B.B.A

1st - year, IInd SEM

Business Statistics.

UNIT - 5

1. Correlation
2. Regression
3. Time - Series

Def. of Correlation :-

If the change in one variable effects a change in the other variable then the two variab. are said to be Correlated.

Ex :- 1. Let us Consider the two variables height and weight these two variables are mutually dependent. If the Height increases weight will also increases.

2. Rainfall & Crop yield.
3. Income & Expenditure.
4. Supply & demand

Types of Correlation :-

(i) Direct or positive Correlation :-

If the increase (or) decrease in one variable results the increase or decrease in the other variable then the Correlation between the two variables is said to be direct or +ve correlation.

Ex :- Height & weight, Rainfall & crop yield, Income & expenditure

(ii) Indirect or Negative Correlation :- If the increase (or decrease) in one variable results the decrease (or) Increase in the other variable, then the Correlation between the two variables is said to be indirect (or) -ve Correlation.

Ex) -

1. If we will increase the time of the worker then the efficiency of the work will decrease. i.e time and efficiency are -ve Correlated

2. volume & pressure..

(ii) un Correlated :-

If the change in one variable is independent of change in the other variable. Then such variables are said to be uncorrelated.

Ex:-

1. The age of a student and quality of a product

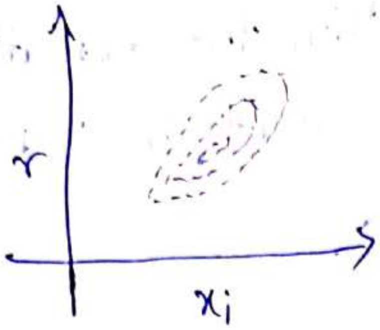
2. Knowledge of a student and taste of an ice cream

Scattered Diagram :-

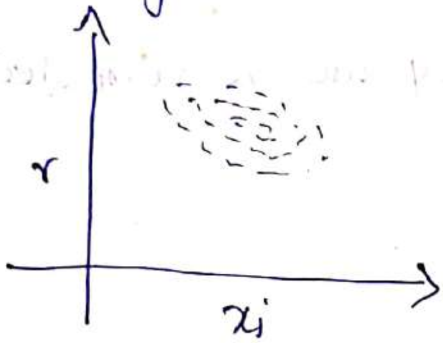
It is the simplest way of the diagramatic representation of Bivariate data. The diagramatic representation of bivariate data is known as a Scattered diagram.

1. If x & y are positively Correlated then its Scattered diagram is given by

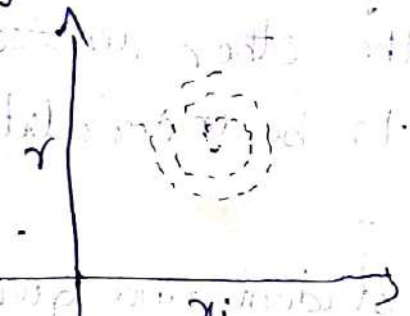
Scanned with CamScanner



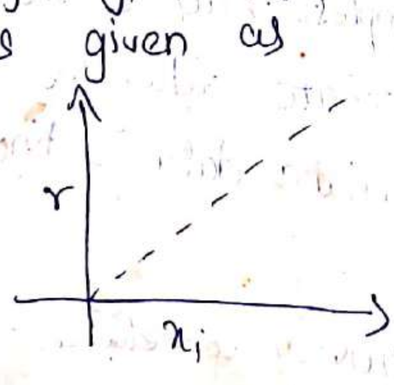
2. If x & y are negatively Correlated then its Scattered diagram as follows.



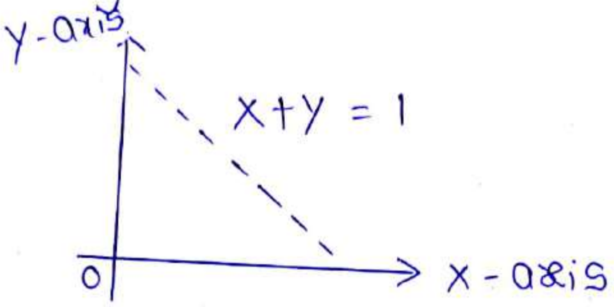
3. If x & y are uncorrelated then its diagram is



4. If x & y are in the Same direction with Constant proportion, than the two variables are Said to be perfectly +vely Correlated and its Scattered diagram is given as.



If X and Y are in the opposite direction with constant proportional then the two variables are said to be perfectly negatively correlated and it's scattered diagram is



Karlpearsons Correlation Coefficient :

Correlation coefficient is a quantitative measure of linear relationship between two variables is called coefficient of correlation it was developed by Karl Pearson's in the year 1890.

According to karl pearson's it is denoted by $r(X,Y)$ or r_{XY} or simply r. and is given by

$$r(x,y) = \frac{cov(x,y)}{\sigma_X \cdot \sigma_Y}$$

(or)

$$r(x,y) = \frac{\frac{1}{n}\Sigma XY - \bar{X}\cdot\bar{Y}}{\sqrt{\left[\frac{\Sigma X^2}{n} - (\bar{X})^2\right]\left[\frac{\Sigma Y^2}{n} - (\bar{Y})^2\right]}}$$

Problem :

Calculate Karl pearson's coefficient of correlation between expenditure on advertising and sales from the data given below.

Advertising Expenses('000 Rs.) :	39	65	62	90	82	75	25	98	36	78
sales (lakhs Rs.) :	47	53	58	86	62	68	60	91	51	84

Solution:- Let the advertising Expenses (in '000 Rs) be denoted by the variable x and the sales (in lakhs) be denoted by the variable y.

Calculations for correlation coefficient

X	Y	$d_x = X - \bar{x}$ $X - 65$	$d_y = Y - \bar{y}$ $Y - 66$	$d_x d_y$	d_x^2	d_y^2
39	47	39 − 65 = −26	47 − 66 = −19	494	676	361
65	53	65 − 65 = 0	53 − 66 = −13	0	0	169
62	58	62 − 65 = −3	58 − 66 = −8	24	9	64
90	86	25	20	500	625	400
82	62	17	−4	−68	289	16
75	68	10	2	20	100	4
25	60	−40	−6	240	1600	36
98	91	33	25	825	1089	625
36	51	−29	−15	435	841	225
78	84	13	18	234	169	324
$\Sigma X = 650$	$\Sigma Y = 660$	$\Sigma d_x = 0$	$\Sigma d_y = 0$	$\Sigma d_x d_y = 2704$	$\Sigma d_x^2 = 5398$	$\Sigma d_y^2 = 2224$

$\bar{x}$ (mean of x) $= \frac{\Sigma X}{n}$ $\bar{y}$ (mean of y) $= \frac{\Sigma Y}{n}$

$= \frac{650}{10}$ $= \frac{660}{10}$

$\bar{x} = 65$ $\bar{y} = 66$

Since mean of x and mean of Y are integers Hence the co-efficient of correlation is obtained by the following formula.

$$r(x,y) = \frac{\Sigma d_x d_y}{\sqrt{\Sigma d_x^2 \cdot \Sigma d_y^2}} = \frac{2704}{\sqrt{5398 \times 2224}}$$

$$= \frac{2704}{\sqrt{12005152}} = \frac{2704}{3464.8451} = 0.7804$$

$$\therefore r(x,y) = 0.7804.$$

Hence there Exist a positive correlation between Advertising Expenses and sales of a product.

NOTE: Since mean of x and mean of Y are integers the correlation co-efficient is obtained by the following formula.

$$r(x,y) = \frac{\Sigma d_x d_y}{\sqrt{\Sigma d_x^2 \cdot \Sigma d_y^2}}$$

NOTE 2: Since the mean of X and mean of Y are not integers Hence the co-efficient of correlation is obtained by the following formula.

$$r(U,V) = \frac{n\Sigma UV - (\Sigma U)(\Sigma V)}{\sqrt{n\Sigma U^2 - (\Sigma U)^2 \cdot n\Sigma V^2 - (\Sigma V)^2}}$$

Problem: Calculate the co-efficient of correlation for the ages of husbands and wives.

Age of Husband : 23 27 28 29 30 31 33 35 36 39

Age of wife : 18 22 23 24 25 26 28 29 30 32

Solution : Let the Age of Husband denoted by the variable X and Age of wife denoted by the variable Y.

Calculations for correlation co-efficient

X	Y	$U = X - 31$ ($X - A$)	$V = Y - 25$ ($Y - B$)	UV	U^2	V^2
23	18	23-31 = -8	18-25 = -7	56	64	49
27	22	27-31 = -4	22-25 = -3	12	16	9
28	23	-3	-2	6	9	4
29	24	-2	-1	2	4	1
30	25	-1	0	0	1	0
31	26	0	1	0	0	1
33	28	2	3	6	4	9
35	29	4	4	16	16	16
36	30	5	5	25	25	25
39	32	8	7	56	64	49
		$\Sigma U = 1$	$\Sigma V = 7$	$\Sigma UV = 179$	$\Sigma U^2 = 203$	$\Sigma V^2 = 163$

A = Asumed mean of x is 31 and

w B = Assumed mean of y is 25.

co-efficient of correlation between u and v is given by

$$r(u,v) = \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{\sqrt{[n\Sigma u^2 - (\Sigma u)^2][n\Sigma v^2 - (\Sigma v)^2]}}$$

$$= \frac{(10 \times 179) - (1 \times 7)}{\sqrt{[10 \times 203 - (1)^2][10 \times 163 - (7)^2]}}$$

$$= \frac{1790 - 7}{\sqrt{[2030 - 1][1630 - 49]}}$$

$$= \frac{1783}{\sqrt{2029 \times 1581}}$$

$$= \frac{1783}{1790.79}$$

$$r(u,v) = 0.9956$$

Hence there exist a positive correlation between Age of husband (x) and Age of wife (y)

Scanned with CamScanner

Probable Error :- If r is the observed Correlation coefficient in a sample of n pairs of observations then its standard Error usually denoted by $S.E(r)$ is given by

$$S.E(r) = \frac{1 - r^2}{\sqrt{n}}$$

Probable Error of the correlation coefficient is given by

$$P.E(r) = 0.6745 \times S.E(r)$$

$$= 0.6745 \left(\frac{1 - r^2}{\sqrt{n}} \right)$$

NOTE 1 : If $r < P.E(r)$ then correlation is not at all significant.

NOTE 2 : If $r > 6 \cdot P.E(r)$ then correlation is significant.

Problem : If $r = 0.9$ and $n = 10$ then calculate $P.E(r)$ and test the significant of coefficient of correlation.

Sol :- Given That :

$$r = 0.9 \quad \text{and} \quad n = 10.$$

$$P.E(r) = 0.6745 \left(\frac{1 - r^2}{\sqrt{n}} \right)$$

$$0.6745\left[\frac{1-(0.9)^2}{\sqrt{10}}\right]$$

$$0.6745\left[\frac{1-0.81}{3.1623}\right]$$

$$0.6745\left[\frac{0.19}{3.1623}\right]$$

$$0.6745 \times 0.0601$$

$$\boxed{\therefore P.E(r) = 0.0405}$$

Significance of r: $r = 0.9$ and

$6 \times P.E(r) = 6 \times 0.0405 = 0.2430$

$\therefore r = 0.9 > 6 \times P.E(r) = 0.2430$ Hence

Coefficient of correlation is significant.

Rank Correlation :- Rank correlation is denoted by ρ (roh) and is given by

$$\rho = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$ where,

d = Difference between Rank of x and Rank of y.

n = number of paired observations.

Problem 1: Ten Competitors in a beauty contest are ranked by three Judges in the following order

1^{st} Judge :	1	6	5	10	3	2	4	9	7	8
2^{nd} Judge :	3	5	8	4	7	10	2	1	6	9
3^{rd} Judge :	6	4	9	8	1	2	3	10	5	7

Use the rank correlation coefficient to determine which pair of Judges has the nearest approach to common tastes in beauty.

Sol :- Let R_1, R_2 and R_3 denote the ranks given by the first, second and third Judges respectively.

Calculation of Rank correlation Coefficient

R_1	R_2	R_3	$d_{12} = R_1 - R_2$	$d_{13} = R_1 - R_3$	$d_{23} = R_2 - R_3$	d_{12}^2	d_{13}^2	d_{23}^2
1	3	6	-2	-5	-3	4	25	9
6	5	4	1	2	1	1	4	1
5	8	9	-3	-4	-1	9	16	1
10	4	8	6	2	-4	36	4	16
3	7	1	-4	2	6	16	4	36
2	10	2	-8	0	8	64	0	64
4	2	3	2	1	-1	4	1	1
9	1	10	8	-1	-9	64	1	81
7	6	5	1	2	1	1	4	1
8	9	7	-1	1	2	1	1	4
			$\Sigma d_{12} = 0$	$\Sigma d_{13} = 0$	$\Sigma d_{23} = 0$	$\Sigma d_{12}^2 = 200$	$\Sigma d_{13}^2 = 60$	$\Sigma d_{23}^2 = 214$

→ Rank correlation between Judges first and second.

$$\rho_{12} = 1 - \frac{6\Sigma d_{12}^2}{n(n^2-1)} \Rightarrow 1 - \frac{6 \times 200}{10(10^2-1)}$$

$$= 1 - \frac{1200}{10 \times 99} \Rightarrow 1 - \frac{120\cancel{0}}{99\cancel{0}}$$

$$= 1 - \frac{120}{99} \Rightarrow 1 - 1.2121$$

$$\boxed{\rho_{12} = -0.2121}$$

∴ The Ranks assigned by the Judges first and second are negatively correlated.

→ Rank correlation between Judges second and third

$$\rho_{23} = 1 - \frac{6\Sigma d_{23}^2}{n(n^2-1)} \Rightarrow 1 - \frac{6 \times 214}{10(10^2-1)}$$

$$= 1 - \frac{1284}{10 \times 99} \Rightarrow 1 - \frac{1284}{990}$$

$$= 1 - 1.2970$$

$$\boxed{\rho_{23} = -0.2970}$$

∴ The Ranks assigned by the Judges second and third are negatively correlated.

→ Rank Correlation between Judges first and Third.

$$\rho_{13} = 1 - \frac{6\Sigma d_{13}^2}{n(n^2-1)} \Rightarrow 1 - \frac{6\times 60}{10(10^2-1)}$$

$$= 1 - \frac{360}{10\times 99} \Rightarrow 1 - \frac{360}{990}$$

$$= 1 - \frac{36}{99} \Rightarrow 1 - 0.3636$$

$$\boxed{\rho_{13} = 0.6364}$$

∴ The Ranks assigned by the Judges first and Third are positively correlated.

Hence the pair of first and Third Judges has the nearest approach to common taste in beauty.

When Ranks are Not Given :-

Problem : Calculate rank correlation coefficient between advertisement cost and sales from the following data.

Advertisement cost : ('000 Rs)	39	65	62	90	82	75	25	98	36	78
Sales : (lakhs Rs.)	47	53	58	86	62	68	60	91	51	84

Sol :- Let X denote the advertisement cost ('000 Rs) and Y denote the sales (lakhs Rs.)

Calculation of Rank correlation Coefficient

X	Y	Rank of X R_X	Rank of Y R_Y	$d = R_X - R_Y$	d^2
39	47	8	10	-2	4
65	53	6	8	-2	4
62	58	7	7	0	0
90	86	2	2	0	0
82	62	3	5	-2	4
75	68	5	4	1	1
25	60	10	6	4	16
98	91	1	1	0	0
36	51	9	9	0	0
78	84	4	3	1	1
				$\Sigma d = 0$	$\Sigma d^2 = 30$

Rank correlation between X and Y

$$\rho = 1 - \frac{6 \Sigma d^2}{n(n^2-1)} \Rightarrow 1 - \frac{6 \times 30}{10(10^2-1)}$$

$$\rho = 1 - \frac{180}{10 \times 99} \Rightarrow 1 - \frac{18\cancel{0}}{99\cancel{0}}$$

$$\rho = 1 - \frac{18}{99} \Rightarrow 1 - 0.1818$$

$$\boxed{\rho = 0.8182}$$

Hence There exist a positive correlation between advertisement Expenditure (X) and sales (Y).

NOTE: The highest observation is given Rank 1 and the next highest observation is given Rank 2 and so on.

Case 3: Repeated Ranks

An observation is Repeated more than once either in X-series or in Y-series or both the series in that case common Ranks are assigned to the Repeated observations.

In the Rank correlation Formula we add the factor $\frac{m(m^2-1)}{12}$ to Σd^2 where 'm' is the number of times an observation is Repeated. Now the Rank correlation coefficient become

$$\rho = 1 - \frac{6\left[\Sigma d^2 + T_x + T_y\right]}{n(n^2-1)}$$

where, T_x is the total correction Factor in X-series

T_y is the total correction Factor in Y-series.

Problem : From the following data Calculate the Rank correlation coefficient after making adjustment for tied Ranks.

X :	10	20	10	30	20	10
Y :	50	60	60	70	80	90.

Sol :- Calculation of Rank correlation

X	Y	Rank of X R_x	Rank of Y R_y	$d = R_x - R_y$	d^2
10^4	50	$\frac{4+5+6}{3} = 5$	6	$5-6=-1$	1
20^2	60^5	$\frac{2+3}{2} = 2.5$	$\frac{4+5}{2} = 4.5$	-2	4
10^5	60^4	5	4.5	0.5	0.25
30	70	1	3	-2	4
20^3	80	$\frac{2+3}{2} = 2.5$	2	0.5	0.25
10^6	90	5	1	4	16
					$\Sigma d^2 = 25.50$

X - Series :

→ The value 20 is Repeated '2' times

i.e., $m = 2$

The Common Rank given to this value is 2.5 which is the average of 2 and 3.

Now the correction Factor $= \frac{m(m^2-1)}{12} = \frac{2(2^2-1)}{12}$

$= 0.5$

→ The value 10 is Repeated '3' times

i.e., $m = 3$

The common Rank given to this value is 5 which is the average of 4, 5 and 6.

Now the correction Factor $= \frac{m(m^2-1)}{12} = \frac{3(3^2-1)}{12}$

$= 2$

Hence the Total correction Factor in X-series is $0.5 + 2 = 2.5$

$\therefore T_X = 2.5$

Y - series :

→ The Value 60 is Repeated '2' times

i.e., $m = 2$.

The common Rank given to this value is 4.5

which is the average of 4 and 5.

Now the correction factor $= \frac{m(m^2-1)}{12} = \frac{2(2^2-1)}{12}$

$$= 0.5$$

Hence the correction factor in Y - series is 0.5

$$\therefore \; T_y = 0.5$$

Rank correlation between x and y is

$$\rho = 1 - \frac{6\left[\Sigma d^2 + T_x + T_y\right]}{n(n^2-1)}$$

$$= 1 - \frac{6\left[25.5 + 2.5 + 0.5\right]}{6(6^2-1)}$$

$$= 1 - \frac{6\left[28.5\right]}{6 \times 35}$$

$$= 1 - \frac{171}{210}$$

$$= 1 - 0.8143$$

$$\boxed{\rho = 0.1857}$$

Regression

Definition :- Regression analysis is a mathematical measure of the average relationship between two or more variables. in terms of the original units of the Data.

Lines of Regression :-

Line of Regression of Y on X : Line of regression of Y on X is the line which gives the best estimate for the value of Y for any specified value of X. and the equation form is

$$Y = a + bX$$

Line of Regression of X on Y : Line of regression of X on Y is the line which gives the best estimate for the value of X for any specified value of Y and the equation form is

$$X = a + bY$$

Coefficients of Regression :-

Coefficient of regression of Y on X : The slope of the line of regression of Y on X is called the coefficient of regression of

y on x. it is denoted by b_{yx} and is given by

$$b_{yx} = \frac{r\sigma_y}{\sigma_x} = \frac{\operatorname{cov}(x,y)}{\sigma_x^2}$$

<u>Coefficient of regression of x on y</u> : The slope of the line of regression of x on y is called the coefficient of regression of x on y. it is denoted by b_{xy} and is given by

$$b_{xy} = \frac{r\sigma_x}{\sigma_y} = \frac{\operatorname{cov}(x,y)}{\sigma_y^2}$$

<u>NOTE 1 :</u> The equation of the line of regression of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$y - \bar{y} = \frac{r\sigma_y}{\sigma_x}(x - \bar{x})$$

<u>NOTE 2 :</u> The equation of the line of regression of x on y is

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - \bar{x} = \frac{r\sigma_x}{\sigma_y}(y - \bar{y})$$

Problem 1 :

1) From the data given below find :

a) The two regression coefficients.

b) The two regression equations.

c) The coefficient correlation between the marks in Economics and Statistics

d) The most likely marks in statistics when marks in Economics are

marks in Economics:	25 28 35 32 31 36 29 38 34 32
marks in statistics:	43 46 49 41 36 32 31 30 33 39

Sol: Let us denote the marks in Economics and the marks in statistics by the variable x and y respectively.

Regression Coefficients :

Regression coefficient of y on x :

$$b_{yx} = \frac{\Sigma d_x d_y}{\Sigma d_x^2}$$

Regression coefficient of x on y :

$$b_{xy} = \frac{\Sigma d_x d_y}{\Sigma d_y^2}$$

where $d_x = x - \bar{x}$ and

$d_y = y - \bar{y}$

Calculations for Regression Equations.

X	Y	$d_x = X - \bar{x} = X - 32$	$d_y = Y - \bar{Y}$ $Y - 38$	$d_x d_y$	d_x^2	d_y^2
25	43	$25-32=-7$	$43-38=5$	-35	49	25
28	46	$28-32=-4$	$46-38=8$	16	16	64
35	49	3	11	33	9	121
32	41	0	3	0	0	9
31	36	-1	-2	2	1	4
36	32	4	-6	-24	16	36
29	31	-3	-7	21	9	49
38	30	6	-8	-48	36	64
34	33	2	-5	-10	4	25
32	39	0	1	0	0	1
$\Sigma X = 320$	$\Sigma Y = 380$	$\Sigma d_x = 0$	$\Sigma d_y = 0$	$\Sigma d_x d_y = -93$	$\Sigma d_x^2 = 140$	$\Sigma d_y^2 = 398$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{32\cancel{0}}{1\cancel{0}} = 32$$

$$\bar{Y} = \frac{\Sigma Y}{n} = \frac{38\cancel{0}}{1\cancel{0}} = 38$$

a) Regression Coefficients :-

→ Regression coefficient of y on x :

$$b_{yx} = \frac{\Sigma d_x d_y}{\Sigma d_x^2} = \frac{-93}{140} = -0.6643$$

→ Regression coefficient of x on y :

$$b_{xy} = \frac{\Sigma d_x d_y}{\Sigma d_y^2} = \frac{-93}{398} = -0.2337$$

b) Regression Equations :-

→ Regression Equation of y on x :

$$y - \bar{y} = b_{yx} (x - \bar{x})$$

$$y - 38 = -0.6643 (x - 32)$$

$$y - 38 = -0.6643x + (0.6643 \times 32)$$

$$y - 38 = -0.6643x + 21.2576$$

$$y = -0.6643x + 21.2576 + 38$$

$$y = -0.6643x + 59.2576$$

∴ Regression Equation of y on x is

$$\boxed{y = 59.2576 - 0.6643x}$$

→ Regression Equation of x on y :

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - 32 = -0.2337(y - 38)$$

$$x - 32 = -0.2337y + (0.2337 \times 38)$$

$$x - 32 = -0.2337y + 8.8806$$

$$x = -0.2337y + 8.8806 + 32$$

$$x = -0.2337y + 40.8806$$

∴ Regression Equation of x on y is

$$\boxed{x = 40.8806 - 0.2337y.}$$

c) Correlation Coefficient :

we have,

$$r^2 = b_{yx} \cdot b_{xy}$$

$$= (-0.6643) \times (-0.2337)$$

$$r^2 = 0.1552$$

$$\therefore r = \pm\sqrt{0.1552}$$

$$= \pm 0.394$$

$$\boxed{\text{Hence } r = -0.394.}$$

NOTE: Since both the Regression Coefficients are negative Hence coefficient of correlation must be negative.

---

2) The data about the sales and advertisement expenditure of a firm is given below :

	Sales (in crores of Rs.)	Advertisement Expenditure (in crores of Rs.)
Means	40	6
Standard deviations	10	15

Coefficient of correlation $= r = 0.9$

a) Estimate the likely sales for a proposed advertisement expenditure of Rs. 10 crores.

b) what should be the advertisement expenditure if the firm proposes a sales target of 60 crores of rupees?

Sol:- Let us denote the sales (in crores of Rs) and the advertisement Expenditure (in crores of Rs.) X and Y respectively.

Given That :

$\bar{x} = 40 \quad \bar{y} = 6 \quad \sigma_x = 10$

$\sigma_y = 1.5$ and $r = 0.9$

a) To Estimate the sales when advertisement Expenditure of 10 crores

i.e., we have to find x when $y = 10$.

We use the Regression Equation of x on y.

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - \bar{x} = r\frac{\sigma_y}{\sigma_x}(y - \bar{y})$$

$$x - 40 = \frac{0.9 \times 10}{1.5}(y - 6)$$

$$x - 40 = \frac{9}{1.5}(y - 6)$$

$$x - 40 = 6(y - 6)$$

$$x - 40 = 6y - 36$$

$$x = 6y - 36 + 40$$

$$x = 6y + 4$$

$$x = (6 \times 10) + 4 \quad (\text{since } y = 10)$$

$$x = 60 + 4$$

$$\boxed{\therefore \quad X = 64 \text{ crores of Rs.}}$$

b) we have to estimate the advertisement expenditure (in crores of Rs.) when the sales target of 60 crores of Rs.)

i.e., when $X = 60$ then $Y = ?$

we use the Regression Equation of Y on X.

$$Y - \bar{Y} = b_{yx} (X - \bar{x})$$

$$Y - \bar{Y} = r \cdot \frac{\sigma_{\cancel{x}}}{\sigma_X} (X - \bar{X})$$

$$Y - 6 = \frac{0.9 \times 1.5}{10} (X - 40)$$

$$Y - 6 = 0.135 (X - 40)$$

$$Y - 6 = 0.135 X - (0.135 \times 40)$$

$$Y - 6 = 0.135 X - 5.4$$

$$Y = 0.135 X - 5.4 + 6$$

$$Y = 0.135 X + 0.6$$

$$Y = (0.135 \times 60) + 0.6 \quad \text{since } X = 60$$

$$Y = 8.1 + 0.6$$

$$\boxed{Y = 8.7 \text{ crores of Rs.}}$$

3) Given below the following information about advertisement Expenditure and sales:

	Advertisement Expenditure (Rs. crores)	Sales (Rs. crores)
Mean	20	120
S.D	5	25

Correlation coefficient = 0.8

a) Calculate the two Regression equations

b) Find the likely sales when advertisement Expenditure is Rs. 25 crores.

c) what should be the advertisement Expenditure if company wants to attain sales target of Rs. 150 crores.

Sol :- Let us denote the advertisement Expenditure and sales by the variables x and y respectively.

Given That

$\bar{x} = 20$, $\bar{y} = 120$, $\sigma_x = 5$

$\sigma_y = 25$ and $r = 0.8$

a) we have to calculate the two Regression Equations.

→ Regression Equation of y on x :

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$y - \bar{y} = r\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

$$y - 120 = \frac{0.8 \times 25}{5}(x - 20)$$

$$y - 120 = \frac{20}{5}(x - 20)$$

$$y - 120 = 4(x - 20)$$

$$y - 120 = 4x - (4 \times 20)$$

$$y - 120 = 4x - 80$$

$$y = 4x - 80 + 120$$

$$y = 40 + 4x \quad \text{—— (1)}$$

∴ The Regression Equation of y on x is

$$\boxed{y = 40 + 4x}$$

→ Regression Equation of x on y :

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - \bar{x} = r\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

$$x - 20 = \frac{0.8 \times 5}{25}(y - 120)$$

$$X - 20 = 0.16\,(Y - 120)$$

$$X - 20 = 0.16Y - (0.16 \times 120)$$

$$X - 20 = 0.16Y - 19.2$$

$$X = 0.16Y - 19.2 + 20$$

$$X = 0.16Y + 0.8 \quad \text{---} \quad (2)$$

∴ The Regression Equation of X on Y is

$$\boxed{X = 0.8 + 0.16\,Y}$$

b) we have to find the likely sales when advertisement expenditure is Rs. 25 crore.

∴ when $X = 25$ then $Y = ?$

~~we use~~ To get Y we use Regression Equation of Y on X.

From equation (1) we have

$$Y = 40 + 4X$$

Given That, $X = 25$

$$Y = 40 + (4 \times 25)$$

$$Y = 40 + 100$$

$$\boxed{Y = 140 \text{ (crores)}}$$

c) we have to find the advertisement Expenditure if the company wants to attain sales of Rs. 150 crore

i.e., when $y = 150$ then $x = ?$

To find 'x' we use Regression Equation of X on Y.

From equation ② we have

$$X = 0.8 + 0.16\,Y$$

Given That, $Y = 150$

$$X = 0.8 + (0.16 \times 150)$$

$$= 0.8 + 24$$

$$\boxed{X = 24.8 \text{ (~~cost~~ crore)}}$$

---

4) Given That,

$\sigma_x = 60$, $\sigma_y = 20$ and $r = 0.6$

Then Find The two Regression coefficients.

Sol :- Regression Coefficient of Y on X :

$$b_{yx} = \frac{r\,\sigma_y}{\sigma_x} = \frac{0.6 \times 2\cancel{0}}{6\cancel{0}} = \frac{1.2}{6} = 0.2$$

$$\therefore b_{yx} = 0.2$$

Scanned with CamScanner

→ Regression Coefficient of X on Y :

$$b_{XY} = \frac{r\,\sigma_X}{\sigma_Y} = \frac{0.6 \times \cancel{60}^{3}}{\cancel{20}} = 1.8$$

$$\therefore \; b_{XY} = 1.8$$

Time Series

Definition :- A time series is an arrange-ment of statistical data in accordance with its time of occurence. (or) The functional relation-ship ~~between two vario~~

$$Y = f(t)$$ where,

Y is the value of the variable under consideration at time 't'

For Example :

1. The population (Y) of a country in different years (t)
2. The number of Deaths (Y) in different months (t) of the year
3. The temperature of a place on different days of the week.
4. The sale (Y) of a departmental store in different months (t) of the year.

Fitting Of Linear TREND :-

1) Below are given the figures of production (in thousand Kilograms) of sugar factory :

years	: 1984	1985	1986	1987	1988
production	: 77	88	94	85	56

a) Fit a straight line trend

b) Calculate the Trend values for all the years also calculate the Trend value for the year 1990.

Sol :- Let the straight line equation is to be fitted to a given set of time period is

$$Y = a + bX \quad \text{—— (1)}$$

The Normal equations of above straight line are

$$\Sigma Y = Na + b\Sigma X \quad \text{—— (2)}$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 \quad \text{—— (3)}$$

Here $N = 5$ (odd) Hence we shift the origin to the middle year i.e., 1986

Let $U = X - 1986$.

Now the transformed straight line equation is

$$y = a + bU$$

Computation of straight line Trend.

Years(x)	production(y)	$U = X - 1986$	U^2	UY	Trend values
1984	77	$1984 - 1986 = -2$	4	-154	89
1985	88	$1985 - 1986 = -1$	1	-88	84.5
1986	94	0	0	0	80
1987	85	1	1	85	75.5
1988	56	2	4	112	71
	$\Sigma Y = 400$	$\Sigma U = 0$	$\Sigma U^2 = 10$	$\Sigma UY = -45$	

when $\Sigma U = 0$ then

$$a = \frac{\Sigma Y}{N} = \frac{400}{5} = 80$$

$$b = \frac{\Sigma UY}{\Sigma U^2} = \frac{-45}{10} = -4.5$$

Now substitute the estimated values of 'a' and 'b' in equation ①.

$\therefore$ The fitted straight line equation is

$$y = 80 - 4.5x \quad \text{—— (4)}$$

b) Now substitute different values of U in equation (4) we get the required Trend values.

for $U = -2$ then $Y = 80 - (4.5 \times 2) = 80 + 9 = 89$

for $U = -1$ then $Y = 80 - (4.5 \times (-1)) = 80 + 4.5$
$= 84.5$

For $U = 0$ then $Y = 80 - (4.5 \times 0) = 80$

for $U = 1$ then $Y = 80 - (4.5 \times 1) = 80 - 4.5$
$= 75.5$

for $U = 2$ then $Y = 80 - (4.5 \times 2) = 80 - 9 = 71$

Also we have to calculate the Trend value for the year 1990.

$U = x - 1986$

$U = 1990 - 1986$

$U = 4$

for $U = 4$ then $Y = 80 - (4.5 \times 4) = 80 - 18 = 62$

Hence the Trend value for the year 1990 is 62

NOTE: If $\Sigma U = 0$ to find the values of a and b we have,

$$a = \frac{\Sigma Y}{N} \quad \text{and} \quad b = \frac{\Sigma UY}{\Sigma U^2}.$$

Scanned with CamScanner

2) Below are given the figures of production of a sugar factory: [Even case problem]

Year :	1989	1990	1991	1992	1993	1994
Production :	77	88	94	85	91	98

a) Fit a straight line by the method of Least squares.

b) Calculate Trend values for all the years

c) what is the yearly increase in production.

Sol :- Let the straight line equation is to be fitted to a given set of time period is

$$Y = a + bX \quad \text{—— (1)}$$

The Normal equations for estimating a and b are

$$\Sigma Y = Na + b\Sigma X \quad \text{—— (2)}$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 \quad \text{—— (3)}$$

Here N = 6 (Even) Hence we shift the origin to the mean of two middle years

$$U = \frac{X - \text{mean of two middle years}}{\frac{1}{2} \times \text{Interval}}$$

$$U = \frac{X - \frac{1991+1992}{2}}{\frac{1}{2} \times 1}$$

$$U = \frac{X - 1991.5}{\frac{1}{2}}$$

$$U = 2(X - 1991.5)$$

Now the Transformed straight line equation is

$$Y = a + bU \quad \text{—— (4)}$$

Computation of straight line Trend

Year (X)	production (Y)	$U = 2(X - 1991.5)$	U^2	UY	Trend values
1989	77	$2(1989 - 1991.5) = -5$	25	-385	
1990	88	$2(1990 - 1991.5) = -3$	9	-264	
1991	94	$= -1$	1	-94	
1992	85	1	1	85	
1993	91	3	9	273	
1994	98	5	25	490	
	$\Sigma Y = 533$	$\Sigma U = 0$	$\Sigma U^2 = 70$	$\Sigma UY = 105$	

Scanned with CamScanner

If $\Sigma U = 0$ then

$$a = \frac{\Sigma Y}{N} = \frac{533}{6} = 88.8333$$

$$b = \frac{\Sigma UY}{\Sigma U^2} = \frac{105}{70} = 1.5$$

Now substitute the estimated values of a and b in equation ① we get the required Trend line.

Hence the fitted straight line equation to a given set of time period is [Incomplete]

$\cancel{Y = 88.8333 + 1.5X}$

b) Now we have to find the Trend values substitute different values of 'U' in equation ④ we get the required Trend values.

for $U = -5$ then $Y = 88.8333 + [1.5 \times (-5)]$
$= 88.8333 - 7.5$
$= 81.3333 \simeq 81$

for $U = -3$ then $Y = 88.8333 + [1.5 \times (-3)]$
$= 88.8333 - 4.5$
$= 84.3333 \simeq 84$

for $U = -1$ then $Y = 88.8333 + [1.5 \times (-1)]$
$= 88.8333 - 1.5$
$= 87.3333 \simeq 87$

For $U=1$ then $y = 88.8333 + (1.5 \times 1)$

$= 88.8333 + 1.5$

$= 90.3333 \simeq 90$

For $U=3$ then $y = 88.8333 + (1.5 \times 3)$

$= 88.8333 + 4.5$

$= 93.3333 \simeq 93$

For $U=5$ then $y = 88.8333 + (1.5 \times 5)$

$= 88.8333 + 7.5$

$= 96.3333 \simeq 96.$

c) we have to find the yearly increment of the production.

The yearly increment itself is the slope of the straight line equation (i.e., 'b')

Hence the yearly increment is '3'

a) Continuation of 'a'

Now substitute the estimated values of a, b and $U = 2(x - 1991.5)$ in equation ①

we get

$$y = 88.8333 + 1.5\left[2(x - 1991.5)\right]$$

$$= 88.8333 + 3(x - 1991.5)$$

$$y = 88.8333 + 3x - (3 \times 1991.5)$$

$$y = 88.8333 + 3x - 5974.5$$

$$y = 3x - 5885.6667$$

Hence the straight line equation to a given set of time period is

$$\boxed{y = 3x - 5885.6667.}$$

Conversion of yearly Trend equation into monthly Trend equation :-

$\cancel{y} \quad \cancel{=}$

Annual Trend Equation : $\quad y = a + bx$

monthly Trend Equation ; $\quad y = \dfrac{a}{12} + \dfrac{b}{144} x$

1) Convert the Annual Trend Equation into a monthly trend equation.

$$y = 480 + 36x.$$

Sol: Given That,

The Annual Trend equation is

$$y = 480 + 36x.$$

Now the monthly Trend equation is

$$y = \frac{480}{12} + \frac{36}{144} x$$

$$y = 40 + 0.25 x$$

Hence the monthly Trend equation is

$$\boxed{y = 40 + 0.25 x}$$

<u>Shifting of origin :-</u>

The Annual Trend equation is $y = a + bx$

Now we have to shift the origin to the forward of k years.

i.e., $y = a + b(x + k)$

(or)

Now we have to shift the origin to the backward of k years

i.e., $y = a + b(x - k)$

The trend equation for annual sales of a good is

$y = 102 + 36x$; with 1^{st} January 1990 as origin.

Determine the Trend equation with 1^{st} January 1992 as origin.

Sol: Since 1992 as origin. we have to shift the origin to the forward of 2 years

$$Y = 102 + 36(X+2)$$

$$= 102 + 36X + 72.$$

$$\boxed{Y = 174 + 36X}$$

$\therefore$ The Trend Equation $Y = 174 + 36X$ with 1st January 1992 as the origin.

UNIT - 4

Theory of probability

Definitions :-

Random Experiment :- An Experiment is conducted repeatedly under essentially homogenious conditions, the result is not unique but may be any one of the various possible outcomes is known as Random Experiment.

Example 1 : If a coin is tossed repeatedly, the result is not unique. we may get any of the two faces, head or tail.

Example 2 : If a die is thrown repeatedly, the result is not unique. we may get any of the six faces, 1, 2, 3, 4, 5 or 6 and so on.

Trial : Performing a Random Experiment is called a trial.

Outcome : The results of a Random Experiment is called an outcome.

Event : The outcomes of a Random Experiment is called an Event.

Example 1 : If a coin is tossed getting of a head or tail is an Event.

Example 2 : If a die is thrown getting any one of the faces 1, 2, 3, 4, 5 or 6 is an Event, or getting of an odd number or an Even number or getting a number greater than 4 or less than 3 is an Event.

Simple Event : An Event is called simple if it corresponds to a single possible outcome of a random Experiment.

Example : If a coin is tossed getting a head or tail is simple Event.

Compound or Composite Event : An Event is called composite if it corresponds to more than a single possible outcome of a random Experiment

Example : If a die is thrown getting an Even number is a composite Event.

Mutually Exclusive Events : Two or more Events are said to be mutually Exclusive if the happening of any one of them excludes the happening of all others in the same

Experiment :

Example : In toss of a coin, the events 'head' and 'tail' are mutually Exclusive because if head comes, we cannot get tail and if tail comes we cannot get head.

Equally likely Events :- The outcomes are said to be equally likely if none of the outcome is expected to occur in preference to other.

Example : In throwing of a die all the six faces 1, 2, 3, 4, 5, 6 are equally likely.

Favourable Events :- The number of outcomes of a random Experiment which entail the happening of an Event are termed as the cases favourable to the Event

Example : In a toss of two coins, the number of cases favourable to the event "Exactly one head" is 2.

i.e., HT, TH

(or)

"Exactly Two heads" is one

i.e., HH.

Independent Events : If two or more Events are said to be independent if the happening of any one of the events is not affected by happening of other events.

Example : In tossing of a coin repeatedly the event of getting a head in the first toss is independent of getting a head in second, third ~~and~~ or subsequent tosses.

Exhaustive Events or Exhaustive Cases :

The total number of possible outcomes of a random Experiment is known as Exhaustive cases.

Example : In tossing of a coin there are two Exhaustive cases Head , Tail

Example 2 : In throwing of a die there are six Exhaustive cases 1, 2, 3, 4, 5 or 6.

Example 3: If r cards are drawn from a pack of 52 cards, the Exhaustive number of cases is $^nC_r = \frac{n!}{r!(n-r)!}$

Mathematical or Classical Definition of Probability :-

If a random Experiment results in N Exhaustive, mutually exclusive and equally likely outcomes out of which m are favourable to the happening of an event A then the probability of occurence of A usually denoted by P(A) is given by

$$P(A) = \frac{\text{Favourable number of cases to A}}{\text{Exhaustive number of cases}}$$

$$P(A) = \frac{m}{N}$$

NOTE 1 : probability of any event is a number lying between 0 and 1.

i.e., $0 \leq P(A) \leq 1$

NOTE 2 : The non happening of an Event is Called Complementry Event it is denoted by $\bar{A}$ and is given by

$$P(\bar{A}) = \frac{N-m}{N} = \frac{N}{N} - \frac{m}{N} = 1 - \frac{m}{N}.$$

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

$\therefore \quad P(A) + P(\bar{A}) = 1$

NOTE 3 : If $P(A) = 0$ then A is called an impossible Event.

NOTE 4 : If $P(A) = 1$ then A is called a certain event.

Axiomatic Definition of Probability :-

Suppose 'S' is a sample space and 'A' is an event of a random experiment then the probability of occurence of an event A is defined as a real number P(A) satisfying the following axioms.

Axiom 1 : $P(A) \geq 0$.

Axiom 2 : $P(S) = 1$

Axiom 3 : If $A_1, A_2, \text{-----}, A_n$ are 'n' mutually disjoint events of S then

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

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

Scanned with CamScanner

Problem 1 : If a coin is tossed twice then find the probability of getting atleast one Head.

Sol : If a coin is tossed twice then the sample space is

$$S = \{H, T\} \times \{H, T\}$$

$$S = \{HH, HT, TH, TT\}$$

The Total number of Possible outcomes are 4.

i.e., $n(S) = 4$.

Let 'A' be the event of getting atleast one Head.

The favourable cases for occurence of an event A are HH, HT and TH

i.e., $n(A) = 3$

$$P(\text{getting atleast one Head}) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{3}{4}$$

Problem 2 : If a die is thrown twice then find the probability of the sum of Two faces is 7.

Solution : If a die is thrown twice then the

Sample space is

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

The Total number of possible outcomes are

$\{(1,1)\ (1,2)\ (1,3)\ (1,4)\ (1,5)\ (1,6)$

$(2,1)\ (2,2)\ (2,3)\ (2,4)\ (2,5)\ (2,6)$

$(3,1)\ (3,2)\ (3,3)\ (3,4)\ (3,5)\ (3,6)$

$(4,1)\ (4,2)\ (4,3)\ (4,4)\ (4,5)\ (4,6)$

$(5,1)\ (5,2)\ (5,3)\ (5,4)\ (5,5)\ (5,6)$

$(6,1)\ (6,2)\ (6,3)\ (6,4)\ (6,5)\ (6,6)\}$

$\therefore\ n(S) = 36.$

Let 'A' be the event of the sum of Two Faces is 7.

The Favourable cases for happening of an event A are

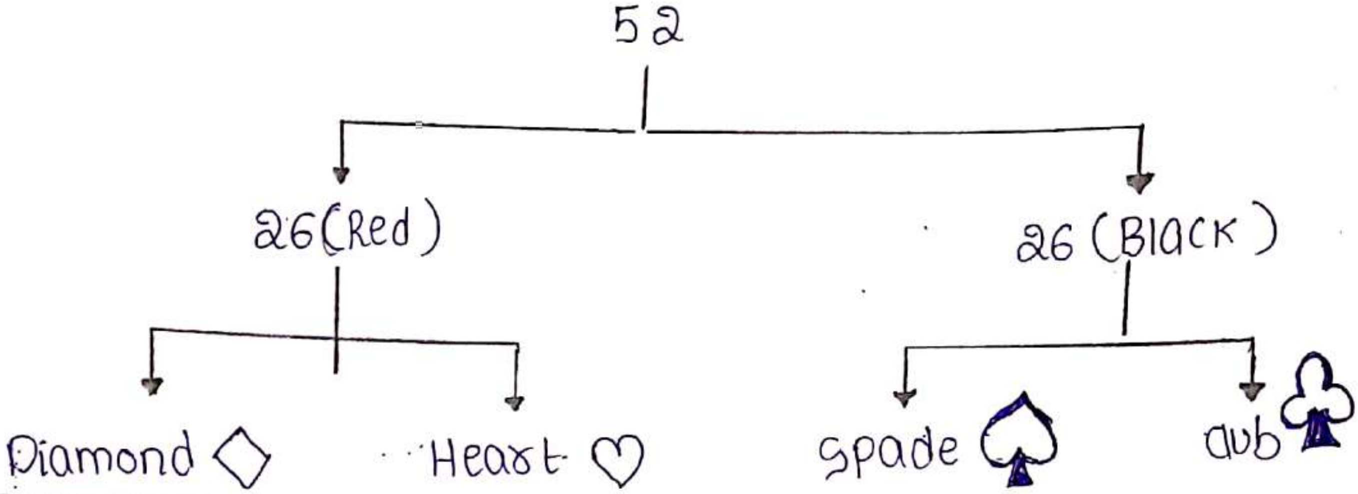
$(1,6)\ (2,5)\ (3,4)\ (4,3)\ (5,2)\ (6,1)$

$\therefore\ n(A) = 6$

$$P(A) = \frac{n(A)}{n(S)} = \frac{6}{36} = \frac{1}{6}$$

$$\therefore\ P(A) = \frac{1}{6}$$

NOTE (Structure of a pack of playing cards)



<u>Problem 3 :</u> If a card is drawn from a pack of 52 cards what is the probability that it is a red card.

<u>Sol :</u> If a card is drawn from a pack of 52 cards. This can be done in 52_{c_1} ways

$$\therefore n(S) = 52_{c_1} = 52.$$

Let A be the Event of the drawn card is Red.

one Red card is drawn out of 26 Red cards

This can be done in 26_{c_1} ways

$$\text{i.e., } n(A) = 26_{c_1} = 26$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{26}{52} = \frac{1}{2}$$

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

problem 4 : If a card is drawn from a pack of 52 cards what is the probability that it is a Diamond card.

Solution : A card is drawn from a pack of 52 cards. This can be done in 52_{c_1} ways.

$$\therefore \ n(S) = 52_{c_1} = 52.$$

Let A be the event of the drawn card is a Diamond.

One Diamond card is drawn out of 13 Diamond cards This can be done in 13_{c_1} ways

$$\therefore \ n(A) = 13_{c_1} = 13$$

$$P(A) = \frac{13}{52} = \frac{1}{4}$$

$$P(A) = \frac{1}{4}$$

Addition Theorem Probability :-

Statement : If A and B are any two events and are not disjoint defined on a sample space then

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

Problem 1 : One card is drawn at Random from a pack of 52 cards. what is the probability that it is either Ace or a Red ?

Solution: One card is drawn at random from a pack of 52 cards is 52_{c_1} ways

i.e., $n(S) = 52_{c_1} = 52$

Let us define the events are

A : The drawn card is Ace

B : The drawn card is Red

One Ace card is drawn out of 4 Ace cards is 4_{c_1} ways

i.e., $n(A) = 4_{c_1} = 4$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{4}{52}$$

$$P(A) = \frac{4}{52}$$

One Red card can be drawn out of 26 Red cards is 26_{c_1} ways

i.e., $n(B) = 26_{c_1} = 26$

$$P(B) = \frac{n(B)}{n(S)} = \frac{26}{52}$$

A a card can be drawn is Red and Ace is 2_{c_1} ways

i.e., $n(A \cap B) = 2$ and $P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{2}{52}$

Now we have to find $P(A \cup B) = ?$

Since from Addition Theorem we have

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

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52}$$

$$= \frac{4 + 26 - 2}{52}$$

$$= \frac{30 - 2}{52}$$

$$= \frac{28}{52}$$

The probability that the drawn card is either Ace or a Red card is $\frac{28}{52}$.

NOTE: If A and B are two mutually disjoint Events then Addition Theorem become

$$P(A \cup B) = P(A) + P(B) \quad / \because P(A \cap B) = 0$$

Problem 2: One card is drawn at random from a pack of 52 cards. what is the probability that it is either a King or a queen.

Solution: One card is drawn at random from a pack of 52 cards This can be done in 52_{C_1} ways.

$$\text{i.e., } n(S) = 52_{C_1} = 52$$

Let us define the events are

A : The drawn card is a King

B : The drawn card is a queen.

One King can be drawn out of 4 King cards This can be done in 4_{C_1} ways

$$\text{i.e., } n(A) = 4_{C_1} = 4$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{4}{52}$$

One queen can be drawn out of 4 queens

This can be done in 4_{C_1} ways.

i.e., $n(B) = 4_{C_1} = 4$

$$P(B) = \frac{n(B)}{n(S)} = \frac{4}{52}$$

Now we have to find $P(A \cup B) = ?$

Hence By Addition theorem.

$P(A \cup B) = P(A) + P(B)$ / $\because$ A and B are disjoint Events

$$= \frac{4}{52} + \frac{4}{52}$$

$$P(A \cup B) = \frac{8}{52}$$

probability that the drawn card is either a King or a queen is $\frac{8}{52}$.

NOTE: De - Morgan's Law.

1. $P(\overline{A \cup B}) = P(\bar{A} \cap \bar{B})$

2. $P(\overline{A \cap B}) = P(\bar{A} \cup \bar{B})$

Problem 3 : If two dice are thrown, what is the probability that the sum is neither 7 nor 11 ?

Solution : If two dice are thrown the sample space is

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

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

i.e., $n(S) = 36$.

Let us define

A be the event of the sum of Two faces is 7

B be the event of the sum of Two faces is 11

∴ The number of outcomes are favourable to an event A are (1,6) (2,5) (3,4) (4,3) (5,2) (6,1)

i.e., $n(A) = 6$

$$\therefore \quad P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{6}{36}$$

and the number of outcomes are favourable to an event B are (5,6) (6,5)

i.e., $n(B) = 2$.

$$\therefore \quad P(B) = \frac{n(B)}{n(S)}$$

$$P(B) = \frac{2}{36}$$

The number of outcomes are favourable to the sum of two faces is 7 and 11 are zero.

i.e., $n(A \cap B) = 0$

$$\therefore \quad P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{0}{36} = 0.$$

Now we have to find $P(\bar{A} \cap \bar{B}) = ?$

we have

$$\begin{aligned} P(\bar{A} \cap \bar{B}) &= P(\overline{A \cup B}) \\ &= 1 - P(A \cup B) \\ &= 1 - [P(A) + P(B) - P(A \cap B)] \\ &= 1 - \left[\frac{6}{36} + \frac{2}{36} - 0\right] \end{aligned}$$

$$P(\bar{A} \cap \bar{B}) = 1 - \frac{8}{36}$$

$$P(\bar{A} \cap \bar{B}) = \frac{28}{36}$$

Conditional probability :- suppose A and B are any two events then the conditional probability of A given B is denoted by $P(A/B)$ and it is defined as the conditional probability of occurence of an event A given that the event B has already happened and is given by

$$P(A/B) = \frac{n(A\cap B)}{n(B)} = \frac{\frac{n(A\cap B)}{n(S)}}{\frac{n(B)}{n(S)}}$$

$$P(A/B) = \frac{P(A\cap B)}{P(B)} \quad / \because P(B) > 0$$

Similarly, suppose A and B are any two events then the conditional probability of B given A is denoted by $P(B/A)$ and it is defined as the conditional probability of occurence of an event B given that the event A has already happened and is given by

$$P(B/A) = \frac{n(A\cap B)}{n(A)} = \frac{\frac{n(A\cap B)}{n(S)}}{\frac{n(A)}{n(S)}}$$

$$\therefore \ P(B/A) = \frac{P(A \cap B)}{P(A)} \ ; \ P(A) > 0$$

Multiplication Theorem for Independent Events:

Statement : Two events A and B are independent if and only if

$$P(A \cap B) = P(A) \cdot P(B)$$

Problem 5 : A problem in statistics is given to 3 students X, Y, Z whose chances of solving it are $1/2$, $3/4$ and $1/4$ respectively. What is the probability that the problem will be solved if all of them try independently.

Solution : Let us define the events

A : The problem is solved by the student X.

B : The problem is solved by the student Y

C : The problem is solved by the student Z.

$A \cup B \cup C$: The problem is solved by atleast one of the student

Given That :

$$P(A) = 1/2 \ , \quad P(B) = 3/4 \text{ and } P(C) = 1/4$$

Now we have to find $P(A \cup B \cup C) = ?$

$$P(A \cup B \cup C) = 1 - P(\overline{A \cup B \cup C})$$

$$= 1 - [P(\bar{A} \cap \bar{B} \cap \bar{C})]$$ / ∵ From Demorgan's Law

$$= 1 - [P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C})]$$ / ∵ A, B and C are Independent

$$= 1 - [\{1 - P(A)\}\{1 - P(B)\}\{1 - P(C)\}]$$

$$= 1 - [\{1 - 1/2\}\{1 - 3/4\}\{1 - 1/4\}]$$

$$= 1 - [1/2 \cdot 1/4 \cdot 3/4]$$

$$= 1 - \frac{3}{32}$$

$$= \frac{32 - 3}{32}$$

$$P(A \cup B \cup C) = \frac{29}{32}$$

The problem will be solved if all of them try independently is $\frac{29}{32}$

Problem 6 : A town has two doctors operating independently. If the probability that doctor X is available is 0.9 and that for Y is 0.8. what is the probability that atleast one doctor is available when needed.

Solution : Let

A be the event of doctor X is available in a town

B be the event of doctor Y is available in a town

Given That :

$$P(A) = 0.9 \quad \text{and} \quad P(B) = 0.8$$

Now we have to find $P(A \cup B) = ?$

By the statement of addition theorem we have

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

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

/∵ A and B are Independent

$$= 0.9 + 0.8 - (0.9 \times 0.8)$$

$$= 1.7 - 0.72$$

$$P(A \cup B) = 0.98$$

The probability that atleast one doctor is available in a town is 0.98

Bayes Theorem :-

Statement : If $E_1, E_2, ----, E_n$ are 'n' mutually disjoint events with $P(E_i) \neq 0 \ \forall$ $i = 1, 2, ---, n$ then for any arbitrary event which is a subset of union of n Events then

$$P(E_i/A) = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)} \ \forall i=1,2,---,n.$$

Problem : The contents of urns I, II and III are as follows :

urn 1 : 6 white 4 black

urn 2 : 3 white 7 black

urn 3 : 1 white 9 black

one urn is chosen at random and a ball drawn. They happen to be white. what is the probability that they came from urn III

Solution: Let

E_1 denote the event of choosing urn I

E_2 denote the event of choosing urn II and

E_3 denote the event of choosing urn III and

also A be the event that the ball drawn

from the selected urn is white.

The we write.

$P(E_1) = 1/3$, $P(E_2) = 1/3$ and $P(E_3) = 1/3$

$$P(A/E_1) = \frac{{}^{6}C_1}{{}^{10}C_1} = \frac{6}{10}$$

$$P(A/E_2) = \frac{{}^{3}C_1}{{}^{10}C_1} = \frac{3}{10}$$

$$P(A/E_3) = \frac{{}^{1}C_1}{{}^{10}C_1} = \frac{1}{10}$$

Now we have to find $P(E_3/A) = ?$

From the statement of Bayes theorem we have.

$$P(E_3/A) = \frac{P(E_3) \cdot P(A/E_3)}{\sum_{i=1}^{3} P(E_i) \cdot P(A/E_i)}$$

$$= \frac{P(E_3) \cdot P(A/E_3)}{P(E_1) \cdot P(A/E_1) + P(E_2) P(A/E_2) + P(E_3) \cdot P(A/E_3)}$$

$$= \frac{\frac{1}{3} \cdot \frac{1}{10}}{\left(\frac{1}{3} \times \frac{6}{10}\right) + \left(\frac{1}{3} \times \frac{3}{10}\right) + \left(\frac{1}{3} \times \frac{1}{10}\right)}$$

$$= \frac{\frac{1}{30}}{\frac{6}{30} + \frac{3}{30} + \frac{1}{30}}$$

$$P(E_3/A) = \frac{\frac{1}{30}}{\frac{10}{30}} = \frac{1}{10}$$

$$\therefore \quad P(E_3/A) = \frac{1}{10}.$$

Problem : A Company has two plants for manufacturing scooters. plant I manufactures 80% of the scooters and plant II manufactures 20%. At the plant I 85% scooters are rated to be of standard quality and at plant II 65% scooters are rated to be of standard quality. one scooter was selected at random. what is the probability that it is manufactured by plant II which is of standard quality

Solution ; Let us Define

E_1 be the event of scooters are manufactured by plant - I

E_2 be the event of scooters are manufactured by plant - II

A be the events of scooters are rated to be of standard quality.

Given That :

$$P(E_1) = \frac{80}{100} \quad , \quad P(E_2) = \frac{20}{100} \; ,$$

$$P(A/E_1) = \frac{85}{100} \quad \text{and} \quad P(A/E_2) = \frac{65}{100}$$

Now we have to find $P(E_2/A)$

By the statement of Bayes theorem we have

$$P(E_2/A) = \frac{P(E_2) \cdot P(A/E_2)}{\sum_{i=1}^{2} P(E_i) \cdot P(A/E_i)}$$

$$= \frac{P(E_2) \cdot P(A/E_2)}{P(E_1) \cdot P(A/E_1) + P(E_2) \cdot P(A/E_2)}$$

$$= \frac{\frac{20}{100} \times \frac{65}{100}}{\left(\frac{80}{100} \times \frac{85}{100}\right) + \left(\frac{20}{100} \times \frac{65}{100}\right)}$$

$$\frac{\frac{1300}{10000}}{\frac{6,800}{10,000} + \frac{1300}{10,000}}$$

Scanned with CamScanner

$$= \frac{\frac{1300}{10,000}}{\frac{8,100}{10,000}}$$

$$= \frac{1300}{8,100}$$

$$P(E_2/A) = \frac{13}{81}$$

<u>problem :3</u> The chances of X, Y and Z becoming managers of a certain company are 4:2:3. The probabilities that bonus schemes will be introduced if X, Y and Z become managers are 0.3, 0.5 and 0.8 respectively. if the bonus scheme has been introduced, what is the probability that X introduces it.

<u>Solution :</u> Let us Define

E_1 be the Event of X becoming manager

E_2 be the Event of Y becoming manager

E_3 be the Event of Z becoming manager

A be the Event of bonus scheme has be introduced

Given That :

$$P(E_1) = \frac{4}{4+2+3} = \frac{4}{9}$$

$$P(E_2) = \frac{2}{4+2+3} = \frac{2}{9}$$

$$P(E_3) = \frac{3}{4+2+3} = \frac{3}{9}$$

$$P(A/E_1) = 0.3 = \frac{3}{10}$$

$$P(A/E_2) = 0.5 = \frac{5}{10}$$

$$P(A/E_3) = 0.8 = \frac{8}{10}.$$

Now we have to find $P(E_1/A)$

By the statement of Bayes theorem we have

$$P(E_1/A) = \frac{P(E_1) \cdot P(A/E_1)}{\sum_{i=1}^{3} P(E_i) \cdot P(A/E_i)}$$

$$= \frac{P(E_1) \cdot P(A/E_1)}{P(E_1) \cdot P(A/E_1) + P(E_2) P(A/E_2) + P(E_3) \cdot P(A/E_3)}$$

$$= \frac{\frac{4}{9} \times \frac{3}{10}}{\left(\frac{4}{9} \times \frac{3}{10}\right) + \left(\frac{2}{9} \times \frac{5}{10}\right) + \left(\frac{3}{9} \times \frac{8}{10}\right)}$$

$$= \frac{\frac{12}{90}}{\frac{12}{90} + \frac{10}{90} + \frac{24}{90}}$$

$$= \frac{\frac{12}{\cancel{90}}}{\frac{46}{\cancel{90}}}$$

$$= \frac{12}{46}$$

$$P(E_1/A) = \frac{6}{23}$$

Scanned with CamScanner