Question No 1 State whether the following statements are true or false and also give the reason in support of your answer:

(a) If $x_1, x_2, x_3 \ldots x_n$ and $y_1, y_2, y_3 \ldots y_n$ are the variate values of two variables X and Y, and their geometric means and G_1 and G_2 respectively, then geometric mean of (x_i/y_i); $i = 1, 2 \ldots n$ will be $\left(\frac{G_1}{G_2}\right)$.

Solution (a) $\boxed{\text{The statement is TRUE}}$

If $x_1, x_2, \ldots x_n$ are the n observations of a variable 'x' then their geometric mean is

$$G_1 = (x_1 \cdot x_2 \cdot x_3 \cdots x_n)^{1/n} \quad -(1)$$

If $\frac{1}{1}, \frac{1}{2}, \frac{1}{3} \ldots \frac{1}{n}$ are the n observations of a variable 'y' then their geometric mean is

$$G_2 = (y_1 \cdot y_2 \cdot y_3 \ldots y_n)^{1/n} \quad -(2)$$

Now, we will determine geometric mean of

$$\frac{x_1}{y_1}, \frac{x_2}{y_2}, \ldots \frac{x_n}{y_n}$$

$$G = \left[\frac{x_1}{y_1}, \frac{x_2}{y_2}, \frac{x_3}{y_3} \ldots \frac{x_n}{y_n}\right]^{1/n}$$

$$G = \frac{(x_1 \cdot x_2 \cdot x_3 \ldots x_n)^{1/n}}{(y_1 \cdot y_2 \cdot y_3 \ldots y_n)^{1/n}} = \frac{G_1}{G_2}$$

$$\boxed{G = \frac{G_1}{G_2}}$$

(b) If X and Y are two independent variables and the variable $U = X + Y$ and $V = X - Y$, then the

$$r(U,V) = \frac{\sigma_X^2 - \sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}$$

Solution (b) $\boxed{\text{The Statement is TRUE}}$

Given that $U = X + Y$ then $\bar{U} = \bar{X} + \bar{Y}$

$V = X - Y$ $\quad\quad$ $\bar{V} = \bar{X} - \bar{Y}$

To check, $r(U,V) = \dfrac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2 + \sigma_y^2}$

$$\begin{aligned}
Cov(U,V) &= E(U-\bar{U})(V-\bar{V}) \\
&= E[(X+Y-\bar{X}-\bar{Y})(X-Y-\bar{X}+\bar{Y})] \\
&= E[\{(X-\bar{X})+(Y-\bar{Y})\}\{(X-\bar{X}) - (Y-\bar{Y})\}] \\
&= E[(X-\bar{X})^2 - (Y-\bar{Y})^2] \\
&= E(X-\bar{X})^2 - E(Y-\bar{Y})^2
\end{aligned}$$

$Cov(U,V) = \sigma_x^2 - \sigma_y^2$ — (3)

[$\because$ If X & Y are two independent variable then $E[(X-\bar{X})(Y-\bar{Y})] = 0$ — eqn (2)]

$$\begin{aligned}
Var(U,V) &= E[(U-\bar{U})^2] \\
&= E[X+Y-\bar{X}-\bar{Y}]^2 \\
&= E[(X-\bar{X})+(Y-\bar{Y})]^2 \\
&= E[(X-\bar{X})^2+(Y-\bar{Y})^2+2(X-\bar{X})(Y-\bar{Y})] \\
&= E(X-\bar{X})^2 + E(Y-\bar{Y})^2 + 2E(X-\bar{X})(Y-\bar{Y})
\end{aligned}$$

Using eqn (2)

$$Var(u,v) = E(x-\bar{x})^2 + E(y-\bar{y})^2$$

$$Var(u,v) = \sigma_x^2 + \sigma_y^2 \quad -(4)$$

$$r(u,v) = \frac{Cov(u,v)}{Var(u,v)}$$

from eqn (3) & (4)

$$\boxed{r(u,v) = \frac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2 + \sigma_y^2}}$$

(c) If each value of X is divided by 2 and of Y is multiplied by 2, then b'_{yx} will be same as b.

Solution:- The Statement is FALSE

Let us take an example.

X	10	30	60	80	90
Y	3	12	18	12	3

$$\text{Mean } (\bar{x}) = \frac{\Sigma x}{n} = \frac{270}{5} = 54$$

$$\text{Mean } (\bar{y}) = \frac{\Sigma y}{n} = \frac{48}{5} = 9.6.$$

$\bar{x} = 54$ is not an integer; use assumed mean $A = 60$

$\bar{y} = 9.6$ is not an integer, use assumed mean $B = 10$.

x	y	$dx = \frac{x-60}{10}$	$dy = \frac{y-10}{1}$	dx^2	dy^2	$dx \cdot dy$
10	3	-5	-7	25	49	35
30	12	-3	2	9	4	-6
60	18	0	8	0	64	0
80	12	2	2	4	4	4
90	3	3	-7	9	49	-21
$\Sigma x = 270$	$\Sigma y = 48$	$\Sigma dx = -3$	$\Sigma dy = -2$	$\Sigma dx^2 = 47$	$\Sigma dy^2 = 170$	$\Sigma dx \cdot dy = 16$

$$b_{yx} = \frac{n \Sigma dx dy - (\Sigma dx)(\Sigma dy)}{n \Sigma dx^2 - (\Sigma dx)^2}$$

$$= \frac{(5 \times 12) - (-3 \times -2)}{5 \times 47 - (-3)^2} = \frac{54}{226} = 0.2389.$$

After dividing each term by 2 in X & multiplying each term by 2 in Y.

x'	5	15	30	40	45
y'	6	24	36	24	6

$$\bar{x}_1 = \frac{\Sigma x'}{n} = \frac{135}{5} = 27 \quad \& \quad \bar{y}' = \frac{\Sigma y'}{n} = \frac{96}{5} = 19.2.$$

$\bar{x}' = 27$ is not an integer, use assumed mean $A = 27$.

$\bar{y}' = 19.2$ is not an integer, use assumed mean $B = 19$.

x'	y'	$dx' = x' - A = x - 27$	$dy' = y' - B = y - 19$	dx'^2	dy'^2	$dx'dy'$
5	6	-22	-13	484	169	286
15	24	-12	5	144	25	-60
30	36	3	17	9	289	51
40	24	13	5	169	25	65
45	6	18	-13	324	169	-234
135	96	0	1	1130	677	108

$$b'_{yx} = \frac{n\Sigma dx'd'y - (\Sigma dx')'(\Sigma dy)'}{n\Sigma dx'^2 - (\Sigma dx)^2}$$

$$= \frac{540}{5650}$$

$$= 0.0956$$

Clearly $b'_{yx} \neq b_{yx}$

(d) The mean and standard deviation of a set of values are 25 and 5 respectively. If a constant value 5 is added to each value, the coefficient of variation of the new set of value is equal to 100%.

Solution $\boxed{\text{The Statement is FALSE}}$

Suppose there are n values x_i $(i=1,2,\ldots n)$ of the variable x Mean $(\mu) = \frac{\Sigma x_i}{n} = 25$ — (1)

& standard deviation $\sigma = \sqrt{\frac{\Sigma(x_i-\mu)^2}{n}} = 5$ — (2)

Now, 5 is added to each value of x_i

So, New mean $= \mu' = \frac{\Sigma(x_i+5)}{n} = \frac{\Sigma x_i + 5n}{n} = 30$ from (1)

Now, Standard deviation

$$= \sigma' = \sqrt{\frac{\Sigma[(x_i+5)-(\mu+5)]^2}{n}} = 5 \text{ from (2)}$$

New Coefficient of variation $= \frac{\sigma'}{\mu'} \times 100$

$$= \frac{5}{30} \times 100 = 16.67\%$$

New Coefficient of variation $= 16.67\%$

(c) If $(A) = 90$, $(AB) = 40$, $N = 150$ and $(\beta) = 80$ then $(\alpha\beta) = 30$

Solution: $\boxed{\text{The Statement is TRUE}}$

Given:- $(A) = 90$ $\quad N = 150$ $\quad$ To check $(\alpha\beta) = 80$

$(AB) = 40$ $\quad (\beta) = 80$

using $N = (B) + (\beta)$

$150 = (B) + 80$

$(B) = 70$

$(\alpha\beta) = N - (A) - (B) + (AB)$

$= 150 - 90 - 70 + 40$

$= 30$

END OF QUESTION NO 1

Question No 2 (a) The numbers 3.2, 5.8, 7.9 and 4.5 have frequencies Y, (Y+2), (Y-3) and (Y+6), resp. If the arithmetic mean is 4.876. find the value of Y and write the whole series.

Solution

x	f	xf
3.2	y	3.2y
5.8	y+2	5.8y + 11.6
7.9	y-3	7.9y - 23.7
4.5	y+6	4.5y + 27
	$\Sigma f = 4x+5$	$\Sigma xf = 21.4x + 14.9$

$$\bar{x} = \frac{\Sigma fx}{\Sigma f}$$

$$4.876 = \frac{21.4x + 14.9}{4x+5}$$

$$4.876(4x+5) = 21.4x + 14.9$$

$$-1.896x = -9.48$$

$$y = 9.48/1.896$$

$$\boxed{y = 5}$$

Whole series

3.2, 3.2, 3.2, 3.2, 3.5, 5.8, 5.8, 5.8, 5.8, 5.8, 5.8, 5.8, 7.9, 7.9, 4.5, 4.5, 4.5, 4.5, 4.5, 4.5, 4.5, 4.5, 4.5, 4.5, 4.5

Question 2 (b). The following is the distribution of age (in years) of 800 workers:

Age Group	No. of workers	Cumulative frequency.
20-25	50	50
25-30	70	120
30-35	100	220
35-40	180	~~400~~ 400
40-45	150	550
45-50	120	670
50-55	70	~~740~~ 740
55-60	60	800

(i) $\sum_{i=1}^{k} f_i = N = 800 \Rightarrow \frac{N}{2} = \frac{800}{2} = 400$

$$\text{Median} = L + \frac{\left(\frac{N}{2} - C\right)}{f} \times h$$

$$= 35 + \frac{\left(\frac{800}{2} - 220\right)}{400} \times 5$$

$$\boxed{\text{Median} = 37.25}$$

where

L = lower class limit of median class = 35

N = total frequency = 800

C = cumulative frequency of the pre median class = 220

f = frequency of median class = 400

h = width of the class = 5

(ii) Calculating Quartile deviation.

$$Q_1 = L + \frac{\left(\frac{iN}{4} - C\right)}{f} \times h \qquad \frac{N}{4} = \frac{800}{4} = 200$$

For first quartile

$L = 30,\ f = 100,\ C = 70,\ h = 5$

$$Q_1 = 30 + \frac{(200 - 70)}{100} \times 5 = 36.5$$

For third quartile $\Rightarrow \frac{3N}{4} = \frac{3 \times 800}{4} = 600$

$L = 45,\ f = 120,\ C = 150,\ h = 5$

$$Q_3 = 45 + \left(\frac{600 - 150}{120}\right) \times 5 = 63.75$$

Hence Quartile deviation

$$= \frac{(Q_3 - Q_1)}{2} = 13.625$$

(iii) Calculating Coefficient of Quartile deviation.

$$= \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$= \frac{63.75 - 36.5}{63.75 + 36.5} = \frac{27.25}{100.25}$$

$$= 0.27.$$

Question No 3 (a) The value of Spearman's rank correlation coefficient of a set of non repeating values was found to be 2/3. The sum of the squares of difference between the corresponding ranks was 55. Find the number of pairs.

Solution:- Spearman's rank correlation coefficient is given by

$$\boxed{r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}}$$

$$r_s = \frac{2}{3}, \quad \sum_{i=1}^{n} d_i^2 = 55$$

$$r_s = 1 - \frac{6 \times 55}{n(n^2-1)} = \frac{2}{3}$$

$$\Rightarrow \frac{2}{3} - 1 = \frac{-6 \times 55}{n(n^2-1)}$$

$$n(n^2-1) = \frac{6 \times 55 \times 3}{2}$$

$$n^3 - n - 990 = 0$$

Put $n = 10$

$$1000 - 10 - 990 = 0$$

$\therefore n = 10$

Number of pairs = 10.

Question 3(b) Calculate Karl Pearson's coefficient of correlation between X and Y for the following data:

$N = 12, \sum X = 120, \sum Y = 130, \sum (X-8)^2 = 150,$

$\sum (Y-10)^2 = 200$ and $\sum (X-8)(Y-10) = 50.$

Solution

Karl's pearson's coefficient is given by

$$\boxed{r = \text{Cor}(X,Y) = \frac{N\sum dx\,dy - \sum dx \sum dy}{\sqrt{\{N\sum dx^2 - (\sum dx)^2\}\{N\sum dy^2 - (\sum dy)^2\}}}}$$

— eqn (1)

$\sum dx\,dy = \sum (x-8)(Y-10) = 50$ — (A)

$\sum dx^2 = \sum (x-8)^2 = 150$ — (B)

$\sum dy^2 = \sum (Y-10)^2 = 200$ — (C)

Here $A_x = 8$ and $A_y = 10$

$\sum dx = \sum (x - A_x)$: Sum of deviation from assumed mean A_x in x-series

Also, $\sum dx = \left(\frac{\sum x}{N} - A_x\right) \times N$

$\Rightarrow \sum x - N \times A_x$

$\Rightarrow 120 - (12 \times 8) = 24$ — (D)

$\Sigma dy = \Sigma(x - Ay)$: Sum of deviation from assumed mean Ay in Y-series

Also, $= \Sigma y - N \times Ay$

$\Sigma dy = 130 - (12 \times 10) = 10$ (E)

Putting value of (A), (B), (C), (D) and (E) in eqn (1)

$$r = \frac{(12 \times 50) - (24 \times 10)}{\sqrt{[(12 \times 150) - (24)^2][12 \times 200 - (10)^2]}}$$

$$r = \frac{360}{\sqrt{(1224 \times 2300)}}$$

$$r = \frac{360}{1677.85}$$

$$\boxed{r = 0.2145}$$ — Ans.

Question 4(a) The following table shows the information as:

Statistical Measures	Advertisement Expenditure (X) (Rs. Lakh)	Sales (Y) Rs. Lakh
Mean	20	100
Standard Deviation	03	12

$r(X,Y) = 0.8$ Then find

(i) the expected advertising expenditure of the company if sale is Rs. 125 lakhs

(ii) the expected sales of the company if the advertising expenditure is Rs. 32 lakhs.

Solution:- $\bar{X} = 20$ $\bar{Y} = 100$

$\sigma_x = 3$ $\sigma_y = 12$

(a) Line of regression of Y on X is given by:

$$\boxed{y - \bar{y} = b_{yx}(x - \bar{x})} \quad \text{— eqn (1)}$$

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = 0.8 \times \frac{12}{3} = 3.2$$

$$y - 100 = 3.2(x - 20)$$

i.e $y = 3.2x - 36$ — eqn (2)

And line of regression of X on y is

$$\boxed{x - \bar{x} = b_{xy}(y - \bar{y})} \text{ - eqn (3)}$$

Here $b_{xy} = \frac{r\sigma_x}{\sigma_y}$

$= 0.8 \times \frac{3}{12} = 0.2$

$x - 10 = 0.2(y - 90)$

$\boxed{x = 0.2y - 8}$

(ii) The expected advertising expenditure of the company if the sale is Rs. 125 lakhs.

x when y = 125 lakhs

$x = 0.2y - 8$

$x = 0.2 \times 125 - 8$

$\boxed{x = 17 \text{ lakhs}}$

(iii) The expected sales of the company if the advertising expenditure is Rs. 32 lakhs

Y when x = 32

$y = 3.2 \times 32 - 36$

$\boxed{Y = 66.4 \text{ lakhs}}$

(b) Given the following data

$r_{12} = 0.8$

$r_{13} = 0.6$

$r_{23} = 0.4$

then find

(i) $r_{12.3}$

we have
$$\boxed{r_{12.3} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}}$$

$$r_{12.3} = \frac{0.8 - (0.6 \times 0.4)}{\sqrt{(1-(0.6)^2)(1-(0.4)^2)}} = \underline{\underline{-0.76}}$$

(ii) $r_{13.2}$

we have
$$\boxed{r_{13.2} = \frac{r_{13} - r_{12} r_{32}}{\sqrt{(1-r_{12}^2)(1-r_{23}^2)}}}$$

$$r_{13.2} = \frac{0.6 - (0.8 \times 0.4)}{\sqrt{(1-0.8^2)(1-0.4^2)}} = 0.38$$

(iii) $r_{23.1}$
$$\boxed{r_{23.1} = \frac{r_{23} - r_{12} r_{13}}{\sqrt{(1-r_{12}^2)(1-r_{13}^2)}}}$$

$$r_{23.1} = \frac{0.4 - (0.8 \times 0.6)}{\sqrt{(1-0.8^2)(1-0.4^2)}} = -0.109$$

(iv) $R_{1.23}$

we have $\boxed{R_{1.23} = \sqrt{1-(1-r_{12}^2)(1-r_{13.2}^2)}}$

$$R_{1.23} = \sqrt{1-(1-(0.8)^2)(1-0.38^2)}$$

$$= \sqrt{1-0.36 \times 0.856}$$

$$\boxed{R_{1.23} = 0.692}$$

Question No 5(a) An investigation of 23713 households was made in an urban and rural mixed locality. Of these 1618 were farmers, 2015 well to do and 770 families were having at least one graduate. Of these graduate families 335 were those of farmers and 428 were well to do, also 587 well to do families were those of farmers and out of them only 156 were having at least one graduate.

Solution Let A = Farmers

B = well to do

C = Having at least one graduate.

To find ABC, $AB\gamma$, $A\beta C$, $A\beta\gamma$, αBC, $\alpha B\gamma$, $\alpha\beta C$, $\alpha\beta\gamma$

We have $A = 1618$ $\quad AB = 587$

$B = 2015$ $\quad AC = 335$

$C = 770$ $\quad BC = 428$

We know that

$$\boxed{(AB) = ABC + (AB\gamma)}$$

$\Rightarrow (AB\gamma) = (AB) - (ABC)$

$AB\gamma = 587 - 156 = 431$

$$\boxed{AB\gamma = 431}$$

$$\boxed{(AC) = ABC + (A\beta C}$$

$A\beta C = (AC) - (ABC) = 335 - 156$

$$\boxed{A\beta C = 176}$$

$$\boxed{(A\beta) = (A\beta C) + (A\beta\gamma)}$$

$(A\beta\gamma) = (A\beta) - (A\beta C)$

$= (A) - (AB) - (AC) + (ABC)$

$= 1618 - 587 - 335 + 156$

$$\boxed{A\beta\gamma = 852}$$

$$\boxed{(\alpha B) = (\alpha BC) + (\alpha B\gamma)}$$

$\alpha BC = (BC) - (ABC)$

$= 428 - 156 = 272.$

$$\boxed{\alpha BC = 272}$$

$(\alpha B) = (\alpha BC) + (\alpha B\gamma)$

$(\alpha B\gamma) = (\alpha B) - (\alpha BC)$

$= (B) - (AB) - (BC) + (ABC)$

$= 2015 - 587 - 428 + 156 = 1156$

$$\boxed{\alpha\beta c = (C) - (AC) - (BC) + (ABC)}$$

$$= 770 - 335 - 428 + 156$$

$$\boxed{\alpha\beta c = 163}$$

$$\boxed{\alpha\beta\gamma = N - (A) - (B) - (C) + (AB) + (BC) + (AC)}$$

$$= 23713 - 1618 - 2015 - 770 + 587 + 335 + 428 - 156$$

$$\boxed{\alpha\beta\gamma = 20504}$$

<u>Question No 5. (b)</u> Can vaccination be regarded as a preventive measure for smallpox from the given data:

(i) Of 1482 persons in a locality exposed to a smallpox, 368 in all were attacked.

(ii) Of 1482 persons, 343 had been vaccinated and of these only 35 were attacked.

<u>Solution</u> Let A denotes attribute of vaccination and B that of attack.

Then the given data

$N = 1482$, $(B) = 368$, $(A) = 343$, $(AB) = 35$

Now,

$$(\alpha\beta) = N - (A) - (B) + (AB)$$

$$= 1482 - 343 - 368 + 35 = 806$$

$$\boxed{\alpha\beta = 806}$$

$$A\beta = (A) - AB) = 343 - 35$$

$$(A\beta) = 308$$

$$\alpha B = (B) - (AB) = 368 - 35$$

$$\boxed{\alpha B = 333}$$

Yule's coefficient of association

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) - (A\beta)(\alpha B)}$$

$$= \frac{(35 \times 806) - (308 \times 333)}{(35 \times 806) + (308 \times 333)}$$

$$= \frac{-74354}{130774} = -0.57$$

Thus, we find that vaccination and small pox are in a high degree of negative association. Hence vaccination can be regarded as a preventive measure of small pox.

Question 6. (a) In a statistical study relating to the prices (in ₹) of two shares, X and Y. The following two regression lines were found as $8X-10Y+70=0$ and $20X-9Y-65=0$. The standard deviation of $X=3$, then find (i) the values of X and Y (ii) r (X, Y) (iii) Standard deviation of Y.

Solution

Given:- $8X-10Y+70=0$

$20X-9Y-65=0$

(i) the value of X and Y

$8X-10Y+70=0$ — (1)

multiply eqn (1) by 5

$40X-50Y+350=0$ — (2)

Now, $20X-9Y-65=0$ — (3)

multiply eqn (3) by 2

$40X-18Y-130=0$ — (4)

from eqn (2) and (4)

$$\boxed{\begin{array}{l} X=10 \\ Y=15 \end{array}}$$

(ii) Let the equation $8X-10Y+70=0$ be the regression equation of Y on X. then

$$Y=0.8X+7$$

Comparing it with

$$Y=b_{YX}X+a, \text{ we get } b_{YX}=0.8$$

Let the equation $20X-9Y-65=0$ be the regression equation of X on Y then

$$X=0.45X+3.25$$

Comparing it with $X=b_{XY}Y+a'$, we get $b_{XY}=0.45$

$$r=\pm\sqrt{b_{YX}\cdot b_{XY}}$$

Since b_{XY}, b_{YX} both are positive, then r is also positive

$$r=\sqrt{0.8\times 0.45}=0.6$$

(iii)

$$b_{YX}=r\frac{\sigma_Y}{\sigma_X}$$

$$\sigma_Y=b_{YX}\frac{\sigma_X}{r}$$

$$\sigma_Y=0.8\left(\frac{3}{0.6}\right)$$

$$\boxed{\sigma_Y=4}$$

Question No 6 (b) Suppose X and Y are the two variables having the correlation coefficient 0.85. If two new variables x' and y' are obtained by adding 50 to each value of X and 100 to each value of Y resp, Calculate the correlation coefficient between x' and y' using the above data. Also, compare the results. The following are the value they have.

Solution

INITIAL TABLE

X	Y
10	40
30	30
50	70
60	80

$x' = X + 50$

$y' = Y + 100$

FINAL TABLE

x'	y'
60	140
80	130
100	170
110	180

Calculating correlation coefficient of new value using Karl pearson's coefficient

$$r = Corr(X, Y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left\{\sum (x_i - \bar{x})^2\right\}\left\{\sum (y_i - \bar{y})^2\right.}}$$

$\sum x' = 350 \quad \bar{x'} = 87.5$

$\sum y' = 620 \quad \bar{y'} = 155$

x'	y'	$(x'-\bar{x})$	$(x'-\bar{x})^2$	$(y'-\bar{y})$	$(y'-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
60	140	-27.5	756.25	-15	225	412.5
80	130	-7.5	56.25	-25	625	187.5
100	170	12.5	156.25	15	225	187.5
110	180	22.5	506.25	25	625	562.5
$\Sigma x'$ = 350	$\Sigma y'$ = 620	$\Sigma(x'-\bar{x})$ = 0	= 1499.75		= 1700	= 1350

$$r = \text{corr}(x', y') = \frac{1350}{\sqrt{1499.75 \times 1700}}$$

$$\boxed{r = 0.85}$$

$r_{initial} = r_{new} = 0.85$

This satisfy the Second property of correlation coefficient i.e Correlation coefficient is independent of change of origin and scale.

Question No 7 (a) 50% of item have characteristics A and B both, 35% have A, but not B, 25% have B but not A. Show that there must be some misprints in the report.

Solution

Since values are in percentage therefore

$N = 100$

A and B both $(AB) = 50$

A not B $(A\beta) = 35$

B not A $(\alpha B) = 25$

Conditions for consistency of the data.

$$\boxed{(AB) \geq (A) + (B) - N} \quad - eq^n (1)$$

~~50 ≥ 35+25-100~~

$A = (AB) + (A\beta)$

$= 50 + 35 = 85$

$B = (AB) + (\alpha B)$

$= 50 + 25 = 75$

using eq^n (1)

$50 \geq 85 + 75 - 100$

$50 \geq 60$

which is not true.

Therefore, there is misprint or mistake of any sort in the report.

Question 7 (b) In the given data, two frequencies are missing and its mean is found to be 1.46. Find the missing frequencies.

Solution

No. of Accidents (x)	Frequencies (f)	$x_i f_i$
0	46	0
1	? = x	x
2	? = y	2y
3	25	75
4	10	40
5	5	25
Total	200	

Let the first missing value be x

Second missing value be y

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

$$1.46 = \frac{x + 2y + 75 + 40 + 25}{200} = \frac{140 + 2y + x}{200}$$

$$\boxed{x + 2y = 152} \quad \text{- eqn (1)}$$

Also, $46 + x + y + 25 + 10 + 5 = 200$

$$\boxed{x + y = 114} \quad \text{- eqn (2)}$$

using eqn (1) and (2)

$\boxed{y = 38}$ → Ans

$\boxed{x = 76}$ → Ans.