iv) If the groups (or strata) cannot be defined clearly, the strata may overlap, reducing the accuracy of the results.

Systematic Sampling

Suppose there is a population of N units and we want to draw a sample of n units. Suppose all the N units of the population are serially numbered from 1 to N.

According to this technique first of all we shall select one unit at random from first k-units, where $k = \dfrac{N}{n}$ and then every kth unit should be selected. The sample obtained in this way is called Systematic Random Sample.

CS CamScanner

Systematic sampling can be used for populations that are in some kind of order, such as listed populations or populations of records in a file.

Let us understand this technique by the following practical example;

Suppose, there is a population of 50 students and we want to draw a systematic sample of 10 students. Suppose the students in the population are numbered from 1 to 50. Now find k, where $k = \frac{N}{n} = \frac{50}{10} = 5$. Therefore, we shall select first student from first 5 students at random and then every 5th student should be included in the required sample.

Advantages and Disadvantages of Systematic Sampling

(i) It is very simple to use and much speedy.

(ii) It distributes the sample more evenly over the listed population. Therefore it gives more accurate results than simple random sampling.

(iii) It is very economical.

(iv) Systematic sampling has the drawback of numbering or arranging the units in a large population.

CamScanner

EXERCISE – 5

5.1 Explain the following terms:

(i) Population (ii) Sample (iii) Parameter (iv) Statistic

5.2 Explain the difference between the following:

(a) Population and Sample (b) Statistic and Parameter

(c) Census and Sample Survey.

5.3 Differentiate between Census and Sampling.

Karachi Board 1990

5.4 What is the difference between Census and Sampling? Explain with examples.

Karachi Board 1988

5.5 Distinguish between Sample Survey and Population Census.

Karachi Board 2001

5.6 What is the difference between Census and Sample Survey? Explain with examples.

5.7 What are the advantages of sampling?

Karachi Board 1989

5.8 Define Sampling. What are the advantages of Sampling?

Karachi Board 1987, 1988, 1992, 1997, 2002 & 2003

5.9 (a) What is meant by sampling? What are its main objectives.

(b) Define Simple Random Sampling and describe how a simple random sample is selected.

Karachi Board 1999

CamScanner

5.10 **(a)** What do you understand by

(i) Population (ii) Sample (iii) Sampling

(b) Write the advantages of Sampling.

(c) Explain Simple Random Sampling.

Karachi Board 1991

5.11 What is meant by sample and sampling. Describe in detail simple random sampling and give its merits and demerits.

Karachi Board 1986

5.12 **(a)** Define Simple Random Sampling and give its examples.

(b) Draw all possible samples each of size 2 from the population 2, 4, 6, 8 and 10 by using sampling without replacement. Find the mean for each sample and verify that mean of all sample means is equal to population mean.

Karachi Board 1998

5.13 Draw all possible samples of size 2 without replacement from the population 8, 2, 10, 4 and show that $E(\bar{x}) = \mu$

Karachi Board 2000

5.14 When is Simple Random Sampling technique preferred to other techniques.

Karachi Board 1992

5.15 A population consists of five numbers 0, 3, 6, 9 and 12.

(i) List all possible samples of size 3 that can be drawn from this population without replacement.

(ii) Verify that $E(\bar{x}) = \mu$

Karachi Board 2003

5.16 Explain briefly why sampling methods are preferred to Census methods.

Karachi Board 1992

CamScanner

5.17 A population has 5 numbers 8, 10, 12, 15, 20. Take all possible samples of size 2 without replacement. Show that:

(i) $E(\bar{x}) = \mu_{\bar{x}} = \mu$ (ii) $V(\bar{x}) = \sigma^2_{\bar{x}} = \dfrac{\sigma^2}{n} \times \dfrac{N-n}{N-1}$

Karachi Board 1988

5.18 A population consists of the numbers 2, 3, 5, 7, 9

(a) Write down all possible samples of size 2 when sampling is done without replacement.

(b) Find mean of these samples and then verify that:

$E(\bar{x}) = \mu$ and $V(\bar{x}) = \dfrac{\sigma^2}{n} \times \dfrac{N-n}{N-1}$

5.19 (a) Explain the following terms:

(i) Population (ii) Sample

(b) Draw all possible samples of size 3 without replacement from the given population 9, 0, 15, 12, 18 and prove that mean of all sample means is equal to the population mean.

Karachi Board 2004

5.20 (a) Draw all possible samples of size 3 without replacement from the population 0, 3, 12, 18, 30 show that:

(i) Mean of $\bar{x} = \mu$ (ii) $V(\bar{x}) = \dfrac{\sigma^2}{n} \times \dfrac{N-n}{N-1}$

(b) Draw all possible samples of size 3 without replacement from the population 6, 4, 10, 8 and 2. Verify that $E(\bar{x}) = \mu$ and $V(\bar{x}) = \dfrac{\sigma^2}{n} \times \dfrac{N-n}{N-1}$

Karachi Board 2002

5.21 (a) What is a Random Sample? How is it selected?

(b) Draw all possible samples of size
5, 7, 9, 11. Find the mean for e
sample means is equal to popula

5.22 (a) Draw all possible samples of siz
Verify that:

(i) $E(\bar{x}) = \mu$ (ii) $V(\bar{x}) =$

(b) Draw all possible samples of si
15, 21. Show that the mean of a
variance of all sample means is

5.23 (a) What do you understand by Po

(b) A population consists of five nu
size 2 that can be drawn from th

5.24 (a) Differentiate between Simple
Sampling.

(b) Consider a population of size N
were outpatients in a hospital. T

$x_1 = 8, x_2 = 10, x_3 = 12,$

Obtain Standard Error of mean
population with replacement (w

5.25 (a) What are the steps involved in

(b) Define stratified random sam
applicable.

CamScanner

(b) Draw all possible samples of size 2 without replacement from the population 1, 3, 5, 7, 9, 11. Find the mean for each sample and verify that the mean of all the sample means is equal to population mean.

Karachi Board 1988

5.22 (a) Draw all possible samples of size 2 with replacement form the population 9, 15. Verify that:

(i) $E(\bar{x}) = \mu$ (ii) $V(\bar{x}) = \dfrac{\sigma^2}{n}$

(b) Draw all possible samples of size 2 with replacement from the population 9, 11, 15, 21. Show that the mean of all sample means is equal to population mean and variance of all sample means is half of the population variance.

5.23 (a) What do you understand by Population, Sample and Parameter.

(b) A population consists of five numbers 4, 7, 9, 12, 15. List all possible samples of size 2 that can be drawn from the population with and without replacement.

Karachi Board 1998

5.24 (a) Differentiate between Simple Random Sampling and Stratified Random Sampling.

(b) Consider a population of size $N = 5$ consisting of the ages of five children, who were outpatients in a hospital. Their ages are as follows:

$$x_1 = 8, x_2 = 10, x_3 = 12, x_4 = 14, x_5 = 16$$

Obtain Standard Error of mean when a random sample of size 3 is drawn from the population with replacement (without drawing the samples from the population).

Karachi Board 2001

5.25 (a) What are the steps involved in drawing a Stratified Random Sample.

Karachi Board 2001

(b) Define stratified random sampling and describe the situation in which it is applicable.

Karachi Board 1990

CS CamScanner

5.26 (a) Distinguish between Simple Random Sampling and Stratified Random Sampling. When does Simple Random Sampling fail?

(b) Draw all possible Simple Random Samples of size 2 without replacement from the population 1, 2, 3, 4, 5. Also compute $E(\bar{x})$ and $Var(\bar{x})$ and show that $E(\bar{x}) = \mu$.

Karachi Board 2001

5.27 Explain the difference between Simple Random Sampling and Stratified Random Sampling bringing out the advantages of Stratified Random Sampling.

Karachi Board 1989

5.28 Distinguish between Simple Random Sampling and Stratified Random Sampling.

Karachi Board 1984

5.29 (a) Define the terms Sampling and Sample.

(b) Explain Simple Random Sampling and Stratified Random Sampling.

Karachi Board 1991

5.30 (a) What are the advantages of Sampling.

(b) How will you take a sample from the population using stratified random sampling technique.

Karachi Board 1988

5.31 (a) Explain briefly the difference between Simple Random Sampling and Stratified Random Sampling.

(b) Draw all possible samples of size 3, without replacement, from the population 3, 6, 15, 18, 27 and show that

(i) $E(\bar{x}) = \mu$ (ii) $V(\bar{x}) = \dfrac{\sigma^2}{n} \times \dfrac{N-n}{N-1}$

Karachi Board 1999

5.32 (a) What is a Simple Random Sample.

(b) Explain the various methods of drawing random samples.

CS CamScanner

(c) Give one example when stratified random sampling is preferred to simple random sampling.

Karachi Board 1991, 2000

5.33 What are the steps involved in drawing a Stratified Random Sample.

Karachi Board 2001

5.34 (a) Define Sampling. What are the different types of Sampling.

(b) Draw all possible samples of size 3 without replacement from the population $\{0, 6, 3, 12, 15, 19\}$ and verify that $E(\bar{x}) = \mu$ and $V(\bar{x}) = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

Karachi Board 2003

5.35 (a) Define Census and Survey.

(b) Define Stratified Random Sampling.

Karachi Board 2000

5.36 Define Simple Random Sampling and Stratified Random Sampling. Under what situations is stratified random sampling preferred to simple random sampling.

Karachi Board 1988

5.37 (a) What are the methods of selecting Random Samples?

(b) Define Systematic Sampling. When is this method used?

Karachi Board 2002

5.38 (a) What is Random Sampling? What are its advantages?

(b) Describe all the steps taken in drawing the following Samples:

(i) Simple Random Samples

(ii) Systematic Samples

Also compare their advantages and draw-backs.

Karachi Board 2000

CamScanner

5.39 Define Sampling and give its advantages. Give the different types of sampling techniques.

Karachi Board 1985

5.40 (a) What are the advantages of sampling and illustrate it with an example.

(b) Explain Systematic Sampling and illustrate it with an example.

Karachi Board 2004

5.41 What are the different methods of drawing random samples.

Karachi Board 1992

5.42 Write short notes on the following

(i) Simple Random Sampling

(ii) Stratified Random Sampling

(iii) Systematic Sampling

(iv) Parameter

(v) Statistic.

5.43 Draw all possible samples of size 2 from the population 2, 4, 6, 8 using sampling with replacement. Verify that mean of all sample means is equal to the population mean.

Karachi Board 2011 (Supp.)

CS CamScanner

Chapter – 6

Theory of Attributes & Test of Independence

Before going to discuss the format of tests of independence, we give a brief description of the theory of attributes.

Attributes

A characteristic which cannot be measured numerically but only its presence or absence can be described is called an Attribute. The examples of attribute are colour of the eyes of an animal, blood group of a person, religion of a person, sex of a baby, etc. The attributes cannot be measured accurately but they can be divided into classes and their numbers in each class can be counted. If the data are divided into two distinct mutually exclusive (disjoint) classes by a single attribute as for example, the population of students is divided into males and females, the process is called dichotomy (cutting into two). If a class is divided into more than two subclasses, such division or classification is known as manifold classification.

Notations and Terminology

Usually, we use capital letters A, B, C, etc. to denote the presence of attributes and the Greek letters α, β, γ etc. to denote the absence of these attributes respectively. For example, if A represents male then α would represent female. Similarly, if B represents literate then β would represent illiterate. The combinations of different attributes are denoted by AB, $A\beta$, αB and $\alpha\beta$.

where: AB = Males and literates

$A\beta$ = Males and illiterates

CamScanner

αB = Females and literates

$\alpha\beta$ = Females and illiterates

The frequencies of attributes and different combinations of attributes are represented by (A), (α), (B), (β), (AB), (Aβ), (αB). ($\alpha\beta$) and N. Frequency data can be presented in a tabular form called Contingency table.

Attribute	B	β	Total
A	(AB)	(Aβ)	(A)
α	(αB)	($\alpha\beta$)	(α)
Total	(B)	(β)	N

where: $(AB) + (A\beta) = (A)$

$(\alpha B) + (\alpha\beta) = (\alpha)$

$(B) + (\beta) = N$

$(AB) + (\alpha B) = (B)$

$(A\beta) + (\alpha\beta) = (\beta)$

$(A) + (\alpha) = N$

Extraction of More Information

In this sub-section we shall illustrate with the help of example, how the tabular representation of class frequencies or relations between them can be used to extract more information from a given set of class frequencies.

Example 6.1

One thousand persons selected from a certain area revealed that there were 600 males and 430 literates while 250 were illiterate females.

Find the number of (i) females (ii) illiterates and (iii) literate males in the group

Method - I: (Using Tabular Representation)

We firstly prepare a two-way table and write all the given data in the table. The other class frequencies are found and are indicated by the circles.

CamScanner

Attribute	Literates	Illiterates	Total
Males	280	320	600
Females	150	250	400
Total	430	570	1000

Thus we have

(i) Number of females in the group = 400

(ii) Number of illiterates in the group = 570 and

(iii) Number of literate males = 280

Method - II: (Using Relations Between Class Frequencies)

Let A : Males, α : Females, B : Literates and β : Illiterates, then the given data can be written as

$$N = 1000,\ (A) = 600,\ (B) = 430 \text{ and } (\alpha\beta) = 250$$

We want to find (i) No. of females = (α) , (ii) No. of illiterates = (β) and (iii) No. of literate males = (AB)

We have $(\alpha) = N - (A) = 1000 - 600 = 400$

$(\beta) = N - (B) = 1000 - 430 = 570$

$(\alpha B) = (\alpha) - (\alpha\beta) = 400 - 250 = 150$

and $(AB) = (B) - (\alpha B) = 430 - 150 = 280$

Thus, we get

(i) No. of females = $(\alpha) = 400$

(ii) No. of illiterates = $(\beta) = 570$ and

(iii) No. of literate males = $(AB) = 280$

Example 6.2

For two attributes A and B, we have (AB) = 35, (A) = 55, N = 100 and (B) = 65. Calculate the missing values.

CS CamScanner

Solution:

$$(\alpha) = N - (A) = 100 - 55 = 45$$

$$(A\beta) = (A) - (AB) = 55 - 35 = 20$$

$$(\alpha B) = (B) - (AB) = 65 - 35 = 30$$

$$(\beta) = N - (B) = 100 - 65 = 35$$

$$(\alpha\beta) = (\beta) - (A\beta) = 35 - 20 = 15$$

Association of Attributes

Association of attributes, in general, is divided into two parts:

(i) Positive Association

(ii) Negative Association

Positive Association

If the two attributes A and B are so related that

$$(AB) > \frac{(A)\times(B)}{N}$$

it will be known as positive association. For example; association between illiteracy and poverty.

Negative Association

If the two attributes A and B are so related that:

$$(AB) < \frac{(A)\times(B)}{N}$$

it will be known as negative association. For example; association between vaccination and attack of a disease.

Note: It should be remembered that negative association is different from independent because in Independence

$$(AB) = \frac{(A)\times(B)}{N} \quad \text{or} \quad (\alpha\beta) = \frac{(\alpha)\times(\beta)}{N}$$

but in negative association $(AB) < \frac{(A)\times(B)}{N}$

Coefficient of Association

The strength of association between the two attributes A and B is computed by the following formula:

$$Q = \frac{(AB)(\alpha\beta)-(A\beta)(\alpha B)}{(AB)(\alpha\beta)+(A\beta)(\alpha B)}$$

It is also called Yule's Coefficient of association and lies between –1 and +1.

If Q = 0, it means there is no association between the attributes A and B.

OR

A and B are independent

If Q > 0, it means there is +ve association between the attributes A and B.

If Q < 0, it means there is –ve association between the attributes A and B.

Example 6.3

For two attributes A and B, we have (AB) = 8, (A) = 18, (αβ) = 5 and N = 35. Calculate the coefficient of association.

Solution:

Putting the given values in the tabular form, we have

2 × 2 Contingency Table

Attribute	A	α	Total
B	(AB) = 8	(αB) = 12	(B) = 20
β	(Aβ) = 10	(αβ) = 5	(β) = 15
Total	(A) = 18	(α) = 17	N = 35

The table suggests

$$(\alpha) = N - (A) = 35 - 18 = 17$$

$$(A\beta) = (A) - (AB) = 18 - 8 = 10$$

$$(\beta) = (A\beta) + (\alpha\beta) = 10 + 5 = 15$$

CamScanner

$$(B) = N - (\beta) = 35 - 15 = 20$$

$$(\alpha\beta) = (B) - (AB) = 20 - 8 = 12$$

Therefore,

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

$$Q = \frac{8 \times 5 - 10 \times 12}{8 \times 5 + 10 \times 12} = \frac{40 - 120}{40 + 120} = -0.5$$

Example 6.4

From the following data, prepare the 2 × 2 table and using Yule's coefficient discuss whether there is association between literacy and unemployment:

Illiterate unemployed	:	220 persons
Literate employed	:	20 persons
Illiterate employed	:	180 persons
Total number of persons	:	500

Solution:

Let A denote the attribute of literacy and B that of unemployment. Hence, α will denote illiteracy and β employment.

$$(\alpha B) = 220 \text{ , } (A\beta) = 20 \text{ , } (\alpha\beta) = 180 \text{ and } N = 500$$

	A	α	Total
B	$(AB) = 80$	$(\alpha B) = 220$	$(B) = 300$
β	$(A\beta) = 20$	$(\alpha\beta) = 180$	$(\beta) = 200$
Total	$(A) = 100$	$(\alpha) = 400$	$N = 500$

Yule's formula:

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

$$Q = \frac{80 \times 180 - 220 \times 20}{80 \times 180 + 220 \times 20} = \frac{14400 - 4400}{14400 + 4400} = \frac{10000}{18800} = +0.532$$

This shows that literacy and unemployment are moderately associated.

Contingency Table and Test of Independence

A Contingency table is a two way table in which frequencies of various classes of two attributes are classified in rows and columns.

For example, college students may be classified according to class status and smoking habits, a sample of adults may be classified according to sex and the colour preference for a product, etc.

Let A is an attribute classified into r - rows and B is another attribute classified into c - columns. Then the following is a general $r \times c$ contingency table:

$A \backslash B$	B_1	B_2	......	B_j	......	B_c	Total
A_1	O_{11}	O_{12}	......	O_{1j}	......	O_{1c}	R_1
A_2	O_{21}	O_{22}	......	O_{2j}	......	O_{2c}	R_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_i	O_{i1}	O_{i2}	......	O_{ij}	......	O_{ic}	R_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_r	O_{r1}	O_{r2}	......	O_{rj}	......	O_{rc}	R_r
Total	C_1	C_2	......	C_j	......	C_c	G

Where O_{ij} = Observed frequency of ith row and jth column.

and $\sum_{i=1}^{r}\sum_{j=1}^{c} O_{ij} = \text{G} = \text{Grand Total}$

Note: The simplest form of a contingency table is the 2 × 2 table.

The data presented in a contingency table can be used to know (or to test) whether the attributes (or Factors) are independent. The testing procedure is as under:

1. Null Hypothesis: H_O:

The two attributes are independent (or there is no association between the two factors)

CamScanner

2. Alternative Hypothesis: H_1:

The two attributes are not independent (or there is an association between the two factors)

3. Level of Significance (α):

The commonly used levels are at $\alpha = 0.05, 0.01$

4. Test Statistic (Chi-Square formula):

$$\chi^2 = \sum\left[\frac{(o_{ij} - e_{ij})^2}{e_{ij}}\right]$$

where o_{ij} = observed frequency of ith row and jth column

$$e_{ij} = \frac{R_i \times C_j}{G} = \text{Expected/Estimated frequency of } i\text{th row and } j\text{th column}$$

and R_i = Total of ith row

C_j = Total of jth column

G = Grand Total

Value obtained by the above formula of χ^2 is called Calculated value of χ^2.

Hence $\chi^2_{Cal} = \sum\left[\frac{(o_{ij} - e_{ij})^2}{e_{ij}}\right]$

5. Tabulated Value:

Tabulated value of χ^2 depends upon α and the degrees of freedom. Where $d.f. = (r-1)(c-1)$

Hence $\chi^2_{tab} = \chi^2_{\alpha,\ (r-1)(c-1)}$

(Table of Chi-Square is enclosed at the end of this book)

6. Conclusion

Reject H_o if $\chi^2_{Cal} > \chi^2_{tab}$

CamScanner

Note: In testing of independence where the null hypothesis is that two criteria of classification are independent. Suppose that the sample size is large. Then, if the null hypothesis of independence is true, the random variable

$$\chi^2 = \sum\left[\frac{(o_{ij} - e_{ij})^2}{e_{ij}}\right]$$

has approximately a chi-square distribution. The number of degrees of freedom is

$$\text{d.f.} = (r - 1)(c - 1)$$

where r and c are the number of rows and columns in the contingency table respectively.

Example 6.5

1600 families were selected at random in a city to test the belief that high income families usually send their children to private schools and low income families often send their children to Government schools. The following results were obtained:

Income \ School	Private	Govt.	Total
Low	494	506	1000
High	162	438	600
Total	656	944	1600

Test whether income and type of schools are independent at $\alpha = 0.05$

Solution:

1. **H_o:** Income and type of schools are independent
2. **H_1:** Income and type of schools are not independent
3. $\alpha = 0.05$
4. **Test Statistic:** $\chi^2 = \sum\left[\frac{(o_{ij} - e_{ij})^2}{e_{ij}}\right]$

Since $o_{11} = 494$, $o_{12} = 506$, $o_{21} = 162$, $o_{22} = 438$

CamScanner

Then $e_{11} = \frac{R_1 \times C_1}{G} = \frac{(1000)(656)}{1600} = 410$

$$e_{12} = \frac{R_1 \times C_2}{G} = \frac{(1000)(944)}{1600} = 590$$

$$e_{21} = \frac{R_2 \times C_1}{G} = \frac{(600)(656)}{1600} = 246$$

$$e_{22} = \frac{R_2 \times C_2}{G} = \frac{(600)(944)}{1600} = 354$$

o_{ij}	e_{ij}	$\frac{(o_{ij} - e_{ij})^2}{e_{ij}}$
$o_{11} = 494$	$e_{11} = 410$	$\frac{(494-410)^2}{410}$
$o_{12} = 506$	$e_{12} = 590$	$\frac{(506-590)^2}{590}$
$o_{21} = 162$	$e_{21} = 246$	$\frac{(162-246)^2}{246}$
$o_{22} = 438$	$e_{22} = 354$	$\frac{(438-354)^2}{354}$
1600	1600	77.78

Then $\chi^2_{Cal} = 77.78$

5. Tabulated Value:

d.f. = (r − 1)(c − 1) = (2 − 1)(2 − 1) = 1

$\chi^2_{tab} = \chi^2_{0.05;1} = 3.84$

6. Conclusion:

Since $\chi^2_{Cal} > \chi^2_{tab}$, we reject H_o

Note: It is to be noted that, the natural application of the contingency table analysis is for cases in which each observation is measured by two qualitative variables (i.e. by two

CamScanner

attributes). However quantitative variables may also be used to classify the observations into rows and columns or both.

Example 6.6

The following table shows the relation between the number of accidents in 1 year and the age of the driver in a random sample of 500 drivers between 18 and 50. Test at $\alpha = 0.01$, the hypothesis that the number of accidents is independent of driver's age.

		Age of Driver			Total
		18 – 25	26 – 40	over 40	
No. of Accidents	0	75	115	110	300
	1	50	65	35	150
	2	25	20	5	50
Total		150	200	150	500

Solution:

1. H_0: There is no association between age of driver and number of accidents
2. H_1: There is an association
3. $\alpha = 0.01$
4. Test Statistic: $\chi^2 = \sum\left[\frac{(o_{ij} - e_{ij})^2}{e_{ij}}\right]$

where $e_{ij} = \frac{R_i C_j}{G}$ with d.f. $v = (r-1)(c-1)$

then $e_{11} = \frac{R_1 C_1}{G} = \frac{(300)(150)}{500} = 90$ $\qquad e_{12} = \frac{R_1 C_2}{G} = \frac{(300)(200)}{500} = 120$

$e_{13} = \frac{R_1 C_3}{G} = \frac{(300)(150)}{500} = 90$ $\qquad e_{21} = \frac{R_2 C_1}{G} = \frac{(150)(150)}{500} = 45$

$e_{22} = \frac{R_2 C_2}{G} = \frac{(150)(200)}{500} = 60$ $\qquad e_{23} = \frac{R_2 C_3}{G} = \frac{(150)(150)}{500} = 45$

CamScanner

$$e_{31} = \frac{R_3C_1}{G} = \frac{(50)(150)}{500} = 15 \qquad e_{32} = \frac{R_3C_2}{G} = \frac{(50)(200)}{500} = 20$$

$$e_{33} = \frac{R_3C_3}{G} = \frac{(50)(150)}{500} = 15$$

Then compute the value of χ^2 as

	o_{ij}	e_{ij}	$\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$
	75	90	2.5
	115	120	0.2
	110	90	4.4
	50	45	0.6
	65	60	0.4
	35	45	2.2
	25	15	6.7
	20	20	0
	5	15	6.7
Total	500	500	23.7

Then χ^2 - Calculated = 23.7

5. **Tabulated Value:** Tabulated value is computed as

   d.f. = $\nu = (3-1)(3-1) = 4$

   Hence $\chi^2_{tab} = \chi^2_{0.01,4} = 13.28$

6. **Conclusion:**

   Since $\chi^2_{cal} > \chi^2_{tab}$ we reject H_o

   where $\chi^2_{cal} = 23.7$ and $\chi^2_{tab} = 13.28$

Example 6.7

Consider the following 2 × 2 contingency table

Attribute	A_1	A_2
B_1	a	b
B_2	c	d

Calculate the value of χ^2

Solution:

Attribute	A_1	A_2	Total
B_1	a	b	$a+b$
B_2	c	d	$c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

Since $o_{11}=a, o_{12}=b, o_{21}=c, o_{22}=d$ and $G=a+b+c+d$

under the null hypothesis of independence, we calculate the expected frequencies as below:

$$e_{11}=\frac{R_1C_1}{G}=\frac{(a+b)(a+c)}{G} \qquad e_{12}=\frac{R_1C_2}{G}=\frac{(a+b)(b+d)}{G}$$

$$e_{21}=\frac{R_2C_1}{G}=\frac{(c+d)(a+c)}{G} \qquad e_{22}=\frac{R_2C_2}{G}=\frac{(c+d)(b+d)}{G}$$

Now the statistic is:

$$\chi^2=\sum\left[\frac{(o_{ij}-e_{ij})^2}{e_{ij}}\right]$$

and therefore

$$\chi^2=\frac{(o_{11}-e_{11})^2}{e_{11}}+\frac{(o_{12}-e_{12})^2}{e_{12}}+\frac{(o_{21}-e_{21})^2}{e_{21}}+\frac{(o_{22}-e_{22})^2}{e_{22}}$$

CamScanner

$$\chi^2 = \frac{\left[a - \frac{(a+b)(a+c)}{G}\right]^2}{\frac{(a+b)(a+c)}{G}} + \frac{\left[b - \frac{(a+b)(b+d)}{G}\right]^2}{\frac{(a+b)(b+d)}{G}} + \frac{\left[c - \frac{(c+d)(a+c)}{G}\right]^2}{\frac{(c+d)(a+c)}{G}} + \frac{\left[d - \frac{(c+d)(b+d)}{G}\right]^2}{\frac{(c+d)(b+d)}{G}}$$

$$\chi^2 = \frac{(ad-bc)^2}{G}\left[\frac{1}{(a+b)(a+c)} + \frac{1}{(a+b)(b+d)} + \frac{1}{(c+d)(a+c)} + \frac{1}{(c+d)(b+d)}\right]$$

$$\chi^2 = \frac{(ad-bc)^2}{G}\left[\frac{(a+b+c+d)^2}{(a+b)(c+d)(b+d)(a+c)}\right] = \frac{(ad-bc)^2(a+b+c+d)}{(a+b)(c+d)(b+d)(a+c)}$$

$$\chi^2 = \frac{(ad-bc)^2 G}{(a+b)(c+d)(b+d)(a+c)}$$

where $G = (a+b+c+d)$

Example 6.8

A random sample of 90 adults are classified according to sex and the number of hours they watch TV during a week.

	Male	Female
over 25 hours	15	29
under 25 hours	27	19

Using $\alpha = 0.01$ level of significance, test the hypothesis that the time spent watching TV is independent of whether one is a male or a female.

Solution:

1. H_o: There is no association between sex and time watching T. V.
2. H_1: There is an association between sex and time watching T. V.
3. $\alpha = 0.01$

CS CamScanner

4. **Test Statistic:** $\chi^2 = \sum\left[\frac{(o_{ij} - e_{ij})^2}{e_{ij}}\right]$

but for a 2 × 2 contingency table, the test statistic becomes:

$$\chi^2 = \frac{(ad - bc)^2 G}{(a+b)(c+d)(b+d)(a+c)}$$

where $G = (a + b + c + d)$ and

Since $a = 15, b = 29, c = 27, d = 19$ and $G = 90$

$$\chi^2 = \frac{(15 \times 19 - 29 \times 27)^2 (90)}{(44)(48)(48)(42)} = 5.47$$

Then χ^2 - calculated = 5.47

5. **Tabulated Value:**

d.f. = (2 – 1)(2 – 1) = 1

Hence $\chi^2_{tab} = \chi^2_{0.01,1} = 6.635$

6. **Conclusion:**

Since calculated value of χ^2 is less than the tabulated value of χ^2, therefore we accept H_o

Yates' Correction for Continuity

It is important to remember that the statistic on which we base our decision has a distribution that is only approximated by the chi-square distribution. The computed χ^2 values depend on the cell frequencies and consequently are discrete.

The Continuous chi-square distribution seems to approximate the discrete sampling distribution of χ^2 very well, provided that the number of degrees of freedom is greater than 1. In a 2 × 2 contingency table, where we have only 1 degree of freedom, a correction called Yates' Correction for Continuity is applied. The corrected formula then becomes

$$\chi^2_{(Corrected)} = \sum\left[\frac{\left(|o_{ij} - e_{ij}| - 0.5\right)^2}{e_{ij}}\right]$$

CamScanner

If the expected cell frequencies are large the corrected and uncorrected results are almost the same. When the expected frequencies are between 5 and 10, Yates' Correction should be applied.

Goodness of Fit Test

A goodness of fit test is used to know whether or not a given set of data follows a specific probability distribution. For this purpose we use following test statistic

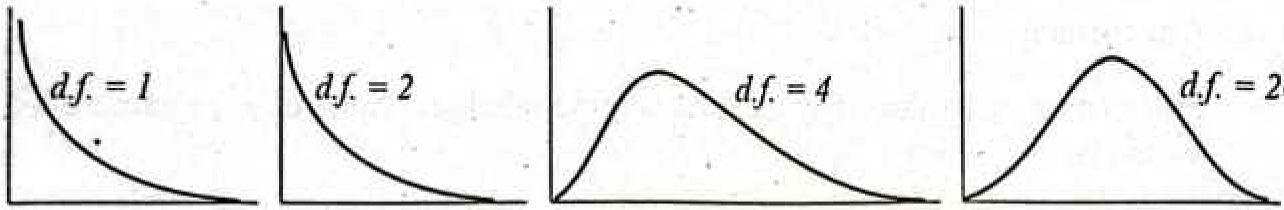
$$\chi^2 = \sum\left[\frac{(o_i - e_i)^2}{e_i}\right]$$

where o_i and e_i represent the observed and expected frequencies respectively of a set of k categories or classes or cells.

Degrees of Freedom in Goodness of Fit Test

The number of degrees of freedom in chi-square goodness of fit test is equal to the number of categories minus the number of quantities obtained from the observed data, which are used in the calculations of the expected frequencies.

The shapes of χ^2 - distribution for various degrees of freedom are given below:



As the degrees of freedom increases the shape become symmetrical.

Note: It is important to note that the χ^2 approximation described here should not be used unless each of the expected frequency is at least equal to 5. If the expected frequency of a category is not at least 5, this category can be combined with one or more other categories until the condition is satisfied.

The Testing procedure involves the following steps:

1. H_0: Fit is Good or (Sample data obtained from specified distribution)
2. H_1: Fit is not Good or (Sample data not obtained from specified distribution)
3. Choose a level of significance equal to α

4. Test Statistic: $\chi^2 = \sum\left[\frac{(o_i - e_i)^2}{e_i}\right]$ with $d.f. = \nu = k - 1$

where k is the No. of cells or categories or classes.

5. Tabulated Value: $\chi^2_{tab} = \chi^2_{\alpha, k-1}$

6. Conclusion: Reject H_o, if the calculated value of χ^2 is greater than the tabulated value of χ^2 (i.e. if $\chi^2_{cal} > \chi^2_{tab}$), otherwise accept H_o

Example 6.9

A die is tossed 180 times with the following results

Dots on die (x)	1	2	3	4	5	6
Frequency (o)	28	36	36	30	27	23

Is this a fair die? Use a 0.01 level of significance.

Karachi Board 2009

Solution:

It is important to note that when we hypothesize that the die is honest/balance or fair, which is equivalent to testing the hypothesis that the distribution of outcomes is uniform.

The testing procedure is given below:

1. **H_o:** The die is fair (or Distribution is Uniform)

OR

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

2. **H_1:** The die is not fair (or Distribution is not Uniform)

3. $\alpha = 0.01$

4. **Test Statistic:** $\chi^2 = \sum\left[\frac{(o_i - e_i)^2}{e_i}\right]$

CamScanner

Computations are as under:

x	Prob.	o	$e = 180 \times$ Prob.	$\frac{(o_i - e_i)^2}{e_i}$
1	$\frac{1}{6}$	28	$180 \times \frac{1}{6} = 30$	0.13
2	$\frac{1}{6}$	36	$180 \times \frac{1}{6} = 30$	1.20
3	$\frac{1}{6}$	36	$180 \times \frac{1}{6} = 30$	1.20
4	$\frac{1}{6}$	30	$180 \times \frac{1}{6} = 30$	0.00
5	$\frac{1}{6}$	27	$180 \times \frac{1}{6} = 30$	0.30
6	$\frac{1}{6}$	23	$180 \times \frac{1}{6} = 30$	1.63
Total	1	180	180	4.46

Hence χ^2 - calculated = 4.46

Now the only quantity provided by the observed data, in computing expected frequencies is the total frequency. Hence d.f. = 6 – 1 = 5

5. Tabulated Value:

$d.f. = k - 1 = 6 - 1 = 5$

Hence, $\chi^2_{tab} = \chi^2_{0.01,5} = 15.09$

6. Conclusion:

Since $\chi^2_{cal} < \chi^2_{tab}$, we accept H_0 where $\chi^2_{cal} = 4.46$ and $\chi^2_{tab} = 15.09$

Example 6.10

A fire department records in a city show that the last 350 fires were distributed according to day of the week as follows:

CamScanner

Monday	= 35
Tuesday	= 35
Wednesday	= 45
Thursday	= 40
Friday	= 65
Saturday	= 75
Sunday	= 55
Total	= 350

Test the hypothesis that fires are uniformly distributed throughout out the week. That is to test ; $H_o : p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = \frac{1}{7}$ (Use $\alpha = 0.01$)

Solution:

1. **H_o:** Fires are uniformly distributed throughout the week.
2. **H_1:** Fires are not uniformly distributed throughout the week.
3. $\alpha = 0.01$
4. **Test Statistic:** $\chi^2 = \sum\left[\frac{(o-e)^2}{e}\right]$

Therefore observed and expected frequencies with the computation of χ^2 are shown in the following table:

CamScanner

o	Prob.	$e = 350 \times$ Prob.	$\frac{(o-e)^2}{e}$
35	1/7	50	4.5
35	1/7	50	4.5
45	1/7	50	0.5
40	1/7	50	2.0
65	1/7	50	4.5
75	1/7	50	12.5
55	1/7	50	0.5
N = 350	1	350	29.0

Then χ^2 - calculated = 29.0

Now, the only quantity provided by the observed data in computing expected frequencies is the total frequency. Hence $d.f. = 7 - 1 = 6$

5. Tabulated Value:

$d.f. = \nu = 7 - 1 = 6$

Hence $\chi^2_{tab} = \chi^2_{0.01,6} = 16.81$

6. Conclusion:

Since $\chi^2_{cal} > \chi^2_{tab}$, we reject H_o, where $\chi^2_{cal} = 29.0$ and $\chi^2_{tab} = 16.81$

CS CamScanner

EXERCISE – 6

6.1 Explain Attribute with the help of an example.

Karachi Board 2004

6.2 Define a Contingency Table.

Karachi Board 2003

6.3 (a) Discuss the role of Contingency tables in Statistics.

(b) A random sample of size 250, adults are classified according to sex and the number of hours they watch T. V. during week.

	Male	Female
Over 25 Hours	70	55
Under 25 Hours	30	95

Is time spent related to sex. Use $\chi^2_{0.05,1} = 3.841$

6.4 What do you mean by Test of Association.

Karachi Board 1998

6.5 From a human population 1000 individuals are selected at random and classified according to the following table on the basis of sex and colour-blindness as follows

	Male	Female
Normal	452	494
Colour Blind	38	16

Test if there is any association between sex and colour blindness at $\alpha = 0.01$

Karachi Board 1997

CamScanner

6.6 (a) Define Contingency Table and write its importance.

(b) The following table shows the number of recruits taking

(i) a preliminary and (ii) a final test in car driving.

Test whether there is any association between the results of the preliminary and those of the final tests. Use $\alpha = 0.05$

Categories		Preliminary	
		Passing	Failing
Final	Passing	605	135
	Failing	195	65

Karachi Board 1995

6.7 (a) What is Contingency table? Write down its uses.

(b) A certain drug is claimed to be effective in curing common cold. In an experiment on 164 people with common cold half of them were given the drug and half of them were given sugar pills. The patients' reactions to the treatment are recorded in the following table. Test the hypothesis that the drug is no better than the sugar pills for curing the common cold. Use $\alpha = 0.05$

Category	Helped	Harmed	No Effect
Drug	52	10	20
Sugar	44	12	26

Karachi Board 1996

6.8 (a) (i) Define χ^2 Distribution.

(ii) Explain briefly the meaning of Association of Attributes.

(b) Calculate χ^2 from the following table:

	Not attacked	Attacked	Total
Not Inoculated	15	301	316
Inoculated	07	438	445
Total	22	739	761

Karachi Board 1994

6.9 The sales representatives of a company were classified by the number of years of experience in sales before joining the company and their performance during the first year with the company. The data are given below

Performance	Prior Experience in years		
	over 5	2 – 5	below 2
Poor	24	42	67
Satisfactory	70	41	92
Exceptional	36	27	56

At 0.05 level of significance is the sale experience a significant factor in the performance of the sales representative.

Karachi Board 2004

6.10 A random sample of 200 families was classified according to educational level and size of family.

Level of Education	Number of children in family		
	0 – 1	2 – 4	5 and above
Primary	8	40	62
Secondary	12	15	18
College	15	12	18

Test the hypothesis at 5% level of significance that the size of the family is independent of the level of education.

Karachi Board 2004

CamScanner

6.11 The marks of students of an institution obtained in Statistics and Mathematics are given blow

	Ist year	IInd year	Total
Mathematics	15	6	21
Statistics	9	20	29
Total	24	26	50

Test the hypothesis that the subjects and levels (years) are independent at

(i) 5% level of significance (ii) 1% level of significance

Karachi Board 2003

6.12 The following data shows the relation between the number of accidents in 1 year and the age of the driver in a random sample of 500 drivers between 18 and 50 years. Test at $\alpha = 0.01$, the hypothesis that the number of accidents is independent of driver's age.

No. of Accidents	Age of Driver		
	18 – 25	26 – 40	over 40
0	75	115	110
1	50	65	35
2	25	20	5

Karachi Board 2003

6.13 In an experiment to study the dependance of hypertension on smoking habits, the following data were taken from 180 individuals:

	Non-Smokers	Moderate Smokers	Heavy Smokers
Hypertension	21	36	30
No Hypertension	48	26	19

CamScanner

Test the hypothesis that presence or absence of hypertension is independent of smoking habits. Use $\chi^2_{0.05,2} = 5.991$

Karachi Board 2003, 2007 (Supp.)

6.14 (a) What is the role of Contingency Table in Statistics.

(b) The level of job satisfaction and job categories of 800 employees of a large business organization are summarized in the following data

		Job Category			
		I	II	III	IV
Level of Satisfaction	High	46	54	60	41
	Medium	103	81	82	79
	Low	62	52	67	73

Test at 0.05 level of significance that the level of satisfaction is independent of the job category. ($\chi^2_{0.05,6} = 12.592$)

Karachi Board 2002

6.15 The following data relate to two types of treatment, new and conventional, tried on 200 patients

	No. of Patients		Total
	Benefit	No Benefit	
New	60	20	80
Conventional	70	50	120
Total	130	70	200

On the basis of this data, can it be concluded that the new treatment is better than the conventional treatment at 0.05 level of significance.

Karachi Board 2002

6.16 A two-way table containing sample frequencies is given below

CS CamScanner

Preference	Marital Status	
	Single	Married
Brand A	20	10
Brand B	20	50

Test the hypothesis at 5% level of significance that marital status and brand preference are independent. ($\chi^2_{0.01,1} = 3.841$)

Karachi Board 2001

6.17 From a locality 100 persons are randomly selected and are asked about their Education levels. The results are given as under

Sex	Middle School	High School	College
Male	10	15	25
Female	25	10	15

Test the Association between Education and Sex by using Chi-Square Test ($\chi^2_{0.05,2} = 5.99$)

Karachi Board 2001, 2008, 2011 (Supp.)

6.18 Test the hypothesis that there is no Association between total Income and television Ownership using the following data

Owner	A	B	C
Colour T. V.	56	51	93
Black & White T. V.	118	207	375
None	26	42	32

Use $\chi^2 = 9.488$, $\alpha = 5\%$

Karachi Board 1999

6.19 (a) Define Contingency Table.

(b) The following data show the distribution of 500 students according to general ability and mathematical ability:

CS CamScanner

Mathematical Ability	General Ability		
	Good	Fair	Poor
Good	50	29	11
Fair	45	57	78
Poor	41	91	98

Do the data indicate that the general ability and mathematical ability are independent of each other? Use $\alpha = 0.01$

Karachi Board 1992

6.20 (a) Define Contingency table.

(b) Given the following contingency table for hair colour and eye colour:

Eye Colour	Hair Colour		
	Black	Fair	Brown
Brown	10	22	32
Black	15	28	29
Grey	25	20	19

Find out if there is any relationship between hair colour and eye colour, using 5% level of significance.

Karachi Board 1991, 2006

6.21 Define Attributes and give an example.

In a random sample 40 men and 60 women were asked whether or not they liked a certain brand of toothpaste. Their responses were shown in the table:

	Yes	No	Total
Men	15	25	40
Women	22	38	60
Total	37	63	100

CamScanner

Test whether the likings of both the sexes are related. The tabulated value of statistic at 1 d.f. = 3.841

Karachi Board 1988

6.22 (a) Write the importance of χ^2 - test in Statistics.

(b) Define Contingency table and association. Explain the use of the table for testing the Association of Attributes.

Karachi Board 1985

6.23 A random sample of 200 married men classified according to education and their number of children was as under:

Education	Number of children		
	0 – 1	2 – 3	over 3
Elementary	14	37	32
Secondary	19	42	17
College	12	17	10

Test the hypothesis at 5% level of significance that the size of a family is independent of the level of education attained by their father.

Karachi Board 1988

6.24 The following table shows the relation between the performance of students in Mathematics and Physics. Test the hypothesis that the performance in Physics is independent of performance in Mathematics at 0.05 level of significance.

Grades		Mathematics		
		A	B	C
Physics	A	56	71	42
	B	47	163	38
	C	14	42	85

Karachi Board 1984, 2010

CS CamScanner

6.25 The following data given the distribution of 200 school children according to physical defects (P_1, P_2, P_3) and speech defects (S_1, S_2, S_3)

Speech Defects	Physical Defects		
	P_1	P_2	P_3
S_1	24	22	24
S_2	19	16	20
S_3	21	28	26

Do the data indicate that the physical defects and speech defects are independent at $\alpha = 0.05$

Karachi Board 1992

6.26 (a) Discuss a contingency table. Comment on its widespread use in research work.

(b) A certain company wishes to determine whether absenteeism is related to age. A random sample of 200 employees is selected and classified according to age and cause of absenteeism from their offices as follow:

Cause	Age		
	Under 30	30 – 50	Over 50
Illness	40	28	52
Others	20	36	24

Is age related to absenteeism, at $\alpha = 0.01$

Karachi Board 1999

6.27 (a) Define Contingency Table.

(b) Given the following data:

Sex	Primary Education	Secondary Education	Higher Education
Male	220	270	110
Female	120	45	35

Do the data suggest association between Education and Sex? Use 5% level of significance.

Karachi Board 1999

6.28 The following table shows the likings of three colours — pink, white and blue — in a sample of 200 males and females:

Colour	Male	Female
Pink	20	40
White	40	20
Blue	60	20

Test (at $\alpha = 0.05$) whether there is any relation between sex and colour.

Karachi Board 1989

6.29 **(a)** What do you mean by the Test of Association?

(b) A random sample of 200 college students was classified according to class status and taking tea with breakfast. Do the following data indicate association between taking tea and class status?

Where $\chi^2_{(0.05,1)} = 3.841$

	Inter Science	B. Sc.	Total
Not taking tea	100	20	120
Taking tea	30	50	80
Total	130	70	200

Karachi Board 1988

6.30 **(a)** What do you mean by the Test of Association?

(b) A random sample of 200 students was classified according to class status and tea-taking with breakfast. Does the following table indicate the association between tea-taking and class status:

CS CamScanner

	Secondary	Inter	Graduation
No Tea-taking	60	50	10
Tea-taking	20	20	40

Examine it at 5% level of significance.

Karachi Board 1989

6.31 (a) Define a Contingency table.

(b) The table given below shows the data obtained during an epidemic of cholera

	Attacked	Not Attacked
Inoculated	31	469
Not Inoculated	185	1315

Test the effectiveness of inoculation in preventing the attack of cholera at 5% level of significance.

Karachi Board 1990

6.32 (a) What is χ^2 distribution? Explain the test of association.

(b) Is there a significant association between A and B from the following (2×2) table.

Attributes	A_1	A_2	Total
B_1	64	26	90
B_2	21	49	70
Total	85	75	160

Karachi Board 1988

6.33 (a) Explain Contingency Table and its role by example.

(b) For a 2×2 Contingency, derive a formula of "Chi-Square".

Karachi Board 2000

CamScanner

6.34 **A random sample of 250 married men was classified according to education and the number of children**

Education	Number of children		
	0 – 1	2 – 3	over 3
Elementary	29	19	20
Secondary	91	18	27
College	25	10	11

Test the hypothesis at the 0.05 level of significance that the size of a family is independent of the level of education attained by the father.

Karachi Board 2003

6.35 **(a)** Write a brief note on the use of Contingency Table.

(b) Show that in a (2×2) contingency table where the frequencies are a, b, c, and d, χ^2 calculated from the independent frequencies is

$$\chi^2 = \frac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(b+d)(a+c)}$$

Karachi Board 1993, 2011

6.36 **(a)** The grades in Statistics course for a particular semester were as follows

Grade	A	B	C	D	E
Freq.	14	18	32	20	16

Test the hypothesis at 0.05 level of significance, that the distribution of grades is uniform.

(b) Derive an expression for chi-square in a 2×2 Contingency table.

Karachi Board 2001

6.37 Write short notes on the following

(i) Contingency Table. (ii) Test of Independence/Test of Association

(iii) Attribute

CamScanner

PRACTICAL EXAMINATION PAPERS
OF
KARACHI BOARD

CamScanner

BOARD OF INTERMEDIATE EDUCATION KARACHI

H. S. C. PART II ANNUAL EXAMINATION – 2003

Time : 2 Hours Max. Marks: 25
Viva : 5 Marks
Journal : 2 Marks

SET – A

STATISTICS PRACTICAL – II

Instructions: (i) Attempt any two questions.
(ii) All questions carry equal marks.
(iii) Computations are to be presented in tabular form.

Q. 1 Finite populations are given as
9, 2, 12, 14, 5, 20, 22 and 10

(a) Draw all possible samples, each of size 3, using without replacement method.

(b) Find out mean and variance of population and verify that $E(\overline{X}) = \mu$

Q. 2 Following information are given:

X = Marks in Statistics	12	33	39	48	57	68	75	80	85	95
Y = Marks in Economics	5	38	45	45	50	33	30	35	40	30

(a) Find coefficient of correlation between X and Y

(b) Find regression equation of Y on X

(c) Estimate Y if X = 50

Q. 3 Fit a binomial probability distribution to

x	0	1	2	3	4	5	6	7	8
f	11	27	40	49	38	20	9	5	1

Plot observed and fitted frequencies on same graph.

CamScanner

ANNUAL EXAMINATIONS – 2003

SET – B

Q. 1 A population is given as:

9 , 2 , 14 , 5 , 20 , 22 and 18

(a) Draw all possible samples, each of size 2, using with replacement method

(b) Find mean and variance of population and find $E(\overline{X}) = \mu$

Q. 2 Following information are given

X	22	43	49	58	67	78	80	85	95	100
Y	15	48	55	55	60	63	70	75	80	92

(a) Find correlation between X and Y.

(b) Estimate X if Y = 70

Q. 3 In a certain city 1,000 were selected at random and questioned about their smoking and drinking habit.
The results of the survey were:

Drink	Never	Smoke → occasionally	Moderately	Heavily
Never	85	23	56	36
Occasionally	153	44	128	75
Moderately	128	26	101	45
Heavily	34	7	15	44

What condition can be drawn concerning the smoking and drinking habits of people in the city use $\alpha = 0.05$

CamScanner

ANNUAL EXAMINATIONS – 2003

SET – C

Q. 1 Draw all possible samples of size 3 from given population 3, 14, 5, 10, 19, 21, 25 and 30. Find means of all possible samples and show that sample mean is an unbiased estimate of the population mean.

$$E(\overline{X}) = \mu$$

Q. 2 A criminologist conducted a survey to determine whether the incidence of certain types of crime varied from one part of a large city to another. The particular crimes of interest were assault, burglary, larceny and homicide. The following table shows the numbers of crimes committed in four areas of the city during the past year.

TYPE OF CRIME				
District	Assault	Burglary	Larceny	Homicide
1	162	115	451	18
2	310	196	996	25
3	258	193	458	10
4	280	175	390	19

Can we conclude from these data at the 0.01 level of significance that the occurrence of these types of crimes is dependent upon the city district?

Q. 3 A mathematics placement test is given to all entering freshmen at a small college. A student who receive a grade below 35 is denied admission to the regular mathematics course and placed in a remedial class. The placement test scores and the final grades for 20 students who took the regular course were recorded as follows:

Placement Test	50	35	35	40	55	65	35	60	90	35
Course Grade	53	41	61	56	68	36	11	70	79	59
	90	80	60	60	60	40	55	60	65	50
	54	91	48	71	71	47	53	68	57	79

(e) Find the equation of regression line to predict course grades from placement test scores.

(f) Find coefficient of correlation and comment.

CS CamScanner

ANNUAL EXAMINATIONS – 2003

SET – D

Q. 1 Given the following data

x	1	2	3	4	5	6	7	8	9
y	8	11	14	17	20	23	26	30	35

(a) Find out the product of Regression coefficients and coefficient of correlation.

(b) Plot the regression line of y on x on graph paper.

Q. 2 A random sample of 200 students is selected from each of four undergraduates classes. Each is asked to express his preference as a Republican, Democrat independent voter. The results are:

Class	Preferences →		Independent
	Republican	Democrat	
Freshman	51	110	39
Scphoman	74	106	20
Junior	25	124	51
Senior	90	60	50

Are the classes alike in their preference? Use $\alpha = 0.05$

Q. 3 (a) Given for a Normal Distribution

$$\mu = 30 \ , \ \sigma = 5$$

Find $P(X \geq 28), P(X \leq 32)$

(b) Fit binomial probability distribution.

x	0	1	2	3	4	5
f	2	9	16	10	3	1

CS CamScanner

ANNUAL EXAMINATIONS – 2003

SET – E

Q. 1 (a) A Normal probability Distribution has Mean = 69 and S. D. = 3. Find $P(66 \le X \le 68)$ and $P(X \le 70)$

(b) Fit Binomial Prob. Distribution to

x	0	1	2	3	4
f	5	17	30	10	8

Q. 2 A population is given as (9, 11, 13, 15, 17). Select all possible samples, each of size 2, using with replacement method. Also find out $E(\overline{X})$

Q. 3 Given values as following:

X	10	15	20	25	30	35	40	45
Y	100	150	200	250	300	350	400	450

(a) Find coefficient of correlation between X and Y.

(b) Find out both regression equations.

(c) Plot the graph of given X and Y.

CS CamScanner

ANNUAL EXAMINATIONS – 2003

SET – H

Q. 1 (a) Fit Binomial Probability Distribution to:

X	0	1	2	3	4	5
f	15	35	60	40	30	20

(b) If the Probability that a New Born Baby will be a Girl is 0.55 Use Normal Distribution to find probability that between 50 and 60 will be girls out of 100 next birth.

Q. 2 A manufacturing company has 9 years data of its X = Advertisement cost and Y = Sales (in Rs. 1000)

X	45	52	55	59	62	64	68	75	78	80
Y	69	75	80	75	85	84	75	87	89	95

(a) Find out coefficient of correlation between X and Y.

(b) What would be sale if Adv. cost is 90.

Q. 3 A population Consist 9, 11, 13, 15, 17, 19, 21 and 23

(a) Select all possible samples, each of size 3 using without replacement method.

(b) Verify that $E(\overline{X}) = \mu$

CS CamScanner

ANNUAL EXAMINATIONS – 2003

SET – I

Q. 1 A population is given as

5, 10, 8 and 13

(a) Draw all possible samples, each of size 2, using without replacement method.

(b) Find $\overline{X}$ and G. M. of each sample.

(c) Find correlation coefficient between $\overline{X}$ and G. M.

Q. 2 In the accompanying contingency table, 'X' represents a rating given to each of a group of university freshmen on the basis of high school reports and 'Y' represents the final standing in degree examinations for the same group. Discuss the associate between these two attributes.

Y \ X	Fair	Good	Excellent
3rd class	173	167	110
2nd class	164	184	115
1st class	105	124	128

Q. 3 Given information are:

X = Age of Father	20	25	35	40	50	65	70	80
Y = Age of Son	02	02	09	17	28	48	50	60

(a) Plot both regression lines on same graph paper and comment.

(b) Estimate age of son, if the age of father is 45 year.

(c) Find Coefficient of Correlation.

CS CamScanner

ANNUAL EXAMINATIONS – 2004

SET – J

Q. 1 Two examiners gave the following marks to ten students. Calculate the Rank correlation coefficient between the two examiners and comments.

Students	A	B	C	D	E	F	G	H	I	J
Examiner – I	80	50	70	58	52	58	75	70	68	72
Examiner – II	69	46	59	59	48	56	60	56	59	70

Q. 2 A Normal distribution has a mean of 50 and standard deviation of 5. What proportion of the population values lie in these intervals:

(a) from 40 to 60

(b) from 35 to 60

(c) from more than 60

(d) from less than 60

(e) exactly at 60

Q. 3 Test whether the number of defective items produced by two machines is independent of the machine on which they are made. Use $\alpha = 1\%$

	Machine output Defective articles	Effective articles
Machine A	25	375
Machine B	42	558

CamScanner

ANNUAL EXAMINATION – 2004

SET – K

Q. 1 A security analyst wants to be convinced that the efficiency of Capital utilization expressed by the annual turnover of inventory, actually does have an effect on a manufactures earning. A sample of 7 firms is chosen and the followed results are obtained.

Company	A	B	C	D	E	F	G
Inventory Turnover	3	4	5	6	7	8	9.
Earning a %	10	8	12	15	13	16	19

(a) Calculate coefficient of correlation (r)

(b) Show that r = s geometric mean of two regression coefficients

(c) Interpret; r.

Q.2 The incidence of defective items in 200 samples of 6 is Shown in the following Table.

No of defcelines per sample	0	1	2	3	4	5	6
No of samples	36	70	61	25	7	1	0

(a) Fit an appropriate probability distribution.

(b) Plot probability curve on graph paper.

(c) Comment the shape of the curve.

Q. 3 (a) Draw all possible samples of size 3 without replacement using the following population 0, 2, 4, 6, 8, 10, 12.

(b) Verify that $E(\overline{X}) = \mu$ And $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$

ANNUAL EXAMINATIONS – 2004

SET – L

Q. 1 (a) A security analyst wants to find the relation between annual turnover of inventory and the manufactures earning. A sample of five firms is chosen and the following results are obtained.

Company	A	B	C	D	E
Inventory Turnover	3	4	5	6	7
Earning as a Percentage	10	8	12	15	13

(b) If the manufactures earning as a percentage of sale is 18 than how much is the inventory turnover.

Q. 2 Find the area under the Standard Normal Curve in each of the following cases:

(i) $P(0 \le Z \le 1.25)$

(ii) $P(-2.0 \le Z \le 0)$

(iii) $P(1.82 \le Z \le 2.50)$

(iv) $P(-0.80 \le Z \le 1.20)$

(v) $P(Z \ge 1.25)$

(vi) $P(Z \le 1.25)$

(vii) $P(Z = 2.0)$

Q. 3 Test the Hypothesis at 5% level of significance that the presence or absence of hypertension is independent of smoking habits.

	Non-Smokers	Moderate Smokers	Heavy Smokers
Hypertension	25	40	30
No-hypertension	40	23	17

CS CamScanner

ANNUAL EXAMINATIONS – 2004

SET – M

Q. 1 Among diabetics, the fasting blood glucose level X may be assumed to be approximately normally distributed with mean 106 mg and standard deviation of 8 mg.
Find:

(a) $P(X \le 120)$

(b) $P(106 \le X \le 110)$

(c) $P(X \ge 112)$.

(d) $P(X = 106)$

Q. 2 Test whether there is an association between an Eye colour and Hair colour, at 5% level of significance.

Eye Colour	Hair Colour	
	Light	Dark
Blue	32	12
Brown	14	22
Other	6	9

Q. 3 Calculate the strength of relationship between the weights and heights of the men.

X-in inches	60	60	60	62	62	64	64	70	70	70
Y-in lbs	110	135	120	120	140	130	135	150	145	170

CamScanner

ANNUAL EXAMINATIONS – 2004

SET – N

Q. 1 Is there an association between the channel watched most and the region, test the hypothesis at 5% level of significance.

Region	Channel			
	1	2	3	4
North	29	16	42	23
Central	6	11	26	7
South	15	3	12	10

Q. 2(a) Draw all possible samples of size 3 without replacement from the given population 0, 1, 2, 3, 4, 5, 6.

(b) Verify:

(i) $E(\bar{x}) = \mu$

(ii) $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$

Q. 3 Find Area under the Normal curve:

(i) $P(Z \leq 1.56)$

(ii) $P(Z \geq -1.29)$

(iii) $P(-1.29 \leq Z \leq 1.56)$

(iv) $P(Z = 1.0)$

CS CamScanner

ANNUAL EXAMINATIONS – 2004

SET – O

Q. 1 Does the data indicate that consumer rating and product are independent? Use 0.01 level of significance.

Product	Rating		
	Superior	Average	Inferior
Radio	30	40	5
Stereo	10	20	5
B/N (T. V.)	30	20	10
Colour (T. V.)	10	40	10

Q. 2 The mean height at air force base is 69 inches with a standard deviation of 2.5 inches. If the air force base has 300 men, then how many men would you expect to have heights.

(a) More than 72 inches

(b) Less than 72 inches

(c) Equal to 72 inches

Q. 3 (a) Draw all possible samples of size 3 from the population 1, 3, 5, 7, 9, 11, 13 without replacement.

(b) Verify:

(i) $E(\overline{x}) = \mu$

(ii) $\sigma_{\overline{x}}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$

CS CamScanner

ANNUAL EXAMINATIONS – 2004

SET – P

Q. 1 (a) Draw all possible sample of size 3 from the population of 2, 4, 6, 8, 10, 12, 14 without replacement.

(b) Verify that

(i) $E(\bar{x}) = \mu$

(ii) $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$

Q. 2 Fit a Binomial distribution using the following data.

X	0	1	2	3	4
f	30	62	46	10	2

(a) Plot observed and expected frequencies on the same graph paper.

(b) Comment the shape of the curve.

Q. 3 Test is there a relationship between returns selected for audit and preparer. Use 0.05 level of significance.

Preparer	**Audit**	
	Yes	No
Self	32	118
Others	8	42

CS CamScanner

MULTIPLE CHOICE QUESTIONS

CamScanner

MULTIPLE CHOICE QUESTIONS (MCQs)

SET – A

Choose the correct answer for each from the given options:

(i) If n coins are tossed, the number of possible outcomes in the sample space is:
- n^2
- 2n
- n
- 2^n

(ii) If two cards are drawn together at random from a standard deck of cards, the number of possible outcomes is:
- 52 × 52
- 52!
- 1326
- 52 × 51

(iii) In rolling a fair die, the probability of an even number is:
- $\frac{1}{2}$
- $\frac{1}{3}$
- $\frac{1}{4}$
- 1

(iv) The probability of drawing a jack from a standard deck of cards is:
- $\frac{1}{52}$
- $\frac{4}{13}$
- $\frac{1}{4}$
- $\frac{1}{13}$

(v) If $P(A|B)=P(A)$ and $P(B|A)=P(B)$, then the events A and B are said to be:
- Independent
- Dependent
- Mutually Exclusive
- Equally Likely

(vi) For two mutually exclusive events A and B, P(A) = 0.2 and P(B) = 0.4. $P(A \cup B)$ is:
- 0.8
- 0.2
- 0.6
- 0.52

(vii) The parameters of the binomial distribution are:
- n, p
- a, b
- c, d
- np, nqp

(viii) The Mean of the binomial distribution is:
- np
- $\sqrt{np(1-p)}$
- $np(1-p)$
- npq

(ix) If X is a binomial random variable with n = 10 and p = 0.4, the mean is:
- 4
- 6
- 8
- 10

(x) The mean and variance of standard normal variable 'Z' are:
- 0 and 1
- μ and σ^2
- 1 and 0
- −1 and +1

(xi) The complete enumeration of population is called:
- census
- sampling
- survey
- None of these

(xii) Given N = 5 and n = 3 in simple random sampling with replacement, the possible samples are:
- 10
- 125
- 60
- 15

(xiii) A population has $\mu = 10$, $\sigma = 8$ and sample size n = 4. The standard error of $\bar{x}$ will be:
- 40
- 4
- 8
- 2

(xiv) When the two variables tend to move in the opposite directions the coefficient of correlation will be:
- positive
- negative
- zero
- none of these

(xv) The limit of correlation coefficient is always:
- 0 to 1
- −1 to 1
- −1 to 0
- 1 to 100

(xvi) The value of chi-square variable is always:
- positive
- negative
- zero
- none of these

(xvii) For r × c contingency table, the number of degrees of freedom is equal to:
- (r − 1) + (c − 1)
- (r − 1)(c − 1)
- 1
- (r × c) − 1

CamScanner

SET – B

Choose the correct answer for each from the given options:

(i) The height of a person is a:
- continuous random variable
- discrete random variable
- qualitative random variable
- none of these

(ii) The coefficient of correlation is unaffected by:
- scale
- origin
- origin and scale
- none of these

(iii) Goodness of a fit test is:
- χ^2 - test
- z - test
- t - test
- none of these

(iv) $P(A)+P(A')$ is equal to:
- zero
- 2
- 1
- none of these

(v) If P(A) = 0.5, P(B) = 0.6 and $P(A \cap B) = 0.3$, then $P(A \cup B)$ is equal to:
- 0.8
- 0.1
- 0.3
- none of these

(vi) The sum of probabilities of a sample space is:
- zero
- –1
- 0.5
- 1

(vii) Mean, median and mode of a normal distribution are always:
- unequal
- equal
- zero
- positive

(viii) The mean of a binomial distribution is:
- npq
- pq
- np^2
- none of these

(ix) If three coins are tossed together, the total number of elements in the sample space will be:
- 6
- 8
- 9
- 27

(x) The complete enumeration of population is called:
- census
- survey
- sampling
- stratified sampling

(xi) How many ways are there if a pair of dice is rolled together?
- 12
- 26
- 30
- 36

(xii) If a quantity is calculated from sample, it is called:
- parameter
- statistic
- discrete
- primary

(xiii) The dependence of one variable over the other variable is called:
- variation
- correlation
- regression
- association

(xiv) All possible samples of size 2 taken from a population of size 5 with replacement will be:
- 1
- 12
- 15
- 25

(xv) Any characteristic of population is called:
- parameter
- statistics
- statistic
- none of these

(xvi) The limit of rank correlation is from:
- 0 to 1
- –1 to 1
- 0.5 to 1
- 0 to 0.5

(xvii) The square root of the product of two regression coefficients is equal to:
- variation
- correlation
- regression
- association

SET – C

Choose the correct answer for each from the given options:

(i) **A selection of distinct objects without regarding order is called:**
- Permutation
- Combination
- n-factorial
- Tree diagram

(ii) **A device used to list all possibilities of outcomes in a systematic way is called:**
- Venn diagram
- Pie diagram
- Tree diagram
- Bar diagram

(iii) **The probability that a second event will occur given that first has already occurred is called:**
- Joint Probability
- Conditional Probability
- Total Probability
- none of these

(iv) **The events occurring together without affecting each other are called:**
- Mutually Exclusive
- Dependent
- Independent
- Equally likely

(v) **The probability of rolling two dice and getting same digits on them is:**
- $\frac{1}{6}$
- $\frac{1}{12}$
- $\frac{1}{2}$
- $\frac{1}{36}$

(vi) If event A is certain to occur, then P(A) is equal to:
- ½
- 0
- 1
- ¼

(vii) A list of all possible values of a random variable along with their respective probabilities is called:
- random experiment
- probability distribution
- frequency distribution
- skewed distribution

(viii) If $p < q$ in a binomial distribution, then the distribution will be:
- Symmetrical
- Positively skewed
- Negatively skewed
- none of these

(ix) The mean and variance of $(q+p)^3$ are:
- 3p and 3pq
- 3p and 3q
- 3p and $\sqrt{3pq}$
- 3q and pq

(x) If X is a binomial random variable with n = 100 and p = 0.4, then $E\left(\frac{X-10}{6}\right)$ is:
- 5
- 4
- –6
- 1

(xi) The normal distribution is an example of:
- discrete distribution
- skewed distribution
- binomial distribution
- continuous distribution

(xii) For a random variable, the probability of a single value of x is always:
- 0
- 1
- from 0 to 1
- greater than 1

(xiii) If the value of r between x and y is exactly zero this means that:
- x and y have the same mean
- x and y are correlated
- x and y have the same standard deviation
- the points on the scatter diagram form a circle

(xiv) The coefficient of correlation is independent of:
- origin and scale
- origin but not scale
- scale but not origin
- none of these

(xv) If a contingency table with 5 rows has 20 degrees of freedom, then the number of columns is:
- 1
- 4
- 6
- none of these

(xvi) Which of the following sampling is not a form of probability sampling?
- Simple random sampling
- Systematic sampling
- Stratified random sampling
- Judgment sampling

(xvii) $E(\bar{x})$ is equal to:
- μ
- σ^2
- 2μ
- $\frac{\sigma^2}{n}$

CamScanner

SET – D

Choose the correct answer for each from the given options:

(i) A diagram used to find the presence of correlation is called:
- Line diagram
- Scatter diagram
- Histogram
- None of these

(ii) A quantity calculated from population is called:
- Statistic
- Population
- Parameter
- Sample

(iii) The square root of the product of two regression coefficients is equal to:
- Variance
- Correlation
- Association
- None of these

(iv) The complete enumeration of population is called:
- Sample survey
- Census
- Inquiry
- None of these

(v) A variable whose value is to be estimated is called:
- independent variable
- dependent variable
- chance variable
- none of these

(vi) Stratified Random sampling is used when population is:
- Homogenous
- Heterogenous
- Large
- None of these

(vii) Another name of random variable is:
- discrete variable
- continuous variable
- chance variable
- none of these

(viii) The sum of the probability of success and failure in binomial distribution is:
- zero
- 1
- $\frac{1}{2}$
- None of these

(ix) The dependence of one variable over the other variable is called:
- Regression
- Correlation
- Association
- None of these

(x) The coefficient of correlation is unaffected by change of:
- scale
- origin
- origin and scale
- none of these

(xi) A finite population in which sampling is with replacement can be considered:
- finite
- infinite
- discrete
- continuous

(xii) In sampling human populations, the unit might be:
- an individual person
- the household
- all living persons
- all of these

(xiii) Association is said to be positive if two attributes are related in such a way that:
- $(AB)=\frac{(A)(B)}{N}$
- $(AB)>\frac{(A)(B)}{N}$
- $(AB)<\frac{(A)(B)}{N}$
- None of these

(xiv) If two coins are tossed, the probability of getting two heads is equal to:
- $\frac{1}{2}$
- $\frac{3}{8}$
- $\frac{1}{4}$
- none of these

(xv) If n coins are tossed, the number of possible outcomes in the sample space is:
- n^2
- $2n$
- $n!$
- 2^n

(xvi) The degrees of freedom of a contingency table are:
- $(rc - 1)$
- $rc(k - 1)$
- $(r - 1)(c - 1)$
- none of these

(xvii) Given N = 5 and n = 3 in simple random sampling with replacement, the possible samples are:
- 10
- 125
- 60
- 15

CamScanner

SET – E

Choose the correct answer for each from the given options:

(i) The random numbers are obtained in such a way that each digit has the probability:
- one
- zero
- equal
- none of these

(ii) Another name of random variable is:
- discrete variable
- continuous variable
- chance variable
- none of these

(iii) If x is a random variable $E(x-\mu)$ is:
- μ
- zero
- one
- none of these

(iv) The sum of the probability of success and failure in a binomial probability distribution is:
- 1.2
- zero
- 1
- $\frac{1}{2}$

(v) In a binomial probability distribution the skewness is positive for:
- $p<\frac{1}{2}$
- $p=\frac{1}{4}$
- $p=\frac{1}{2}$
- $p>\frac{1}{2}$

(vi) The degrees of freedom of a contingency table are:
- (rc – 1)
- (r – 1)(c – 1)
- k – 1
- none of these

(vii) The association of attributes is:
- positive
- negative
- zero
- all of these

(viii) A diagram that is used to check the presence of correlation is called:
- line diagram
- scatter diagram
- histogram
- none of these

(ix) A variable whose value is to be estimated is called:
- independent variable
- dependent variable
- chance variable
- none of these

(x) A probability function is a:
- positive function
- negative function
- non-negative function
- none of these

(xi) If a fair coin is tossed 4 times, the probability of at least two heads is:
- $\frac{3}{8}$
- $\frac{15}{16}$
- $\frac{11}{16}$
- $\frac{9}{5}$

(xii) The shape of a binomial distribution depends on:
- occurrence of trials
- n and p
- mean
- none of these

(xiii) Stratified random sampling is used when the population is:
- Homogenous
- Heterogenous
- large
- small

(xiv) An inquiry form comprising a number of questions is known as:
- an inquiry form
- data collection form
- questionnaire
- none of these

(xv) If A and B are two mutually exclusive events, $P(A \cap B)$ is:
- $P(A)\cdot P(B|A)$
- $P(A)\cdot P(B)$
- $P(B)\cdot P(A|B)$
- $P(\phi)$

(xvi) Venn diagram is one of the ways of representing an experiment:
- Event
- Sets
- Sample space
- Sample points

(xvii) A random variable can take only a/an:
- integral value
- random number
- discrete value
- none of these

CS CamScanner

SET – F

Choose the correct answer for each from the given options:

(i) The small part taken from a large collection of data is called:
- population
- sample
- statistic
- parameter

(ii) A quantity calculated from population is called:
- statistic
- population
- parameter
- sample

(iii) If sampling is carried out without replacement then:
- $E(\bar{x}) = \mu$
- $E(\bar{x}) = E(x)$
- $E(\bar{x}) = \sigma$
- None of these

(iv) In sampling distribution of $\bar{x}$, $\sum \bar{x} P(\bar{x})$ is:
- 1
- −1
- 0
- None of these

(v) $P(A) + P(A')$ is equal to:
- zero
- 2
- 1
- None of these

(vi) Goodness of fit test is:
- χ^2 - test
- z - test
- t - test
- None of these

(vii) The square root of the product of two regression coefficients is equal to:
- variation
- correlation
- regression
- association

(viii) The complete enumeration of population is called:
- census
- survey
- sampling
- None of these

(ix) The total area under the density function of χ^2 is:
- zero
- one
- less than one
- greater than one

(x) The dependence of one variable over the other is called:
- regression
- correlation
- testing of hypothesis
- None of these

(xi) The limit of correlation coefficient r is:
- $-1 \le r \le 1$
- $0 \le r \le 1$
- $1 \le r \le 2$
- $-1 \le r \le 0$

(xii) The number of parameter(s) in binomial probability distribution is:
- one
- two
- three
- four

(xiii) A null hypothesis is rejected when:
- $\chi^2_{tab} > \chi^2_{cal}$
- $\chi^2_{tab} < \chi^2_{cal}$
- $\chi^2_{tab} = \chi^2_{cal}$
- None of these

(xiv) The range of probability p is:
- $-1 \le p \le +1$
- $0 \le p \le 1$
- $0 \le p \le 2$
- None of these

(xv) A two-way table with r rows and c columns is called:
- frequency
- distribution table
- contingency table
- None of these

(xvi) The Binomial distribution is:
- discrete probability distribution
- continuous probability distribution
- simple probability distribution
- None of these

(xvii) As the sample size increases:
- sampling error increases
- sampling error reduces
- sampling error remains the same
- None of these

CamScanner

SET – G

Choose the correct answer for each from the given options:

(i) The dependence of one variable over the other is termed as:
- Correlation
- Dependence
- Regression
- Association

(ii) If all the plotted points tend to lie near a straight line the relationship is said to be:
- Linear
- Curvilinear
- Quadratic
- None of these

(iii) The rang of correlation coefficient is:
- $-1<r<+1$
- $0\le r\le +1$
- $r\le 1$
- $-1\le r\le +1$

(iv) The mean E(x) of random variable x is equal to:
- $\sum x\cdot P(x)$
- $\sum x^2\cdot P(x)$
- $\sum x\cdot P(x^2)$
- None of these

(v) In case of binomial distribution:
- $\bar{x}>\sigma^2$
- $\sigma^2>\bar{x}$
- $\bar{x}=\sigma^2$
- $\sigma^2\ge\bar{x}$

(vi) Area of normal curve under the limits $\mu\pm\sigma$ is:
- 95.44%
- 68.26%
- 99.73%
- 99.9%

(vii) The value of e = 2.7183 is a:
- Quantitative variable
- Continuous variable
- Discrete variable
- Constant

(viii) The number of parameter(s) in normal distribution is:
- One
- Two
- Three
- Four

(ix) Coefficient of rank correlation is:
- $1-\frac{6\sum di^2}{n(n^2-1)}$
- $\frac{1}{n}-\frac{6\sum di^2}{n(n^3-1)}$
- $1-\frac{6\sum di^2}{n^2(n^2-1)}$
- $1-\frac{6\sum di^2}{n^2(n-1)}$

(x) The correlation coefficient depends:
- on origin and scale
- on origin only
- not on origin and scale
- on scale only

(xi) If A and B are two mutually exclusive events then:
- $A\cap B=0$
- $A\cap B=\phi$
- A & B are overlapping events
- A & B are complementary events

(xii) The distribution used in goodness of fit test is:
- t - distribution
- z - distribution
- f - distribution
- χ^2 - distribution

(xiii) The standard deviation of binomial distribution is:
- np
- npq
- $\sqrt{np}$
- $\sqrt{npq}$

(xiv) The mean of sampling distribution of sample mean in case of sampling without replacement is:
- $\bar{x}$
- μ
- $E(\mu)$
- $\sigma_{\bar{x}}^2$

(xv) The mean and variance of standard normal distribution are:
- same
- μ and σ^2
- $\bar{x}$ and s^2
- 0 and 1

(xvi) If $P(A|B)=\frac{P(A\cap B)}{P(B)}$ then:
- A and B are independent
- A depends on B
- B depends on A
- dependence could not be observed

(xvii) The number of permutations from the letters in EXAM is:
- 2^4
- 4!
- ${}^4P_2$
- ${}^4C_4$

CS CamScanner

SET – H

Choose the correct answer for each from the given options:

(i) Mean of normal distribution is:
- $\bar{x}$
- μ
- $\frac{1}{\mu}$
- None of these

(ii) Coefficient of determination is:
- $\frac{1}{r}$
- r
- r^2
- None of these

(iii) Coefficient of correlation is not affected by shifting of:
- scale
- origin
- origin and scale
- None of these

(iv) Goodness of fit test is:
- χ^2 - test
- z - test
- t - test
- None of these

(v) $P(A') + P(A)$ is equal to:
- zero
- 2
- 1
- None of these

(vi) If P(A) = 0.5, P(B) = 0.6 and $P(A \cap B) = 0.3$ then $P(A \cup B)$ is equal to:
- 0.8
- 0.1
- 0.3
- None of these

(vii) Sum of probabilities of a sample space is:
- zero
- –1
- 0.5
- None of these

(viii) Mean, median and mode of normal distribution are always:
- equal
- unequal
- zero
- positive

(ix) Mean of binomial distribution is:
- npq
- pq
- np^2
- None of these

(x) If three coins are tossed, then the total number of elements in the sample space will be:
- 6
- 27
- 9
- 8

(xi) If three coins are tossed, then the probability of getting at least one head will be:
- $\frac{1}{8}$
- $\frac{5}{8}$
- $\frac{1}{9}$
- $\frac{7}{8}$

(xii) The complete enumeration of a population is called:
- census
- survey
- sampling
- stratified sampling

(xiii) The value calculated from sample is known as:
- Parameter
- Statistic
- Discrete
- Primary

(xiv) If two variables have the same data, then the coefficient of correlation will be:
- 1
- 0.5
- –1
- 0

(xv) Dependence of one variable upon another variable is called:
- variation
- correlation
- regression
- association

(xvi) All possible samples of size 3, from a population of size 5, with replacement will be:
- 12
- 125
- 15
- 1

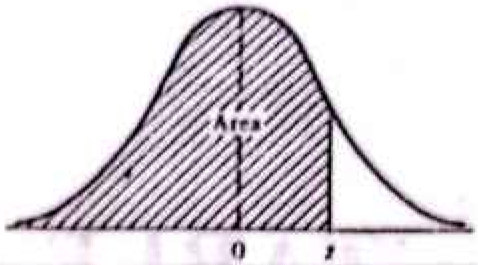
(xvii) The limit of rand coefficient of correlation is:
- from 0 to 1
- from –1 to 1
- from 0.5 to 1
- from 0 to 0.5

CS CamScanner

STATISTICAL TABLES

CamScanner

Z – TABLE

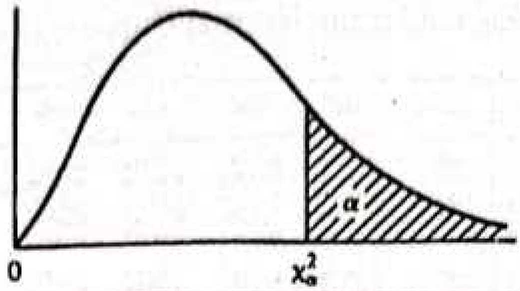


Areas Under the Normal Curve

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

CamScanner

χ^2 – TABLE



Critical Values
of the Chi-Square Distribution

ν	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
	α							
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	3.841	5.024	6.635	7.879
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672

CS CamScanner

ANSWERS

CamScanner

ANSWERS

EXERCISE – 1

1.1 $y = 1.75 + 0.25x$ **1.2** $y = 1.36 + 0.5x$ **1.3** $y = -0.43 - x$

1.4 $y = 1.5 + 0.5x$, $\hat{y} = 6$ **1.5** $y = 11.9 - 0.65x$, $\hat{y} = 4.1$

1.6 $y = 1.02 + 0.73x$, $\hat{y} = 10.51$ **1.7** $y = 1.3 + 0.9x$, $\hat{y} = 6.25$

1.8 $y = -8.7 + 2.1x$, $\hat{y} = 12.3$ **1.9** $y = 1.46 + 1.33x$, $\hat{y} = 11.44$

1.10 **(a)** $y = .51 + 0.35x$ **(b)** $\hat{y} = 9.41$

1.11 **(i)** $y = 0.99 + 2x$ **(ii)** $\hat{y} = 30.99$ **1.12** $x = 3.44 - 0.28y$

1.13 $x = 16.4 - 1.3y$, $\hat{x} = 6.65$ **1.14** $\hat{y} = 11.9 - 0.65x$, $\hat{x} = 16.4 - 1.3y$

1.15 $y = 5.2 + 2.1x$, $x = -2.03 + 0.39y$

1.16 **(a)** $y = 2$, $\hat{y} = 2$ **(b)** $x = 3.17$, $\hat{x} = 3.17$

1.17 **(i)** $\hat{x} = 7.9$ **(ii)** $\hat{y} = 3.9$ **1.18** **(a)** 8.2 tons **(b)** \$ 1.63

1.19 Expected height = 61.7 inches **1.20** $\hat{y} = 31.35$

1.22 **(i)** $y = 52.4 + 0.55x$ **(ii)** $\hat{y} = 79.9$

1.23 **(b)** Estimated No. of persons = 150

1.24 $r = 0.8$ **1.25** $r = -0.68$ **1.26** $r = 0.9$ **1.27** $r = 0.96$

1.28 $r = 0.76$ **1.29** $r = 0.86$ **1.30** $r = -0.92$

1.31 $r = 0.95$, when $r = 1$, correlation is said to be perfect +ve

when $r = 0$, there is no correlation

when $r = -1$, correlation is said to be perfect –ve

1.32 $r = 0.92$ **1.33** $r = -0.26$ **1.34** $r = -0.68$

1.35 $y = 11.3 - 0.65x$, $\hat{y} = 5.8$

CamScanner

1.36 (i) $r = 0.97$ (ii) $y = 19.99 + 3.56x$ (iii) $\hat{y} = 91.19$

1.37 (b) $y = -1 + x$, $\hat{y} = 12$ and $r = 1$

1.38 (b) – (i) $y = 6.89 + 0.62x$ (ii) $\hat{y} = 14.95$

1.39 (b) – (i) $r = 0.98$ (ii) $y = -19.38 + 1.13x$ (iii) $r_{uv} = 0.98$

1.40 (b) $y = 1.36 + 0.5x$ 1.41 (b) $r = 0.84$, $b_{yx} = 0.64$, $b_{xy} = 1.11$

1.42 (c) $r = 0.95$, High degree +ve correlation. 1.43 (c) $r = 0.15$

1.44 (b) $r = 0.67$ 1.45 (i) $x = -204.52 + 1.72y$ (ii) $\hat{y} = 167.25$ (iii) $r = 0.86$

1.46 (b) $r = 0.24$, Correlation is weak

1.47 (b) Estimated cost = 22.05 (thousand rupees)

1.48 (b) – (i) $y = -19.91 + 2.89x$ (ii) $\hat{y} = 110.14$ 1.49 (b) $y = 0.54 + 3.94x$

1.50 (b) $y = 10 - 1.4x$, $\hat{y} = 0.2$ 1.51 (b) $r = 1$, Correlation is perfect +ve.

1.52 (b) – (i) $y = 9 + 0.75x$ (ii) $\hat{y} = 20.25$

1.54 (b) $r = 0.90$, high degree +ve correlation. 1.55 (b) $r = 0.95$

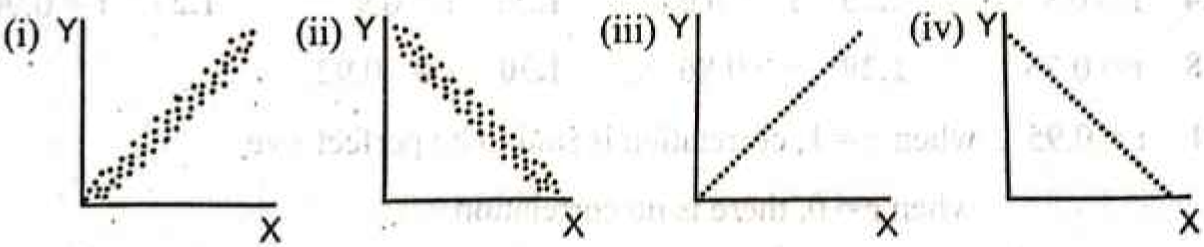
1.56 (i) $r = 0.55$ (ii) $y = 4.88 + 0.86x$ (iii) $r = 0.55$

1.57 (b) $r = 0.65$ 58. (b) $r = -0.615$

1.59 (b) $y = 56.75 + 0.075x$, $x = -176.4 + 4.8y$
Estimated height of Ahmed = $\hat{y} = 71.75$ (inches)
Estimated weight of Kamil = $\hat{x} = 111.6$ (lbs)

1.60 (b) $y = 10 - 1.4x$, $\hat{y} = 0.2$

1.61 (a)



(b) $r = 0.60$

1.62 (b) $y = 4.1 - 0.5x$, $\hat{y} = 2.35$

CamScanner

1.63 (b) – (i) $r = 0.98$ (ii) $y = 15 + 0.16x$, $\hat{y} = 95.00$

1.64 (b) – (i) $y = -78.23 + 2.88x$, $x = 22.20 + 0.28y$ (ii) $\hat{y} = 94.57$, $\hat{x} = 71.2$

(iii) $r = 0.90$ (iv) $r^2 = 0.81$

1.65 (b) $r = -0.15$, $y = 43.17 - 0.24x$ **1.66** $y = 0.9 + 2.94x$, $\hat{y} = 27.36$

1.67 (b) – (i) $\hat{y} = -3.2 + 1.44x$, $\hat{x} = 2.76 + 0.56y$ (ii) $r^2 = 0.81$

1.68 (b) $r = -0.88$, $y = 47.43 - 2.08x$ **1.69** (a) $r = 0.91$ (b) $r = -0.91$

1.70 $b_{yx} = 0.9$, $b_{xy} = 0.98$, $r = 0.94$

1.71 (a) $r = 0.770$ (b) – (i) $r = -0.29$ (ii) $\bar{x} = 1.67$ (iii) $\bar{y} = 10.33$

1.72 (c) $r = -0.24$ **1.73** (a) $y = 4.84 + 1.08x$ (b) $\hat{y} = 16.72$ (c) $r = 0.87$

1.74 (b) – (i) $\bar{x} = 13$, $\bar{y} = 17$ (ii) S. D. of $y = 4$ (iii) $r = 0.6$

1.75 (a) – (i) $y = 2 + 0.11x$ (ii) $\hat{y} = 3.1$

1.76 (b) $y = 114.55 - 1.45x$, $\hat{y} = 60.9$ **1.77** (b) $r = 0.97$ **1.78** (b) $r_s = 0.85$

1.79 $r_s = 0.62$ **1.80** $r_s = 0.71$ **1.81** $r_s = -0.21$ **1.82** $r = 0.77$

EXERCISE – 2

2.1 24 **2.2** 60 **2.3** 468000 **2.4** (i) 24 (ii) 576

2.5 720 **2.6** 720 **2.7** 720 **2.8** 24 **2.9** 60

2.10 60 **2.11** 60 **2.12** (a) 5040 (b) 720

2.13 120 (i) 24 (ii) 6

2.14 (a) 12 (b) – (i) 5040 (ii) 3628800 (iii) 1814400

2.15 (a) 120 (b) 3125 **2.16** (a) 125 (b) 60

2.17 (a) 100 (b) 48 **2.18** (a) 500 (b) – (i) 48 (ii) 100

2.19 120 **2.20** (i) 9 (ii) 6 **2.21** (i) 384 (ii) 576

2.22 (a) – (i) 12 (ii) 9 (iii) 1 (iv) 116280 (b) 40320

CS CamScanner

2.23 60 2.24 (a) 360 (b) 120 2.25 14400

2.26 (i) 48 (ii) 100 2.27 (i) 180 (ii) 75 (iii) 105

2.28 (i) 294 (ii) 144 2.29 504

2.30 (a) 2880 (b) – (i) 120 (ii) 12 (iii) 12 2.31 207360 2.32 1764000

2.33 180 2.34 648 2.35 (a) 180 (b) 75 (c) 105

2.36 (b) 3360 2.37 3360 2.38 840 2.39 180

2.40 29937600 2.41 4989600 2.42 50400 2.43 27720

2.44 27720 2.45 1260 2.46 1441440 2.47 210

2.48 (a) 10 (b) 126 2.49 (i) 120 (ii) 84 (iii) 92378 (iv) 10 (v) 190

2.50 (a) 455 (b) 252 2.51 (b) 210 (c) 720 2.52 (a) 560 (b) 150

2.53 84 2.54 180 2.55 50 2.56 30

2.57 (i) 164 (ii) 1 (iii) 210 2.58 (b) 25

2.59 (i) 126 (ii) 81 2.60 (a) – (i) 30240 (ii) 252 (iii) 2520

2.60 (b) No. of Combinations $= {}^4C_3 = 4$, No. of Permutations $= {}^4p_3 = 24$

2.61 (b) – (i) 100 (ii) 30 (iii) 56 2.62 16 2.63 108336

2.64 (i) 350 (ii) 210 2.65 140

2.66 (a) 364 (i) 286 (ii) 78 (b) 126 2.67 210

2.68 (i) 792 (ii) 220 2.69 280 , 84 2.70 (i) 15 (ii) 18

2.71 (b) – (i) 294 (ii) 168 (c) 56 2.72 S = { H , T }

2.73 (i) S = { (1,1) (1,2) .... (6,6) } (ii) S = { HH , HT , TH , TT }

(iii) S = { (1H) (2H) (3H) (4H) (5H) (6H) (1T) (2T) (3T) (4T) (5T) (6T) }

(iv) S = { HHH , HHT , HTH , THH , HTT , THT , TTH , TTT }

2.74 (i) S = { HHH , HHT , HTH , THH , HTT , THT , TTH , TTT }

(ii) A = { TTH , THT , HTT } (iii) B = { HHT , HTH , THH , HHH }

2.75 (a) S = { (1,1) (1,2) .... (6,6) } (b) A = { (1,1) (1,2) (2,1) (1,3) (3,1) (2,2) }

CamScanner

(c) B = { (4,1) (4,2) (4,3) (4,4) (4,5) (4,6) (1,4) (2,4) (3,4) (5,4) (6,4) }

(d) C = { (3,1) (3,2) (3,3) (3,4) (3,5) (3,6) } (e) D = { } = ϕ

2.76 (a) S = { HHH , HHT , HTH , THH , HTT , THT , TTH , TTT }

(b) A = { HTT , THT , TTH , TTT } (c) B = { HTT , THT , TTH }

2.77 (i) S = { HHHH , HHHT , HHTH , HTHH , THHH , HHTT , HTHT , THTH , TTHH , THHT , HTTH , HTTT , THTT , TTHT , TTTH , TTTT }

(ii) A = { HHTT , HTHT , THTH , TTHH , HTTH , THHT , TTTH , TTHT , THTT , HTTT , TTTT }

(iii) B = { HHTT , TTHH , HTTH , THHT , HTHT , THTH }

2.78 (i) S = { HH , HT , TH , TT }

(ii) S = { HHH , HHT , HTH , THH , HTT , THT , TTH , TTT }

(iii) S = { (H1) (H2) (H3) (H4) (H5) (H6) (T1) (T2) (T3) (T4) (T5) (T6) }

EXERCISE – 3

3.1 P(Even) = $\frac{1}{2}$, P(4 or higher) = $\frac{1}{2}$ 3.2 (i) $\frac{1}{2}$ (ii) $\frac{1}{4}$ (iii) $\frac{3}{4}$

3.3 $\frac{1}{2}$ 3.4 $\frac{1}{4}$ 3.5 (i) Zero (ii) $\frac{1}{9}$ (iii) $\frac{1}{2}$ 3.6 $\frac{1}{4}$

3.7 S = { HHH , HHT , HTH , THH , HTT , THT , TTH , TTT } , P(at least 1 - H) = $\frac{7}{8}$.

3.8 (i) $\frac{1}{6}$ (ii) $\frac{1}{12}$ 3.9 (a) $\frac{5}{12}$ (b) $\frac{1}{6}$

3.10 (i) $\frac{1}{9}$ (ii) $\frac{8}{9}$ (iii) $\frac{35}{36}$ 3.11 (b) $\frac{1}{6}$ 3.12 $\frac{5}{18}$ 3.13 $\frac{5}{36}$

3.14 $\frac{1}{6}$ 3.15 (a) $\frac{1}{6}$ (b) $\frac{1}{6}$ (c) $\frac{5}{6}$ (d) $\frac{5}{18}$ (e) $\frac{35}{36}$

3.16 $\frac{1}{4}$ 3.17 (i) $\frac{1}{6}$ (ii) $\frac{5}{18}$ (iii) $\frac{5}{12}$

CamScanner

3.18 (a) – (i) $\frac{1}{12}$ (ii) $\frac{5}{12}$ (b) – (i) $\frac{7}{12}$ (ii) $\frac{1}{6}$ (iii) $\frac{5}{36}$

3.19 S = { HHH , HHT , HTH , THH , HTT , THT , TTH , TTT } , $\frac{1}{2}$

3.20 (i) $\frac{1}{6}$ (ii) $\frac{5}{6}$ **3.21** $\frac{1}{6}$ **3.22** $\frac{1}{6}$ **3.23** $\frac{1}{3}$

3.24 (i) $\frac{1}{13}$ (ii) $\frac{1}{4}$ **3.25** (a) 0.47 (b) 0.55 **3.26** $\frac{13}{20}$

3.27 (i) 0.13 (ii) 0.36 (iii) 0.51 **3.28** 0.11 **3.29** 0.12 **3.30** 0.12

3.31 $\frac{2}{3}$ **3.32** (i) $\frac{15}{28}$ (ii) $\frac{13}{28}$ **3.33** 0.48 **3.34** (i) $\frac{2}{7}$ (ii) $\frac{1}{7}$ (iii) $\frac{4}{7}$

3.35 (i) 0.045 (ii) 0.27 **3.36** (i) 0.045 (ii) 0.02 (iii) 0.18 (iv) 0.27

3.37 0.42 **3.38** 0.5357 **3.39** $\frac{1}{54145}$ **3.40** 0.43 **3.41** 0.44

3.42 (i) 0.12 (ii) 0.19 **3.43** (i) 0.3 (ii) 0.3

3.44 (i) $\frac{5}{36}$ (ii) $\frac{15}{16}$ (iii) 0.14 **3.45** 0.48

3.46 (i) 0.18 (ii) 0.11 **3.47** 0.23 **3.48** 0.02 **3.49** $\frac{2}{9}$

3.50 (ii) A and C (v) B and D (vi) C and D **3.51** (c) $\frac{7}{26}$

3.52 (a) $\frac{2}{13}$ (c) $\frac{1}{4}$ **3.53** (i) 0.9 (ii) Zero (iii) 0.6

3.54 (i) $\frac{1}{3}$ (ii) $\frac{11}{15}$ (iii) $\frac{2}{5}$ (iv) $\frac{4}{5}$ **3.56** (i) 0.17 (ii) 0.97

3.57 (i) 0.28 (ii) 0.06 (iii) 0.19 **3.59** (i) $\frac{1}{3}$ (ii) $\frac{7}{12}$ (iii) $\frac{2}{3}$

3.60 (a) – (i) $\frac{2}{7}$ (ii) $\frac{17}{42}$ (b) – (i) $\frac{5}{36}$ (ii) $\frac{3}{8}$ (iii) $\frac{8}{663}$ (iv) $\frac{4}{9}$

CS CamScanner

3.61 0.83 **3.62** $\frac{2}{5}$ **3.63** (a) $\frac{1}{13}$ (b) $\frac{1}{4}$ (c) $\frac{2}{13}$ (d) $\frac{3}{13}$

3.64 (a) $\frac{1}{221}$ (b) $\frac{325}{1326}$ (c) $\frac{33}{221}$ **3.65** $\frac{5}{6}$ **3.66** $\frac{3}{5}$

3.67 (a) $\frac{11}{26}$ (b) $\frac{4}{13}$ **3.68** $\frac{1}{12}$ **3.69** (b) $\frac{4}{13}$ **3.70** 0.87

3.71 (b) $\frac{4}{13}$ **3.73** (b) $\frac{7}{13}$ **3.74** (b) 0.925

3.75 (b) – (i) 0.95 (ii) 0.6 (iii) 0.05 **3.76** (b) $\frac{11}{36}$

3.77 $P(B') = 0.4$, $P(A \text{ or } B) = 0.7$ **3.78** (b) – (i) $\frac{1}{6}$ (ii) $\frac{2}{3}$

3.79 (b) 0.77 **3.80** (b) $\frac{13}{120}$ **3.82** (i) 0.4 (ii) 0.65

3.83 $\frac{11}{12}$ **3.84** 0.68 **3.85** (a) $\frac{71}{72}$ (b) – (i) $\frac{1}{2}$ (ii) $\frac{4}{13}$ **3.86** $\frac{4}{5}$

3.87 (a) – (i) 0.6 (ii) 0.6 (b) $P(A) = \frac{5}{36}$, $P(B) = \frac{1}{6}$, $P(A \cap B) = \frac{5}{216}$, $P(A \cup B) = 0.28$

3.88 (i) $\frac{3}{10}$ (ii) $\frac{9}{20}$ **3.89** (i) $\frac{2}{13}$ (ii) $\frac{7}{13}$ **3.90** $\frac{14}{15}$ **3.91** (b) $\frac{14}{45}$

3.92 (a) – (i) Not - Mutually Exclusive (ii) Not - Independent

(b) – (i) $\frac{1}{10}$ (ii) $\frac{3}{50}$ (iii) 1

3.93 (a) – (i) $\frac{1}{12}$ (ii) $\frac{1}{2}$ (b) $\frac{1}{6}$ (c) – (i) $\frac{3}{10}$ (ii) $\frac{1}{2}$

3.94 (a) – (i) 0.42 (ii) 0.88 (b) – (i) $\frac{1}{10}$ (ii) $\frac{1}{25}$ (iii) 1

3.95 (i) $\frac{5}{9}$ (ii) $\frac{1}{2}$ **3.96** (i) $\frac{1}{16}$ (ii) $\frac{1}{17}$ **3.97** (i) $\frac{1}{169}$ (ii) $\frac{1}{221}$

3.98 (i) $\frac{5}{18}$ (ii) $\frac{13}{18}$ **3.99** (a) – (i) $\frac{4}{15}$ (ii) $\frac{8}{15}$ (iii) $\frac{2}{3}$ (b) – (iii) $\frac{4}{663}$

3.100 (i) $\frac{1}{10}$ (ii) $\frac{3}{10}$ **3.101** (a) $P(A|B)=0.6$, $P(A \cup B)=0.76$ (b) 0.65

3.103 (a) – (i) 0.43 (ii) 0.6 **3.104** (i) $\frac{5}{36}$ (ii) $\frac{19}{27}$

3.105 (b) $\frac{5}{18}$ (c) $\frac{17}{32}$ **3.107** (b) $\frac{8}{5525}$ **3.108** (b) 0.32

3.109 (i) $\frac{7}{12}$ (ii) $\frac{3}{4}$

EXERCISE – 4

4.2 (a)

x	0	1	2	3
P(x)	$1/8$	$3/8$	$3/8$	$1/8$

(b)

x	0	1	2	3
P(x)	$1/8$	$3/8$	$3/8$	$1/8$

$E(X) = 1.5$

4.3 (a)

x	0	1	2	3
P(x)	$10/56$	$30/56$	$15/56$	$1/56$

(b)

x	0	1	2	3	4	5
P(x)	$1/32$	$5/32$	$10/32$	$10/32$	$5/32$	$1/32$

$E(X) = 2.5$

4.4 (b) Mean = 2.59 , Variance = 1.44

(b)

x	0	1	2	3
P(x)	$1/8$	$3/8$	$3/8$	$1/8$

$E(X) = 1.5$, $V(X) = 0.75$

CamScanner

4.6 (b)

x	0	1	2	3
P(x)	$4/56$	$24/56$	$24/56$	$4/56$

$E(X) = 1.5$

4.7 (i)

x	1	2	3	4	5	6	7	8
P(x)	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$

(ii) $E(X) = 4.5$, $V(X) = 5.25$

4.8 (a) – (i)

x	0	1	2
P(x)	$\frac{1128}{1326}$	$\frac{192}{1326}$	$\frac{6}{1326}$

(ii) $E(X) = 0.15$ (iii) $E(x - \bar{x}^2) = 0.14$

(c) – (i) Mean = 1.3 (ii) Variance = 0.81

4.9 Mean = 1.4 , Variance = 0.64

4.10

x	0	1	2	3
P(x)	$1/35$	$12/35$	$18/35$	$4/35$

Mean = 1.71 , Variance = 0.49

4.11 (b) Zero **4.12** (b) – (i) 0.67 (ii) 0.01 **4.13** (i) $\frac{7}{2}$ (ii) $\frac{91}{6}$

4.14 (b) Mean = 7 , Variance = $\frac{35}{6}$ **4.15** (b) Mean = 16 , Variance = 20

4.17 (c) $E(X) = 1.6$

4.18

x	1	2	3	4
P(x)	$1/10$	$2/10$	$3/10$	$4/10$

Mean = 3 and $E(2X + 3) = 9$

4.19 (b) – (i) 0.33 (ii)

y	2	5	8	11	14
P(y)	0.8	0.11	0.06	0.02	0.01

$E(3X + 2) = 2.99$

4.21 (b) $E(X) = 0.65$

4.22 (c) Expectation of A = Rs. 72 , Expectation of B = Rs. 36 , Expectation of C = Rs. 18

4.23 (b) $E(X^2 + 2X - 3) = 10.85$ 4.24 (b) $\bar{Y} = 0$ and $V(Y) = 144$

4.25 (b) \$1.5 4.26 (b) Mean = 2.375 , Variance = 1.48

4.27 (b) $k = 2$ (c) – (i) $V(X) = 2.92$ (ii) $V(Y) = 0.5$ (iii) $V(X + Y) = 3.42$

4.28 (c) $V(X) = 0.74$ (d) $E(X) = 3.20, E(Y) = 3.0, E(XY) = 9.6, E(X + Y) = 6.20$

4.29 (d) 0.17 (e) 0.11

4.30 (b) $\left(\frac{2}{5}+\frac{3}{5}\right)^{10} = \left(\frac{2}{5}\right)^{10} + {}^{10}C_1\left(\frac{2}{5}\right)^{9}\left(\frac{3}{5}\right) + {}^{10}C_2\left(\frac{2}{5}\right)^{8}\left(\frac{3}{5}\right)^{2} + \dots\dots\dots + \left(\frac{3}{5}\right)^{10}$

Mean = 4 , Variance = 2.4

4.31 0.22 4.32 (i) 0.001 (ii) 0.80

4.34 (a) 2.4 , 0.96 (b) Mean = $\frac{3}{4}$, Variance = $\frac{9}{16}$

4.35 (a) – (i) Mean = 30 , Variance = 15 (ii) Mean = 10 , Variance = 9

(b) – (i) Binomial (ii) 4.5 (iii) 2.25 (iv) 9 (v) Discrete

4.36 (b) – (i) 0.0055 (ii) 0.29 4.38 (b) 0.89 4.39 (b) 0.06

4.40 $P(x) = {}^{5}C_x\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{5-x}$, $x = 0, 1, 2, 3, 4, 5$, Average No. of Boys = 2.5

4.41 (b) 0.07 4.42 (b) $n = 100$, $p = 0.36$ (c) Not possible

4.43 (a) mean = np and variance = npq (b) 0.81 4.44 (b) 1.5

4.45 (a) – (i) 4 (ii) 2 (iii) 1.4142 (iv) 35.36% (b) Mean = 15 , Variance = 22.5

4.46 (c) – (i) Mean = 10 , Variance = 9 (ii) Mean = 30 , Variance = 15

4.47 (ii) 164

4.48 (a) $E(X) = 1$, $V(X) = 0.75$ (c) $P(x) = {}^{4}C_x\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{4-x}$, $x = 0, 1, 2, 3, 4$

(d) $E(X) = 2$, $V(X) = 1$

4.49 (i) 3.5 , 1.32 (ii) 3.33 , 0.75 4.50 (b) – (i) 0.16 (ii) 0.0039

4.51 0.23 **4.52** (b) 0.23

4.53 (a) – (i) $P(x) = {}^5C_x \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{5-x}$, $x = 0, 1, 2, 3, 4, 5$

(ii) Mean = 3.33 , Variance = 1.11 (c) 0.32

4.54 (b) When $p = \frac{1}{2}$ (c) $n = 100$, $p = \frac{1}{2}$ (d) $n = 17$, $p = 0.6$

4.55 (c) $n = 100$, $p = 0.36$ **4.56** (c) – (i) 0.17 (ii) 0.52

4.57 (b) – (i) 0.9772 (ii) 0.9772 (iii) 0.9848

4.58 (a) – (i) 0.0179 (ii) 0.9192 (b) – (i) 0.0062 (ii) 0.1587 (iii) 0.8664

4.59 (b) – (i) 0.5764 (ii) 0.0808 (iii) 0.1587 **4.60** (b) 0.95

4.61 (b) – (i) 0.0228 (ii) 0.1587 (iii) 0.2119 (iv) 0.9192

4.62 (b) – (i) 0.5461 (ii) 0.1612 (iii) 0.1357

4.63 (b) – (i) 0.1554 (ii) 0.3446 (iii) 0.2119

4.64 (b) – (i) 0.5403 (ii) 0.8413 (iii) 0.2119

4.65 (b) – (i) 0.6826 (ii) 0.0668 (iii) 0.3085

4.66 (b) – (i) 0.1056 (ii) 0.8944 (iii) 0.6301

4.67 (b) – (i) 119 (ii) 484 (iii) 670

4.68 (b) – (i) 227 (ii) 6 (iii) 2 (c) (i) 0.063 (ii) 0.0166 (iii) 0.7314

(e) – (i) 159 (ii) 692 **4.69** (c) 1 **4.70** (b) $\mu = 8$ and $\sigma^2 = 1$

4.71 (b) – (i) 0.9938 (ii) 0.9938 (iii) 0.9876 (iv) 0.5795

(c) – (i) 0.9192 (ii) 0.9821 (iii) 0.6106

(d) – (i) 0.9938 (ii) 0.9938 (iii) 0.95

(e) – (i) 0.3674 (ii) 0.3446 (iii) 0.1587

4.72 (b) – (i) 0.5461 (ii) 1 (c) – (i) 0.5111 (ii) 0.1587

(d) – (i) 0.3674 (ii) 0.1587 (iii) 0.3446

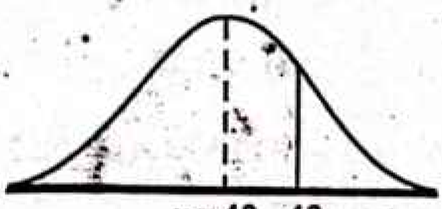
4.73 (c) – (i) 0.6589 (ii) 0.0668 (iii) 0.1587

4.74 (d) – (i) –1.25 (ii) 19.6 (e) –1.25 **4.75** (b) 0.5

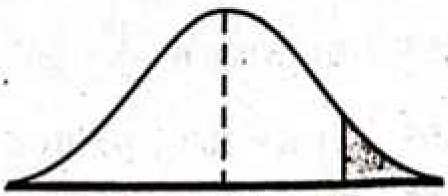
4.77 **(i)** 10 **(ii)** Normal Distribution **(iii)** 25 **(iv)** $\mu = 10$ **(v)** Zero **(vi)** Binomial Distribution **(vii)** $\frac{1}{5\sqrt{2\pi}}$ **(viii)** μ **(ix)** Continuous **(x)** Bell

4.79 **(a)** – (i) 0.9641 (ii) 0.0013

(b) – (i)



(ii)



4.80 **(b)** – (i) 0.0668 (ii) 0.0062 (iii) 0.9198

4.81 **(a)** 0.6826 **(b)** 0.9544

4.82 **(b)** – (i) 0.1446 (ii) 0.0909 **(c)** – (i) 0.0548 (ii) 0.3616 (iii) 0.6898

4.83 **(c)** – (i) 0.5344 (ii) 0.9994 (iii) 0.0876

4.84 **(i)** 0.8643 **(ii)** 0.2978 **(iii)** 0.0796

EXERCISE – 5

5.12 **(b)**

Samples	$\bar{x}$
(2 , 4)	3
(2 , 6)	4
(2 , 8)	5
(2 , 10)	6
(4 , 6)	5
(4 , 8)	6
(4 , 10)	7
(6 , 8)	7
(6 , 10)	8
(8 , 10)	9

$E(\bar{X}) = \mu = 6$

5.13

Samples	$\bar{x}$
(8 , 2)	5
(8 , 10)	9
(8 , 4)	6
(2 , 10)	6
(2 , 4)	3
(10 , 4)	7
Total	36

$E(\bar{x}) = \mu = 6$

5.15 (i)

Samples	$\bar{x}$
(0 , 3 , 6)	3
(0 , 3 , 9)	4
(0 , 3 , 12)	5
(0 , 6 , 9)	5
(0 , 6 , 12)	6
(0 , 9 , 12)	7
(3 , 6 , 9)	6
(3 , 6 , 12)	7
(3 , 9 , 12)	8
(6 , 9 , 12)	9
Total	60

(ii) $E(\bar{x}) = \mu = 6$

5.17 (b) - (i)

Samples	$\bar{x}$
(8 , 10)	9
(8 , 12)	10
(8 , 15)	11.5
(8 , 20)	14
(10 , 12)	11
(10 , 15)	12.5
(10 , 20)	15
(12 , 15)	13.5
(12 , 20)	16
(15 , 20)	17.5
Total	130

$E(\bar{x}) = \mu = 13$

(ii) $V(\bar{x}) = 6.6$

5.18

Samples	$\bar{x}$
(2 , 3)	2.5
(2 , 5)	3.5
(2 , 7)	4.5
(2 , 9)	5.5
(3 , 5)	4
(3 , 7)	5
(3 , 9)	6
(5 , 7)	6
(5 , 9)	7
(7 , 9)	8
Total	52

$E(\bar{x}) = \mu = 5.2$, $V(\bar{x}) = 2.46$

5.19 (b)

Samples	$\bar{x}$
(9 , 0 , 15)	8
(9 , 0 , 12)	7
(9 , 0 , 18)	9
(9 , 15 , 12)	12
(9 , 15 , 18)	14
(9 , 12 , 18)	13
(0 , 15 , 12)	9
(0 , 15 , 18)	11
(0 , 12 , 18)	10
(15 , 12 , 18)	15
Total	108

$E(\bar{x}) = \mu = 10.8$

CamScanner

5.20 (a)

Samples	$\bar{x}$
(0 , 3 , 12)	5
(0 , 3 , 18)	7
(0 , 3 , 30)	11
(0 , 12 , 18)	10
(0 , 12 , 30)	14
(0 , 18 , 30)	16
(3 , 12 , 18)	11
(3 , 12 , 30)	15
(3 , 18 , 30)	17
(12 , 18 , 30)	20
Total	126

(i) $E(\bar{x}) = \mu = 12.6$

(ii) $V(\bar{x}) = 19.44$

(b)

Samples	$\bar{x}$
(6 , 4 , 10)	6.7
(6 , 4 , 8)	6
(6 , 4 , 2)	4
(6 , 10 , 8)	8
(6 , 10 , 2)	6
(6 , 8 , 2)	5.3
(4 , 10 , 8)	7.3
(4 , 10 , 2)	5.3
(4 , 8 , 2)	4.7
(10 , 8 , 2)	6.7
Total	60

$E(\bar{x}) = \mu = 6$

$V(\bar{x}) = 1.334$

5.21 (b)

Samples	$\bar{x}$
(1 , 3)	2
(1 , 5)	3
(1 , 7)	4
(1 , 9)	5
(1 , 11)	6
(3 , 5)	4
(3 , 7)	5
(3 , 9)	6
(3 , 11)	7
(5 , 7)	6
(5 , 9)	7
(5 , 11)	8
(7 , 9)	8
(7 , 11)	9
(9 , 11)	10
Total	90

$E(\bar{x}) = \mu = 6$

5.22 (a)

Samples	$\bar{x}$
(9 , 9)	9
(9 , 15)	12
(15 , 9)	12
(15 , 15)	15
Total	48

$E(\bar{x}) = \mu = 12$

$V(\bar{x}) = 4.5$

(b)

Samples	$\bar{x}$
(9 , 9)	9
(9 , 11)	10
(9 , 15)	12
(9 , 21)	15
(11 , 9)	10
(11 , 11)	11
(11 , 15)	13
(11 , 21)	16
(15 , 9)	12
(15 , 11)	13
(15 , 15)	15
(15 , 21)	18
(21 , 9)	15
(21 , 11)	16
(21 , 15)	18
(21 , 21)	21
Total	224

$E(\bar{x}) = \mu = 14$

$V(\bar{x}) = \frac{\sigma^2}{2} = 10.5$

5.23 (b)

Samples without Replacement	Samples with Replacement		
(4 , 7)	(4 , 4)	(9 , 4)	(15 , 4)
(4 , 9)	(4 , 7)	(9 , 7)	(15 , 7)
(4 , 12)	(4 , 9)	(9 , 9)	(15 , 9)
(4 , 15)	(4 , 12)	(9 , 12)	(15 , 12)
(7 , 9)	(4 , 15)	(9 , 15)	(15 , 15)
(7 , 12)	(7 , 4)	(12 , 4)	
(7 , 15)	(7 , 7)	(12 , 7)	
(9 , 12)	(7 , 9)	(12 , 9)	
(9 , 15)	(7 , 12)	(12 , 12)	
(12 , 15)	(7 , 15)	(12 , 15)	

CS CamScanner

5.24 **(b)** S. E. of $(\overline{X}) = 1.63$

5.26 **(b)**

Samples	$\bar{x}$
(1 , 2)	1.5
(1 , 3)	2
(1 , 4)	2.5
(1 , 5)	3
(2 , 3)	2.5
(2 , 4)	3
(2 , 5)	3.5
(3 , 4)	3.5
(3 , 5)	4
(4 , 5)	4.5
Total	30

$E(\bar{x}) = \mu = 3$, $V(\bar{x}) = 0.75$

5.31 **(b)**

Samples	$\bar{x}$
(3 , 6 , 15)	8
(3 , 6 , 18)	9
(3 , 6 , 27)	12
(3 , 15 , 18)	12
(3 , 15 , 27)	15
(3 , 18 , 27)	16
(6 , 15 , 18)	13
(6 , 15 , 27)	16
(6 , 18 , 27)	17
(15 , 18 , 27)	20

(i) $E(\bar{x}) = \mu = 13.8$ (ii) $V(\bar{x}) = 12.36$

CamScanner

5.34

Samples	$\bar{x}$
(0 , 6 , 3)	3
(0 , 6 , 12)	6
(0 , 6 , 15)	7
(0 , 6 , 19)	8.33
(0 , 3 , 12)	5
(0 , 3 , 15)	6
(0 , 3 , 19)	7.33
(0 , 12 , 15)	9
(0 , 12 , 19)	10.33
(0 , 15 , 19)	11.33
(6 , 3 , 12)	7
(6 , 3 , 15)	8
(6 , 3 , 19)	9.33
(6 , 12 , 15)	11
(6 , 12 , 19)	12.33
(6 , 15 , 19)	13.33
(3 , 12 , 15)	10
(3 , 12 , 19)	11.33
(3 , 15 , 19)	12.33
(12 , 15 , 19)	15.33
Total	183.3

$E(\bar{x}) = \mu = 9.17$, $V(\bar{x}) = 9.02$

EXERCISE – 6

6.3 (b) $\chi^2_{Cal} = 26.6$, Reject H_o

6.5 $\chi^2_{Cal} = 10.43183$, Reject H_o

6.6 (b) $\chi^2_{Cal} = 5.49$, Reject H_o

6.7 (b) $\chi^2_{Cal} = 1.62$, Accept H_o

6.8 (b) $\chi^2_{Cal} = 6.62$

6.9 $\chi^2_{Cal} = 12.65$, Reject H_o

6.10 $\chi^2_{Cal} = 18.78$, Reject H_o

CamScanner

6.11 $\chi^2_{Cal} = 7.96$ (i) Reject at $\alpha = 0.05$ (ii) Reject at $\alpha = 0.01$

6.12 $\chi^2_{Cal} = 23.69$, Reject H_o

6.13 $\chi^2_{Cal} = 14.46$, Reject H_o

6.14 (b) $\chi^2_{Cal} = 9.5691$, Accept H_o

6.15 $\chi^2_{Cal} = 5.86$, Reject H_o

6.16 $\chi^2_{Cal} = 12.698$, Reject H_o

6.17 $\chi^2_{Cal} = 9.928$, Reject H_o

6.18 $\chi^2_{Cal} = 26.606$, Reject H_o

6.19 (b) $\chi^2_{Cal} = 55.06$, Reject H_o

6.20 (b) $\chi^2_{Cal} = 11.69$, Reject H_o

6.21 $\chi^2_{Cal} = 0.00715$, Accept H_o

6.23 $\chi^2_{Cal} = 7.464$, Accept H_o

6.24 $\chi^2_{Cal} = 110.080$, Reject H_o

6.25 $\chi^2_{Cal} = 1.376$, Accept H_o

6.26 (b) $\chi^2_{Cal} = 10.398$, Reject H_o

6.27 (b) $\chi^2_{Cal} = 38.558$, Reject H_o

6.28 $\chi^2_{Cal} = 26.38$, Reject H_o

6.29 (b) $\chi^2_{Cal} = 44.322$, Reject H_o

6.30 (b) $\chi^2_{Cal} = 44.64$, Reject H_o

6.31 (b) $\chi^2_{Cal} = 14.64$, Reject H_o

6.32 (b) $\chi^2_{Cal} = 26.723$, Reject H_o at $\alpha = 0.05$

6.34 $\chi^2_{Cal} = 12.00$, Reject H_o

6.36 (a) $\chi^2_{Cal} = 10$, Reject H_o

CamScanner

ANSWERS TO MCQs

CamScanner

ANSWERS TO MCQs

SET – A

(i)	2^n	(ii)	1326	(iii)	$\frac{1}{2}$	(iv)	$\frac{1}{13}$
(v)	Independent	(vi)	0.6	(vii)	n, p	(viii)	np
(ix)	4	(x)	0 and 1	(xi)	Census	(xii)	125
(xiii)	4	(xiv)	negative	(xv)	–1 to 1	(xvi)	Positive
(xvii)	$(r-1)(c-1)$						

SET – B

(i)	Continuous variable	(ii)	Origin and scale	(iii)	χ^2 - test
(iv)	1	(v)	0.8	(vi)	1
(vii)	equal	(viii)	np	(ix)	8
(x)	Census	(xi)	36	(xii)	Statistic
(xiii)	regression	(xiv)	25	(xv)	Parameter
(xvi)	–1 to 1	(xvii)	Correlation		

CamScanner

SET – C

(i) Combination (ii) Tree diagram (iii) Conditional Probability

(iv) Independent (v) $\frac{1}{6}$ (vi) 1

(vii) probability distribution (viii) Positively skewed

(ix) 3p and 3pq (x) 5 (xi) Continuous distribution

(xii) from 0 to 1 (xiii) The points on the scatter diagram form a circle

(xiv) origin and scale (xv) 6 (xvi) Judgement sampling

(xvii) μ

SET – D

(i) Scatter diagram (ii) Parameter (iii) Correlation

(iv) Census (v) dependent variable (vi) Heterogenous

(vii) Chance variable (viii) 1 (ix) Regression

(x) origin and scale (xi) infinite (xii) the household

(xiii) $(AB) > \frac{(A)(B)}{N}$ (xiv) $\frac{1}{4}$ (xv) 2^n

(xvi) $(r-1)(c-1)$ (xvii) 125

CamScanner

SET – E

(i)	equal	(ii)	chance variable	(iii)	zero
(iv)	1	(v)	$p > \frac{1}{2}$	(vi)	(r – 1)(c – 1)
(vii)	all of these	(viii)	scatter diagram	(ix)	dependent variable
(x)	non-negative function	(xi)	$\frac{11}{16}$	(xii)	n and p
(xiii)	Heterogenous	(xiv)	questionnaire	(xv)	$P(\phi)$
(xvi)	Sample space	(xvii)	random number		

SET – F

(i)	sample	(ii)	parameter	(iii)	$E(\bar{x}) = \mu$
(iv)	None of these	(v)	1	(vi)	χ^2 - test
(vii)	correlation	(viii)	census	(ix)	one
(x)	regression	(xi)	$-1 \le r \le 1$	(xii)	two
(xiii)	$\chi^2_{tab} < \chi^2_{cal}$	(xiv)	$0 \le p \le 1$	(xv)	contingency table
(xvi)	discrete probability distribution	(xvii)	sampling error reduces		

CS CamScanner

SET – G

(i)	Regression	(ii)	Linear	(iii)	$-1 \le r \le 1$
(iv)	$\sum x \cdot P(x)$	(v)	$\bar{x} > \sigma^2$	(vi)	68.26%
(vii)	constant	(viii)	two	(ix)	$1-\dfrac{6\sum d^2}{n(n^2-1)}$
(x)	not on origin and scale	(xi)	$A \cap B = \phi$	(xii)	χ^2 - distribution
(xiii)	$\sqrt{np}$	(xiv)	μ	(xv)	0 and 1
(xvi)	A depends on B	(xvii)	4!		

SET – H

(i)	μ	(ii)	r^2	(iii)	origin and scale
(iv)	χ^2 - test	(v)	1	(vi)	0.8
(vii)	None of these	(viii)	equal	(ix)	np
(x)	8	(xi)	$\dfrac{7}{8}$	(xii)	Census
(xiii)	Statistic	(xiv)	1	(xv)	regression
(xvi)	125	(xvii)	from –1 to 1		

CS CamScanner