Chapter 10

SIMPLE REGRESSION AND CORRELATION

10.6. (a) The necessary calculations for determining the equation of the least squares regression line are shown below:

X	Y	X^2	Y^2	XY
20	5	400	25	100
11	15	121	225	165
15	14	225	196	210
10	17	100	289	170
17	8	289	64	136
19	9	361	81	171
92	68	1496	880	952

The estimated linear regression line of Y on X is

$$\hat{Y} = a + bX,$$

where a and b are the least squares estimates of the parameters α and β respectively and are given by

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

Substituting the sums, we get

$$b = \frac{(6)(952) - (92)(68)}{(6)(1496) - (92)^2}$$

$$= \frac{5712 - 6256}{8976 - 8464} = \frac{-544}{512} = -1.0625$$

$$a = \frac{68}{6} - (-1.0625)\left(\frac{92}{6}\right) = 11.3333 + 16.2917 = 27.625$$

Thus the estimated regression line is

$$\hat{Y} = 27.625 - 1.0625X.$$

(b) The predicted values of Y are found by substituting the X values in the estimated equation. Thus for $X = 10$,

$$\hat{Y} = 27.625 - 1.0625\,(10) = 27.625 - 10.625 = 17.00$$

Similarly, the predicted values of Y for $X = 11$, 15, 17, 19 and 20 are 15.94, 11.69, 9.56, 7.44, 6.38 respectively.

(c) The standard error of estimate is given by

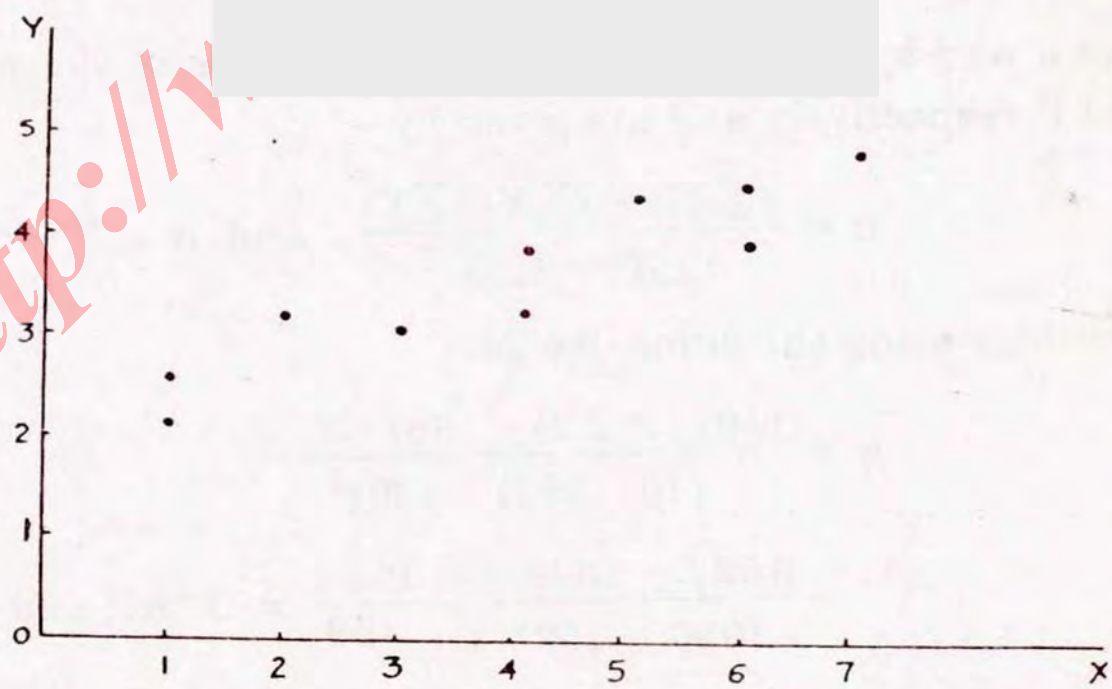
$$s_{Y.X} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n-2}}$$

Now $$\sum(Y-\hat{Y})^2 = (5-6.38)^2 + (15-15.94)^2 + (14-11.69)^2 + (17-17.00)^2 + (8-9.56)^2 + (9-7.44)^2$$

$$= (-1.38)^2 + (-0.94)^2 + (2.31)^2 + 0 + (-1.56)^2 + (1.56)^2 = 12.9913$$

Hence $$s_{Y.X} = \sqrt{\frac{12.9913}{4}} = \sqrt{3.2478} = 1.80$$

10.7. (a) The scatter diagram of the given data appears below:



(b) The necessary computations for the least-squares regression line and residuals are given in the following table:

X_i	Y_i	X_i^2	X_iY_i	$\hat{Y}_i = 2.00 + 0.387X_i$	Residuals $e_i = Y_i - \hat{Y}_i$
1	2.1	1	2.1	2.387	−0.287
1	2.5	1	2.5	2.387	0.113
2	3.1	4	6.2	2.774	0.326
3	3.0	9	9.0	3.161	−0.161
4	3.8	16	15.2	3.548	0.252
4	3.2	16	12.8	3.548	−0.348
5	4.3	25	21.5	3.935	0.365
6	3.9	36	23.4	4.322	−0.422
6	4.4	36	26.4	4.322	0.078
7	4.8	49	33.6	4.709	0.091
39	35.1	193	152.7	35.093	0.007

The estimated linear regression line of Y on X is

$$\hat{Y} = a + bX,$$

where a and b are the least-squares estimates of the parameters α and β respectively and are given by

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

Substituting the sums, we get

$$b = \frac{(10)(152.7) - (39)(35.1)}{(10)(193) - (39)^2}$$

$$= \frac{1527 - 1368.9}{1930 - 1521} = \frac{158.1}{409} = 0.387, \text{ and}$$

$$a = 3.51 - (0.387)(3.9)$$

$$= 3.51 - 1.5093 = 2.00$$

Thus the estimated regression line is

$$\hat{Y} = 2.00 + 0.387X$$

(c) To compute the residuals, we first calculate the estimated values $\hat{Y}$ which appear in fifth column of the above table. The residuals are given in the last column and they add to 0.007. Theoretically the sum is zero and this small difference is due to rounding.

(d) When $X = 10$, the predicted value of Y is

$$\hat{Y} = 2.00 + 0.387\,(10) = 5.87$$

10.8. The estimated regression equation is $\hat{Y} = a + bX$, where

$$b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}, \text{ and}$$

$$a = \bar{Y} - b\bar{X}\,.$$

(a) Now $\quad b = \dfrac{1000 - (10)\,(10)\,(20)}{2{,}000 - (10)\,(10)^2} = \dfrac{-1000}{1000} = -1$, and

$$a = 20 - (-1)\,(10) = 30.$$

Hence the desired estimated regression equation is

$$\hat{Y} = 30 - X.$$

(b) Here $\quad b = \dfrac{(32)\,(193640) - (528)\,(11720)}{(32)\,(11440) - (528)^2}$

$$= \frac{6196480 - 6188160}{366080 - 278784} = \frac{8320}{87296} = 0.095, \text{ and}$$

$$a = \frac{11720}{32} - (0.095)\left(\frac{528}{32}\right)$$

$$= 366.25 - (0.095)\,(16.5) = 364.68$$

Therefore the desired estimated regression equation is

$$\hat{Y} = 364.68 + 0.095X.$$

(c) Here $\quad b = \dfrac{(100)\,(1613) - (1239)\,(79)}{(100)\,(17322) - (1239)^2}$

$$= \frac{161300 - 97881}{1732200 - 1535121} = \frac{63419}{197079} = 0.32, \text{ and}$$

$$a = \bar{Y} - b\bar{X} = \frac{79}{100} - (0.32)\frac{1239}{100}$$

$$= 0.79 - 3.96 = -3.17$$

Here the desired estimated regression equation is

$$\bar{Y} = -3.17 + 0.32X$$

(d) Here $b = \dfrac{(10)(130628) - (1710)(760)}{(10)(293162) - (1710)^2}$

$$= \frac{1306280 - 1299600}{2931620 - 2924100} = \frac{6686}{7523} = 0.8887, \text{ and}$$

$$a = 76 - (0.8887)(171) = 76 - 151.97 = -75.97$$

Thus the desired estimated regression line is

$$\hat{Y} = -75.97 + 0.89X$$

(e) Here $b = \dfrac{\sum(X-\bar{X})(Y-\bar{Y})}{\sum(X-\bar{X})^2} = \dfrac{9871}{2800} = 3.53$, and

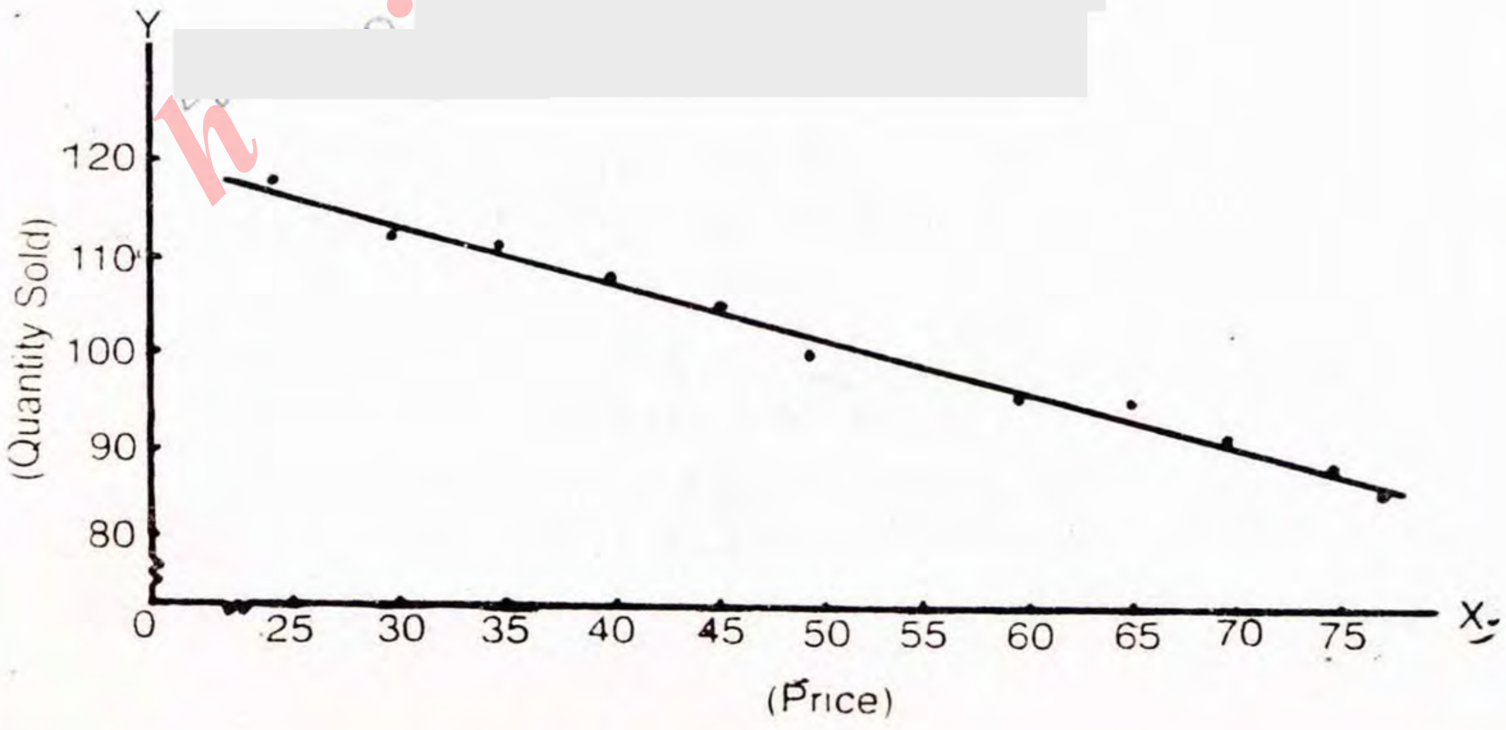
$$a = \bar{Y} - b\bar{X} = 237 - (3.53)(52)$$

$$= 237 - 183.56 = 53.44$$

Hence the desired estimated regression equation is

$$\hat{Y} = 53.44 + 3.53X.$$

10.9. (a) The scatter diagram of the given data appears below:



(b) The equation for the estimated regression line is

$$\hat{Y} = a + bX,$$

where a and b, the least-squares estimates of the parameters α and β, are given as

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

The necessary computations are given in the table below:

Price (X)	Quantity (Y)	X^2	XY	$\hat{Y} = 130.62-0.57X$	$Y_i-\hat{Y}_i$	$(Y_i-\hat{Y}_i)^2$
25	118	625	2950	116.37	1.63	2.6569
45	105	2025	4725	104.97	0.03	0.0009
30	112	900	3360	113.52	−1.52	2.3104
50	100	2500	5000	102.12	−2.12	4.4944
35	111	1225	3885	110.67	0.33	0.1089
40	108	1600	4320	107.82	0.18	0.0324
65	95	4225	6175	93.57	1.43	2.0449
75	88	5625	6600	87.87	0.13	0.0169
70	91	4900	6370	90.72	0.28	0.0784
60	96	3600	5760	96.42	−0.42	0.1764
495	1024	27225	49145	1024.05	−0.05	11.9374

Substituting the sums, we get

$$b = \frac{(10)(49145) - (495)(1024)}{(10)(27225) - (495)^2} = \frac{491450-506880}{272250-245025}$$

$$= \frac{-15430}{27225} = -0.57, \text{ and}$$

$$a = \frac{1024}{10} - (-0.57)\left(\frac{495}{10}\right) = 102.4 + 28.215 = 130.62.$$

Hence the equation of the desired regression line is

$$\hat{Y} = 130.62 - 0.57X.$$

The estimated regression line is shown on the scatterdiagram.

(c) The standard deviation of regression (standard error of estimate) is given by

$$s_{y.x} = \sqrt{\frac{\sum(Y_i - \hat{Y}_i)^2}{n-2}} = \sqrt{\frac{11.9374}{8}} = \sqrt{1.4922} = 1.22.$$

10.10. The necessary computations are given in the following table:

X	Y	XY	X^2	Y^2
3.2	6.5	20.80	10.24	42.25
2.7	5.3	14.31	7.29	28.09
4.5	8.6	38.70	20.25	73.96
1.0	1.2	1.20	1.00	1.44
2.0	4.2	8.40	4.00	17.64
1.7	2.9	4.93	2.89	8.41
0.6	1.1	0.66	0.36	1.21
1.9	3.0	5.70	3.61	9.00
17.6	32.8	94.70	49.64	182.00

(a) The estimated least-squares regression equation for Y values on X values is

$$\hat{Y} = a + bX, \text{ where}$$

$$b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \frac{(8)(94.70) - (17.6)(32.8)}{(8)(49.64) - (17.6)^2}$$

$$= \frac{757.60 - 577.28}{397.12 - 309.76} = \frac{180.32}{87.36} = 2.064, \text{ and}$$

$$a = \bar{Y} - b\bar{X} = \frac{32.8}{8} - (2.064)\left(\frac{17.6}{8}\right)$$

$$= 4.1 - 4.5408 = -0.441$$

Hence the desired regression equation is $\hat{Y} = -0.441 + 2.064X$.

(b) The standard error of estimate, $s_{y.x}$ is

$$s_{y.x} = \sqrt{\frac{\sum Y^2 - a\sum Y - b\sum XY}{n-2}}$$

$$= \sqrt{\frac{182.00-(-0.441)(32.8)-(2.064)(94.70)}{8-2}}$$

$$= \sqrt{\frac{182.00 + 14.4648 - 195.4608}{6}}$$

$$= \sqrt{\frac{1.004}{6}} = \sqrt{0.1673} = 0.41.$$

(c) The estimated least-squares regression equation for X values on Y values is

$$\hat{X} = a_0 + b_0 Y, \text{ where}$$

$$b_0 = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum Y^2 - (\sum Y)^2} = \frac{(8)(94.70) - (17.6)(32.8)}{(8)(182.00) - (32.8)^2}$$

$$= \frac{180.32}{380.16} = 0.474, \text{ and}$$

$$a_0 = \bar{X} - b_0\bar{Y} = 2.2 - (0.474)(4.1) = 0.257$$

Thus the desired least-squares regression equation is

$$\hat{X} = 0.257 + 0.474Y.$$

(d) The standard error of estimate, $s_{x.y}$ is

$$s_{x.y} = \sqrt{\frac{\sum X^2 - a_0\sum X - b_0\sum XY}{n-2}}$$

$$= \sqrt{\frac{49.64 - (0.257)(17.6) - (0.474)(94.70)}{8-2}}$$

$$= \sqrt{\frac{49.64 - 4.5232 - 44.8878}{6}}$$

$$= \sqrt{\frac{0.2290}{6}} = \sqrt{0.0382} = 0.20.$$

10.11. (b) The computation of the co-efficient of determination.

Income (X) (000)	Expenditure (Y) (000)	X^2	Y^2	XY
10	7	100	49	70
20	21	400	441	420
30	23	900	529	690
40	34	1600	1156	1360
50	36	2500	1296	1800
60	53	3600	2809	3180
210	174	9100	6280	7520

The co-efficient of determination, r^2, is given by the ratio

$$r^2 = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{\sum(\hat{Y}-\bar{Y})^2}{\sum(Y-\bar{Y})^2} = 1-\frac{\sum(Y-\hat{Y})^2}{\sum(Y-\bar{Y})^2}$$

Now $\sum(Y-\hat{Y})^2 = \sum Y^2 - a\sum Y - b\sum XY$,

where $b = \dfrac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} = \dfrac{(6)(7520)-(210)(174)}{(6)(9100)-(210)^2}$

$$= \frac{45120-36540}{54600-44100} = \frac{8580}{10500} = 0.8171, \text{ and}$$

$$a = \frac{174}{6} - (0.8171)\left(\frac{210}{6}\right) = 29 - 28.5985 = 0.4015$$

$$\therefore \quad \sum(Y-\hat{Y})^2 = 6280 - (0.4015)(174) - (0.8171)(7520)$$

$$= 6280 - 69.861 - 6144.592 = 65.547, \text{ and}$$

$$\sum(Y-\bar{Y})^2 = \sum Y^2 - (\sum Y)^2/n = 6280 - (174)^2/6$$

$$= 6280 - 5046 = 1234$$

Hence $r^2 = 1 - \dfrac{65.547}{1234} = 1 - 0.053 = 0.947.$

Alternative Method:

$$r^2 = \left[\frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}\right]^2$$

$$= \left[\frac{(6)(7520) - (210)(174)}{\sqrt{[(6)(9100) - (210)^2][(6)(6280) - (174)^2]}}\right]^2$$

$$= \left[\frac{8580}{\sqrt{(10500)(7404)}}\right]^2 = \left[\frac{8580}{8817.1423}\right]^2 = (0.9731)^2 = 0.947$$

This means that 94.7% of the variation in expenditure (Y) is related to the variation in income (X).

10.12. (b) For the data given in question 10.11 (b), we have

(i) Total variation $= \sum(Y-\bar{Y})^2 = \sum Y^2 - \frac{(\sum Y)^2}{n}$

$$= 6280 - \frac{(174)^2}{6} = 1234,$$

(ii) Unexplaiend variation $= \sum(Y-\hat{Y})^2 = \sum Y^2 - a\sum Y - b\sum XY$

$$= 6280-(0.4015)(174)-(0.8171)(7520)$$

$$= 6280-69.861-6144.592 = 65.547, \text{and}$$

(iii) Explaiend variation $=$ Total variation $-$ Unexplained variation

$$= 1234 - 65.547 = 1168.453.$$

10.14. (b) Calculation of the Co-efficient of Correlation.

X	Y	XY	X^2	Y^2
2	18	36	4	324
4	12	48	16	144
5	10	50	25	100
6	8	48	36	64
8	7	56	64	49
11	5	55	121	25
36	60	293	266	706

$$r_{XY} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left\{\sum X^2 - \frac{(\sum X)^2}{n}\right\}\left\{\sum Y^2 - \frac{(\sum Y)^2}{n}\right\}}}$$

$$= \frac{293 - \frac{(36)(60)}{6}}{\sqrt{\left\{266 - \frac{(36)^2}{6}\right\}\left\{706 - \frac{(60)^2}{6}\right\}}}$$

$$= \frac{-67}{\sqrt{(50)(106)}} = \frac{-67}{72.80} = -0.92$$

(c) Computation of the correlation co-efficient, when each X value is multiplied by 2 and 6 is added to each product, and when each Y value is multiplied by 3 and 15 is subtracted from each product, *i.e.* when $u = 2X+6$ and $v = 3Y-15$.

u	v	uv	u^2	v^2
10	39	390	100	1521
14	21	294	196	441
16	15	240	256	225
18	9	162	324	81
22	6	132	484	36
28	0	0	784	0
108	90	1218	2144	2304

$$r_{uv} = \frac{\sum uv - \frac{(\sum u)(\sum v)}{n}}{\sqrt{\left\{\sum u^2 - \frac{(\sum u)^2}{n}\right\}\left\{\sum v^2 - \frac{(\sum v)^2}{n}\right\}}}$$

$$= \frac{1218 - \frac{(108)(90)}{6}}{\sqrt{\left\{2144 - \frac{(108)^2}{6}\right\}\left\{2304 - \frac{(90)^2}{6}\right\}}}$$

$$= \frac{-402}{\sqrt{(200)(954)}} = \frac{-402}{436.81} = -0.92$$

We obtain the same result as in part (b), because the correlation co-efficient is independent of the origin and scale.

10.15. (b) Computation of the correlation coefficient.

X	Y	$u=10X-70$	$v=10Y-60$	u^2	v^2	uv
8.2	8.7	12	27	144	729	324
9.6	9.6	26	36	676	1296	936
7.0	6.9	0	9	0	81	0
9.4	8.5	24	25	576	625	600
10.9	11.3	39	53	1521	2809	2067
7.1	7.6	1	16	1	256	16
9.0	9.2	20	32	400	1024	640
6.6	6.3	−4	3	16	9	−12
8.4	8.4	14	24	196	576	336
10.5	12.4	35	64	1225	4096	2240
Total	--	167	289	4755	11501	7147

$$\therefore \quad r = \frac{\sum uv - \dfrac{(\sum u)(\sum v)}{n}}{\sqrt{\left\{\sum u^2 - \dfrac{(\sum u)^2}{n}\right\}\left\{\sum v^2 - \dfrac{(\sum v)^2}{n}\right\}}}$$

$$= \frac{7147 - \dfrac{(167)(289)}{10}}{\sqrt{\left\{4755 - \dfrac{(167)^2}{10}\right\}\left\{11501 - \dfrac{(289)^2}{10}\right\}}}$$

$$= \frac{2320.7}{\sqrt{(1966.1)(3148.9)}} = \frac{2320.7}{2488.18} = 0.93.$$

10.16. (a) Given $r_{XY} = 0.7$.

(i) Since the co-efficient of correlation is symmetric, therefore

$$r_{YX} = r_{XY} = 0.7.$$

(ii) Here $u = -2X$, $v = 3Y$ so that $\bar{u} = -2\bar{X}$ and $\bar{v} = 3\bar{Y}$. Also $u-\bar{u} = -2(X-\bar{X})$ and $v-\bar{v} = 3(Y-\bar{Y})$

$$\therefore \quad r_{uv} = \frac{\sum(u-\bar{u})(v-\bar{v})}{\sqrt{\sum(u-\bar{u})^2 \sum(v-\bar{v})^2}} = \frac{(-2)(3)\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{(4)(9)\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}}$$

$$= \frac{-6}{\sqrt{36}} r_{XY} = -r_{XY} = -0.7.$$

because the signs of the co-efficient of X and Y are different.

(b) Computation of the co-efficient of correlation for the given sample values.

$$r = \frac{\sum XY - n\bar{X}\bar{Y}}{\sqrt{[\sum X^2 - n\bar{X}^2][\sum Y^2 - n\bar{Y}^2]}} = \frac{404-(20)(2)(8)}{\sqrt{[180-20(2)^2][1424-20(8)^2]}}$$

$$= \frac{404-320}{\sqrt{(100)(144)}} = \frac{84}{120} = 0.70$$

10.17. (i) Computation of the correlation co-efficient.

$$\text{Now} \quad r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left\{\sum X^2 - \frac{(\sum X)^2}{n}\right\}\left\{\sum Y^2 - \frac{(\sum Y)^2}{n}\right\}}}$$

$$= \frac{482788 - \frac{(2433)(4245)}{23}}{\sqrt{\left[281019 - \frac{(2433)^2}{23}\right]\left[841786 - \frac{(4245)^2}{23}\right]}}$$

$$= \frac{33740.83}{\sqrt{(23649.92)(58306.66)}} = \frac{33740.83}{37134.185} = 0.91$$

(ii) The high value of correlation co-efficient indicates that there does exist a relationship between X and Y. Therefore the estimated least-squares line of regression is

$$\hat{Y} = a + bX$$

The two normal equations are

$$\sum Y = na + b\sum X;$$

$$\sum XY = a\sum X + b\sum X^2.$$

Substituting the sums in the normal equations, we get

$$4245 = 23a + 2433b$$

$$482788 = 2433a + 281019b$$

Solving them simultaneously, we obtain

$$a = 33.3 \text{ and } b = 1.43.$$

Hence $\hat{Y} = 33.3 + 1.43X$, is the desired estimated least-squares line of regression.

10.18. Calculation of the co-efficient of correlation between traffic density (X) and accident rate (Y).

	X	Y	X^2	Y^2	XY
	30	2	900	4	60
	35	4	1225	16	140
	40	5	1600	25	200
	45	5	2025	25	225
	50	8	2500	64	400
	60	15	3600	225	900
	70	24	4900	576	1680
	80	30	6400	900	2400
	90	32	8100	1024	2880
$\sum$	500	125	31250	2959	8885

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

$$= \frac{8885 - (500)(125)/9}{\sqrt{31250 - (500)^2/9}\sqrt{2959 - (125)^2/9}}$$

$$= \frac{8885 - 6944.44}{\sqrt{3472.22}\sqrt{1222.89}} = \frac{1940.56}{2060.62} = 0.94$$ $= 0.98$

The correlation between traffic density and accident rate is positive and is 0.94.

The coefficient of determination $= r^2$

$$= (0.94)^2 = 0.8836$$

A value of $r^2 = 0.8836$ indicates that 88.36% of the variability in Y is explained by its linear relationship with the variable X and 11.64% of the variation is due to chance or other factors.

10.19. Calculaton of the correlation co-efficient between the marks of Economics Paper (X) and Physics Paper (Y).

Roll No.	X	Y	D_x $(X-50)$	D_y $(Y-50)$	D_x^2	D_y^2	D_xD_y
1	36	62	−14	+12	196	144	−168
2	56	42	6	−8	36	64	−48
3	41	60	−9	10	81	100	−90
4	46	53	−4	3	16	9	−12
5	59	36	9	−14	81	196	−126
6	46	50	−4	0	16	0	0
7	65	42	15	−8	225	64	−120
8	31	66	−19	16	361	256	−304
9	68	44	18	−6	324	36	−108
10	41	58	−9	8	81	64	−72
11	70	65	20	15	400	225	300
12	36	71	−14	21	196	441	−294
Σ	595	649	−5	49	2013	1599	-1042

$$\therefore \quad r = \frac{\sum D_xD_y - (\sum D_x)(\sum D_y)/n}{\sqrt{\sum D_x^2 - (\sum D_x)^2/n}\sqrt{\sum D_y^2 - (\sum D_y)^2/n}}$$

$$= \frac{-1042 - (-5)(49)/12}{\sqrt{2013 - (-5)^2/12}\sqrt{1599 - (49)^2/12}}$$

$$= \frac{-1042 + 20.42}{\sqrt{(2010.92)(1398.92)}} = \frac{-1021.58}{1679.91} = -0.61.$$

There is negative correlation between the marks in Economics and marks in Physics.

10.20. Calculation of the co-efficient of correlation.

	X	Y	D_x $(X-8)$	D_y $(Y-17)$	D_x^2	Dy^2	D_xD_y
	3	25	−5	8	25	64	−40
	4	24	−4	7	16	49	−28
	5	20	−3	3	9	9	−9
	6	20	−2	3	4	9	−6
	7	19	−1	2	1	4	−2
	8	17	0	0	0	0	0
	9	16	1	−1	1	1	−1
	10	13	2	−4	4	16	−8
	11	10	3	−7	9	49	−21
	12	6	4	−11	16	121	−44
Σ	75	170	−5	0	85	322	−159

Here $\bar{X} = \frac{\sum X}{n} = \frac{75}{10} = 7.5, \quad \bar{Y} = \frac{\sum Y}{n} = \frac{170}{10} = 17,$

$$S_x = \sqrt{\frac{\sum D_x^2}{n} - \left(\frac{\sum D_x}{n}\right)^2}$$

$$= \sqrt{\frac{85}{10} - \left(\frac{-5}{10}\right)^2} = \sqrt{85 - 0.25} = 2.87$$

$$S_y = \sqrt{\frac{\sum D_y^2}{n} - \left(\frac{\sum D_y}{n}\right)^2}$$

$$= \sqrt{\frac{322}{10} - \left(\frac{0}{10}\right)^2} = \sqrt{32.2} = 5.67$$

$$\therefore \quad r = \frac{\frac{\sum D_x D_y}{n} - \frac{\sum D_x}{n} \cdot \frac{\sum D_y}{n}}{S_x S_y}$$

$$= \frac{-15.9 - (-0.5)(0)}{(2.87)(5.67)} = \frac{-15.9}{16.27} = -0.98.$$

The equation to the regression line of Y on X is

$$Y - \bar{Y} = r \frac{S_y}{S_x}(X - \bar{X})$$

or $$Y-17 = -0.98\left(\frac{5.67}{2.87}\right)(X-7.5)$$

$$= -1.94\,(X-7.5)$$

or $$Y = 17 - 1.94X + 14.55 = 31.55 - 1.94X$$

The equation to the regression line of X on Y is

$$X - \bar{X} = r \frac{S_x}{S_y}(Y - \bar{Y})$$

or $$X - 7.5 = -0.98\left(\frac{2.87}{5.67}\right)(Y - 17)$$

$$= -0.50\,(Y - 17)$$

or $$X = 7.5 - 0.50Y + 8.5 = 16.0 - 0.50Y.$$

10.21 (a) Calculation of the correlation co-efficient.

	X	Y	D_x $(X-18)$	D_y $(Y-20)$	D_x^2	D_y^2	D_xD_y
	5	11	−13	−9	169	81	117
	12	16	−6	−4	36	16	24
	14	15	−4	−5	16	25	20
	16	20	−2	0	4	0	0
	18	17	0	−3	0	9	0
	21	19	3	−1	9	1	−3
	22	25	4	5	16	25	20
	23	24	5	4	25	16	20
	25	21	7	1	49	1	7
Σ	156	168	−6	−12	324	174	205

$$\therefore \quad r = \frac{\sum D_x D_y - (\sum D_x)(\sum D_y)/n}{\sqrt{\sum D_x^2 - (\sum D_x)^2/n}\sqrt{\sum D_y^2 - (\sum D_y)^2/n}}$$

$$= \frac{205 - (-6)(-12)/9}{\sqrt{324 - (-6)^2/9}\sqrt{174 - (-12)^2/9}}$$

$$= \frac{205 - 8}{\sqrt{(320)(158)}} = \frac{197}{224.85} = 0.876.$$

(b) Calculation of the co-efficient of correlation between persons employed (X) and cloth manufactured (Y).

	X	Y	D_x $(X-176)$	D_y $(Y-40)$	$D_x D_y$	D_x^2	D_y^2
	137	23	−39	−17	663	1521	289
	209	47	33	7	231	1089	49
	113	22	−43	−18	774	1849	324
	189	40	13	0	0	169	0
	176	39	0	− 1	0	0	1
	200	51	24	11	264	576	121
	219	49	43	9	387	1849	81
Σ	1243	271	31	−9	2319	7053	865

$$\therefore \quad r = \frac{\sum D_x D_y - (\sum D_x)(\sum D_y)/n}{\sqrt{\sum D_x^2 - (\sum D_x)^2/n}\sqrt{\sum D_y^2 - (\sum D_y)^2/n}}$$

$$= \frac{2319 - (31)(-9)/7}{\sqrt{7053 - (31)^2/7}\sqrt{865 - (-9)^2/7}}$$

$$= \frac{2319 + 39.86}{\sqrt{(6915.71)(853.43)}} = \frac{2358.86}{2429.42} = 0.97.$$

Thus the co-efficient of correlation between persons employed and cloth manufactured is 0.97.

10.22. To find relation between age and blindness, we first calculate the numbers of blind per lakh (100,000) and then correlate with the midpoints of age groups.

Age Group	Mid. value X	Blinds per lakh Y	D_x $(X-44.5)$	D_y $(Y-150)$	D_xD_y	D_x^2	Dy^2
0–9	4.5	55	−40	−95	3800	1600	9025
10–19	14.5	67	−30	−83	2490	900	6889
20–29	24.5	100	−20	−50	1000	400	2500
30–39	34.5	111	−10	−39	390	100	1521
40–49	44.5	150	0	0	0	0	0
50–59	54.5	200	10	50	500	100	2500
60–69	64.5	300	20	150	3000	400	22500
70–79	74.5	500	30	350	10500	900	122500
Total	--	--	−40	283	21680	4400	167435

$$\therefore \quad r = \frac{\sum D_xD_y - (\sum D_x)(\sum D_y)/n}{\sqrt{\sum D_x^2 - (\sum D_x)^2/n}\sqrt{\sum D_y^2 - (\sum D_y)^2/n}}$$

$$= \frac{21680 - (-40)(283)/8}{\sqrt{4400 - (-40)^2/8}\sqrt{167435 - (283)^2/8}}$$

$$= \frac{21680 + 1415}{\sqrt{(4200)(157424)}} = \frac{23095}{25713} = 0.898.$$

The co-efficient of correlation is positive and very high, implying that blindness increases with age.

10.23. The corrected sums are calculated first as below:

$$\text{Corrected } \sum X = 125 - 6 - 8 + 8 + 6 = 125$$

$$\text{Corrected } \sum Y = 100 - 14 - 6 + 12 + 8 = 100$$

$$\text{Corrected } \sum X^2 = 650 - (6^2 + 8^2) + (8^2 + 6^2) = 650$$

$$\text{Corrected } \sum Y^2 = 640 - (14^2 + 6^2) + (12^2 + 8^2) = 436$$

$$\text{Corrected } \sum XY = 508 - (6 \times 14 + 8 \times 6) + (12 \times 8 + 6 \times 8)$$

$$= 508 - (84 + 48) + (96 + 48) = 520$$

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

$$= \frac{\frac{520}{25} - \left(\frac{125}{25}\right)\left(\frac{100}{25}\right)}{\sqrt{\frac{650}{25} - \left(\frac{125}{25}\right)^2}\sqrt{\frac{436}{25} - \left(\frac{100}{25}\right)^2}}$$

$$= \frac{20.8 - 20}{\sqrt{(26-25)\ (17.44-16)}} = \frac{0.8}{1.2} = 0.67.$$

10.24. (a) Here $b_{yx} = -0.219$ and $b_{xy} = -0.785$

Since the regression coefficients are negative, therefore r is given by

$$r = -\sqrt{b_{yx} \cdot b_{xy}} = -\sqrt{(-0.219)\ (-0.785)} = -0.415$$

(b) Here $b_{yx} = 0.648$ and $b_{xy} = 0.917$

Since the regression co-efficients are positive, therefore r is given by

$$r = +\sqrt{b_{yx} \cdot b_{xy}} = +\sqrt{(0.648)\ (0.917)} = +0.77$$

(c) Here $b_{yx} = 1.94$ and $b_{xy} = 0.15$

Since the regression co-efficients are positive, so r is given by

$$r = +\sqrt{b_{yx} \cdot b_{xy}} = +\sqrt{(1.94)\ (0.15)} = +0.54$$

(d) Here $b_{yx} = -1.96$, but the line of X on Y is given as $Y = 15.91 - 2.22X$ To determine b_{xy}, we should transform the above equatioin as follows:

$$-2.22X = Y - 15.91$$

or $$X = \frac{-Y}{2.22} + \frac{15.91}{2.22} = -0.45Y + 7.167$$

Thus $b_{xy} = -0.45$

Hence $r = -\sqrt{b_{yx} \times b_{xy}}$, as the regression coefficients are negative.

$$= -\sqrt{(-1.96)\ (-0.45)} = -0.94.$$

10.25. **Taking** $\mathbf{u = \dfrac{X-29.5}{10}}$ **and** $\mathbf{v = \dfrac{Y-24.5}{10}}$**, the arithmetic is arranged in the table below:**

Y_i	X_j / u_j / v_i	9.5	19.5	29.5	39.5	49.5	$f_{i\cdot}$	$f_{i\cdot}v_i$	$f_{i\cdot}v_i^2$	$f_{ij}u_jv_i$
		−2	−1	0	1	2				
4.5	−2	12 3	2 1	--	--	--	4	−8	16	14
14.5	−1	24 12	8 8	0 14	−1 1	--	35	−35	35	31
24.5	0	0 2	0 13	0 40	0 12	0 3	70	0	0	0
34.5	1	---	−3 3	0 40	27 27	14 7	77	77	77	38
44.5	2	---	---	0 6	8 4	16 4	14	28	56	24
$f_{\cdot j}$		17	25	100	44	14	200	62	184	107
$f_{\cdot j}u_j$		−34	−25	0	44	28	13			↑
$f_{\cdot j}u_j^2$		68	25	0	44	56	193			|
$f_{ij}u_jv_i$		36	7	0	34	30	107	←---- Check		|

$$\therefore \quad r = \frac{\sum fuv - (\sum fu)(\sum fv)/n}{\sqrt{\left[\sum fu^2 - (\sum fu)^2/n\right]}\sqrt{\left[\sum fv^2 - (\sum fv)^2/n\right]}}$$

$$= \frac{107 - (13)(62)/200}{\left[\sqrt{193 - (13)^2/200}\right]\left[\sqrt{184 - (62)^2/200}\right]}$$

$$= \frac{107 - 4.03}{\sqrt{(192.155)\ (164.78)}} = \frac{102.97}{177.94} = 0.58.$$

The estimated equation to the regression line of Y on X is

$$Y - \bar{Y} = r\frac{s_y}{s_x}(X - \bar{X}).$$

where $\bar{Y} = b + \frac{\sum fv}{n} \times k = 24.5 + \frac{62 \times 10}{200} = 27.6;$

$$\bar{X} = a + \frac{\sum fu}{n} \times h = 29.5 + \frac{13 \times 10}{200} = 30.15;$$

$$s_x = \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} \times h = \sqrt{\frac{193}{200} - \left(\frac{13}{200}\right)^2} \times 10 = 9.8; \text{ and}$$

$$s_y = \sqrt{\frac{\sum fv^2}{n} - \left(\frac{\sum fv}{n}\right)^2} \times k = \sqrt{\frac{184}{200} - \left(\frac{62}{200}\right)^2} \times 10 = 9.1.$$

Substituting the values, we get

$$Y - 27.6 = 0.58 \left(\frac{9.1}{9.8}\right) (X - 30.15)$$

or $\quad Y - 27.6 = 0.54\,(X - 30.15)$

or $\quad Y = 0.54X - 16.28 + 27.6 = 11.32 + 0.54X,$

which is the desired regression equation of Y on X.

10.26. Let $u = \frac{X-130}{20}$ and $v = \frac{Y-63}{3}$, where X denotes weight in pounds and Y, the height in inches. The calculations are then arranged as in the following table:

Y_i		90	110	130	150	170	190	Total			
	u / v	−2	−1	0	1	2	3	(f)	fv	fv^2	fuv
57	−2	--	--	--	--	$^{-4}$ 1	--	1	−2	4	−4
60	−1	^{16} 8	^{21} 21	^{0} 8	$^{-1}$ 1	--	--	38	−38	38	36
63	0	^{0} 3	^{0} 50	^{0} 57	^{0} 12	--	^{0} 2	124	0	0	0
66	1	$^{-2}$ 1	$^{-24}$ 24	^{0} 54	^{19} 19	^{6} 3	--	101	101	101	−1
69	2	--	$^{-2}$ 1	^{0} 8	^{10} 5	^{12} 3	--	17	34	68	20
72	3	--	--	^{0} 2	--	--	^{9} 1	3	9	27	9
Total (f)		12	96	129	37	7	3	284	104	238	60
fu		−24	−96	0	37	14	9	−60			↑
fu^2		48	96	0	37	28	27	236			|
fuv		14	−5	0	28	14	9	60	←-- Check		|

$$\therefore \quad r = \frac{\sum fuv - (\sum fu)(\sum fv)/n}{\sqrt{\sum fu^2 - (\sum fu)^2/n}\sqrt{\sum fv^2 - (\sum fv)^2/n}}, \text{ where } n = \sum f$$

$$= \frac{60 - (-60)(104)/284}{\sqrt{236 - (-60)^2/284}\sqrt{238 - (104)^2/284}}$$

$$= \frac{60 + 21.97}{\sqrt{(223.32)(199.92)}} = \frac{81.97}{211.30} = 0.39$$

10.27 (b) Here

$$r = \frac{\sum fuv - (\sum fu)(\sum fv)/n}{\sqrt{\sum fu^2 - (\sum fu)^2/n}\sqrt{\sum fv^2 - (\sum fv)^2/n}}$$

$$= \frac{91 - (-4)(-53)/66}{\sqrt{109 - (-4)^2/66}\sqrt{115 - (-53)^2/66}}$$

$$= \frac{87.79}{\sqrt{(108.76)(72.56)}} = \frac{87.79}{88.83} = 0.9883$$

Now $\quad u = \dfrac{x - 1250}{500}$ or $x = 1250 + 500\,u$

$$\therefore \quad \bar{x} = 1250 + 500\left(\frac{-4}{66}\right) = 1250 - 30.30 = 1219.70,$$

And $\quad v = \dfrac{y - 500}{200}$ or $y = 500 + 200\,v$

$$\therefore \quad \bar{y} = 500 + 200\left(\frac{-53}{66}\right) = 500 - 160.61 = 339.39$$

$$s_x = h\sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} = 500\sqrt{\frac{109}{66} - \left(\frac{-4}{66}\right)^2}$$

$$= 500\sqrt{1.6515 - 0.0037} = 500\,(1.2837) = 641.85,$$

$$s_y = k\sqrt{\frac{\sum fv^2}{n} - \left(\frac{\sum fv}{n}\right)^2} = 200\sqrt{\frac{115}{66} - \left(\frac{-53}{66}\right)^2}$$

$$= 200\sqrt{1.7424 - 0.6449} = 200\,(1.0476) = 209.52$$

Regression coefficients are:

$$b_{yx} = r \cdot \frac{s_y}{s_x} = 0.9883 \times \frac{209.52}{641.85} = 0.3226, \text{ and}$$

$$b_{xy} = r \cdot \frac{s_x}{s_y} = 0.9883 \times \frac{641.85}{209.52} = 3.0276.$$

Regression line of Y on X is $Y - \bar{Y} = b_{yx}(X - \bar{X})$,

i.e. $Y - 339.39 = 0.3226\,(X - 1219.70)$

or $Y = 0.3226X - 54.085$

Regression line of X on Y is $X - \bar{X} = b_{xy}(Y - \bar{Y})$

or $X - 1219.70 = 3.0276\,(Y - 339.39)$ or $X = 3.0276Y + 192.163$.

10.28. Let $u = X_1 + X_2$ and $v = X_2 + X_3$. Then

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

Now $u - \bar{u} = (X_1 + X_2) - (\bar{X}_1 + \bar{X}_2)$, $(\because\ \bar{u} = \bar{X}_1 + \bar{X}_2)$

$= (X_1 - \bar{X}_1) + (X_2 - \bar{X}_2)$, and

$v - \bar{v} = (X_2 + X_3) - (\bar{X}_2 + \bar{X}_3)$, $(\because\ \bar{v} = \bar{X}_2 + \bar{X}_3)$

$= (X_2 - \bar{X}_2) + (X_3 - \bar{X}_3)$.

$\therefore\ \sum(u-\bar{u})^2 = [(X_1 - \bar{X}_1) + (X_2 - \bar{X}_2)]^2$

$= \sum(X_1-\bar{X}_1)^2 + \sum(X_2-\bar{X}_2)^2 + 2\sum(X_1-\bar{X}_1)\sum(X_2-\bar{X}_2)$

$= nS^2 + nS^2 + 0$ $(\because\ X_1$ and X_2 are uncorrelated)

$= 2nS^2$,

$\therefore\ \sum(v-\bar{v})^2 = \sum[(X_2 - \bar{X}_2) + (X_3 - \bar{X}_3)]^2$

$= \sum(X_2-\bar{X}_2)^2 + \sum(X_3-\bar{X}_3)^2 + 2\sum(X_2-\bar{X}_2)(X_3-\bar{X}_3)$

$= nS^2 + nS^2 + 0$ $(\because\ X_2$ and X_3 are uncorrelated)

$= 2nS^2$, and

$\therefore\ \sum(u-\bar{u})(v-\bar{v}) = \sum[(X_1-\bar{X}_1) + (X_2-\bar{X}_2)]\,[(X_2-\bar{X}_2) + (X_3-\bar{X}_3)]$

$= \sum(X_1-\bar{X}_1)(X_2-\bar{X}_2) + \sum(X_1-\bar{X}_1)(X_3-\bar{X}_3)$
$+ \sum(X_2-\bar{X}_2)^2 + \sum(X_2-\bar{X}_2)(X_3-\bar{X}_3)$.

$= 0 + 0 + nS^2 + 0 = nS^2$.

Substituting these values, we get

$$r_{uv} = \frac{nS^2}{\sqrt{(2nS^2)(2nS^2)}} = \frac{nS^2}{2nS^2} = \frac{1}{2}$$

10.29. (b) Calculation of the rank correlation co-efficient.

Here $d = 0, -8, 0, 0, 0, -1, 5, 2, 1, -1, -4, 3, -1, 2, -1, 3.$

$\therefore \quad \sum d^2 = 0 + 64 + 0 + 0 + 0 + 1 + 25 + 4 + 1 + 1 + 16 + 9 + 1 + 4 + 1 + 9 = 136$

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

$$= 1 - \frac{6 \times 136}{16(256 - 1)} = 1 - 0.2 = 0.8$$

10.30. (b) Calculations for the product moment co-efficient of correlation.

a	b	a^2	b^2	ab
7.4	8.5	54.76	72.25	62.90
9.0	6.1	81.00	37.21	54.90
11.0	2.4	121.00	5.76	26.40
2.5	6.7	6.25	44.89	16.75
4.6	12.6	21.16	158.76	57.96
6.5	3.3	42.25	10.89	21.45
41.0	39.6	326.42	329.76	258.36

$$\therefore \quad r = \frac{n\sum ab - (\sum a)(\sum b)}{\sqrt{n\sum a^2 - (\sum a)^2}\sqrt{n\sum b^2 - (\sum b)^2}}$$

$$= \frac{(6)(258.36) - (41.0)(39.6)}{\sqrt{(6)(326.42) - (41.0)^2}\sqrt{(6)(329.76) - (39.6)^2}}$$

$$= \frac{1550.16 - 1623.60}{\sqrt{(277.52)(410.40)}} = \frac{-73.44}{337.48} = -0.22$$

Ranking the values and calculation of the rank correlation co-efficient.

		Ranks			
a	b	a	b	$d(=a-b)$	d^2
7.4	8.5	3	2	1	1
9.0	6.1	2	4	−2	4
11.0	2.4	1	6	−5	25
2.5	6.7	6	3	3	9
4.6	12.6	5	1	4	16
6.5	3.3	4	5	−1	1
Σ	--	--	--	--	56

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

$$= 1 - \frac{6 \times 56}{6(36-1)} = 1 - 1.6 = -0.6.$$

10.31. (b) Calculation of the co-efficient of rank correlation.

Laboratory (X)	8	3	9	2	7	10	4	6	1	5	Total
Lecture (Y)	9	5	10	1	8	7	3	4	2	6	--
$d(=X-Y)$	−1	−2	−1	1	−1	3	1	2	−1	−1	--
d^2	1	4	1	1	1	9	1	4	1	1	24

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

$$= 1 - \frac{6 \times 24}{10(100-1)} = 1 - 0.14556 = 0.8545$$

10.32. Denoting the judges by 1, 2 and 3, the calculations are given below:

Judge 1	Judge 2	Judge 3	d_{12}	d_{12}^2	d_{23}	d_{23}^2	d_{13}	d_{13}^2
1	3	6	−2	4	−3	9	−5	25
6	5	4	1	1	1	1	2	4
5	8	9	−3	9	−1	1	−4	16
10	4	8	6	36	−4	16	2	4
3	7	1	−4	16	6	36	2	4
2	10	2	−8	64	8	64	0	0
4	2	3	2	4	−1	1	1	1
9	1	10	8	64	−9	81	−1	1
7	6	5	1	1	1	1	2	4
8	9	7	−1	1	2	4	1	1
$\sum$	--	--	--	200	--	214	--	60

$$\therefore \quad r_{s(12)} = 1 - \frac{6\sum d_{12}^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 200}{10(100-1)} = 1 - 1.21 = -0.21;$$

$$r_{s(23)} = 1 - \frac{6\sum d_{23}^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 214}{10 \times 99} = 1 - 1.30 = -0.30; \text{ and}$$

$$r_{s(13)} = 1 - \frac{6\sum d_{13}^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 60}{10(100-1)} = 1 - 0.36 = +0.64$$

The co-efficient of rank correlation between the first Judge and the third Judge is positive, meaning that the Judges 1 and 3 have the nearest approach to the tastes in beauty.

10.33. Calculation of Spearman's rank correlation coefficients.

Entry	Judge X	Judge Y	Judge Z	d_{xy}	d^2_{xy}	d_{yz}	d^2_{yz}	d_{xz}	d^2_{xz}
A	5	1	6	4	16	−5	25	−1	1
B	2	7	4	−5	25	3	9	−2	4
C	6	6	9	0	0	−3	9	−3	9
D	8	10	8	−2	4	2	4	0	0
E	1	4	1	−3	9	3	9	0	0
F	7	5	2	2	4	3	9	5	25
G	4	3	3	1	1	0	0	1	1
H	9	8	10	1	1	−2	4	−1	1
K	3	2	5	1	1	−3	9	−2	4
L	10	9	7	1	1	2	4	3	9
Σ	--	--	--	--	62	--	82	--	54

Now $$r_{s(xy)} = 1 - \frac{6\sum d^2_{xy}}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 62}{10(100-1)} = 1 - 0.38 = 0.62;$$

$$r_{s(yz)} = 1 - \frac{6\sum d^2_{yz}}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 82}{10(100-1)} = 1 - 0.50 = 0.50;$$

$$r_{s(xz)} = 1 - \frac{6\sum d^2_{xz}}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 54}{10(100-1)} = 1 - 0.33 = 0.67.$$

The rank correlation co-efficient between Judges X and Z is the highest and is positive. We may therefore conclude that the pair of Judges (X, Z) has the nearest approach to common tastes.

10.34. (b) We observe that both the sets of ranks contain ties. The co-efficient of rank correlation is therefore calculated as below:

X	8	3	6.5	3	6.5	9	3	1	5	Total
Y	8	9	6.5	2.5	4	5	6.5	1	2.5	--
$d(=X-Y)$	0	−6	0	0.5	2.5	4	−3.5	0	2.5	0
d^2	0	36	0	0.25	6.25	16	12.25	0	6.25	77

For ties, we add the following quantity to $\sum d^2$

$$\frac{1}{12}(3^3-3)+\frac{1}{12}(2^3-2)+\frac{1}{12}(2^3-2)+\frac{1}{12}(2^3-2)$$

i.e. $\quad 2 + 0.5 + 0.5 + 0.5$ or 3.5

Hence $r_s = 1 - \dfrac{6(77+3.5)}{9(81-1)} = 1 - \dfrac{483}{720} = 1 - 0.67 = 0.33.$

10.35. Calculation of the co-efficient of concordance.

X	1	2	3	4	5	6	7	8	9	10
Y	7	10	4	1	6	8	9	5	2	3
Z	9	6	10	3	5	4	7	8	2	1
Total	17	18	17	8	16	18	23	21	13	14

Here $\quad m = 3$ and $n = 10$.

$\therefore \quad$ Mean $= \dfrac{m(n+1)}{2} = \dfrac{3(10+1)}{2} = 16.5$, and

$$S = (17-16.5)^2 + (18-16.5)^2 + \ldots + (13-16.5)^2 + (14-16.5)^2$$
$$= (0.5)^2 + (1.5)^2 + (0.5)^2 + (-8.5)^2 + (-0.5)^2 + (1.5)^2$$
$$+ (6.5)^2 + (4.5)^2 + (-3.5)^2 + (-2.5)^2 = 158.50.$$

Hence the co-efficient of concordance, W, is

$$W = \frac{12 \times S}{m^2(n^3-n)} = \frac{12 \times 158.50}{9(1000-10)} = \frac{634}{2970} = 0.21.$$

●●●●●●●●●