30/5/22 PROBABILITY & STATISTICS 2: ASSIGNMENT 2 (Roll no. 87)

Ans 1)

~~$\Sigma x_i y_i$~~ x_i	y_i	$x_i y_i$	x_i^2
40	56	2240	1600
10	62	620	100
100	195	19500	10000
110	240	26400	12100
120	170	20400	14400
150	270	40500	22500
20	48	960	400
90	196	17640	8100
80	214	17120	6400
130	286	37180	16900
850	1737	182560	92500

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{850}{10} = 85$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{1737}{10} = 173.7$$

$$\hat{\beta} = \frac{\Sigma x_i y_i - n\bar{x}\bar{y}}{\Sigma x_i^2 - n\bar{x}^2} = \frac{182560 - 10(85)(173.7)}{92500 - 10(85)^2}$$

$$= \frac{34915}{20250} = 1.7241$$

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 173.7 - (1.724)(85)$
$= 27.16$

$\therefore y = \cancel{1.724x+} \; 27.16 + 1.724x$

Ans 2) i) $\hat{\beta} = \dfrac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \dfrac{38191.41 - 12(\cancel{385.2}\;32.1)(\cancel{91.87}\;96.875)}{12666.58 - 12(\cancel{1055.84}\;(32.1)^2)}$

$= \dfrac{875.16}{301.66} = 2.9011$

$\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{385.2}{12} = 32.1$

$\bar{y} = \dfrac{\sum y_i}{n} = \dfrac{1162.5}{12} = 96.875$

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$
$= 96.875 - (2.9011)(32.1)$
$= 3.74969$

$y = 3.749 + 2.9011x$

ii) $\dfrac{\hat{\beta} - \beta}{\sqrt{\sigma^2/S_{xx}}} \sim t_{n-2}$

$= 95\%$ CI $\Rightarrow \hat{\beta} + t_{n-2,\,0.025/2.5\%}\sqrt{\dfrac{\sigma^2}{S_{xx}}}$

$= 2.9011 \pm t_{10,\,2.5\%}\sqrt{\dfrac{\sigma^2}{S_{xx}}}$

$S_{yy} = \Sigma y^2 - n\bar{y}^2 = 119026.9 - 12(96.875)^2 = 6409.7$

$$\sigma^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right)$$

$$= \frac{1}{10}\left(6409.7 - \frac{(87.5.16)^2}{301.66}\right)$$

$$= \cancel{6155.816} \quad 387.07$$

$$\Rightarrow 2.9011 \pm 2.228\sqrt{\frac{\cancel{6155.816}\ 387.07}{301.66}}$$

$$= \cancel{(-7.163, 12.9657)} \qquad (0.3773, 5.4248)$$

Ans 3) i) The variation in city C is low while variation in city D is large.

ii) The population are normally distributed & iid.

iii) $SS_T = 1495 - \frac{(103)^2}{28} = 1116.107$

$$SS_B = \cancel{\frac{1495}{28} - \frac{(103)^2}{28}}$$

$$= \frac{(64^2 + 44^2 + 8^2 + (-13)^2)}{7} - \frac{(103)^2}{28}$$

$$= 516.107$$

$$SS_R = SS_T - SS_B$$

$$= 1116.107 - 516.107$$

$$= 599.9998 \approx 600$$

Ans 4) a) $y_{ij} \sim N(\mu + \tau_i, \sigma^2)$, $i = 1, 2 \dots k$
$j = 1, 2 \dots n$

$n = \sum_{i=1}^{k} n_i$ with $\sum_{i=1}^{k} n_i \tau_i = 0$

b) $SS_T = \sum y^2 - \frac{(y)^2}{n}$

$= \cancel{174376} - \cancel{\frac{1856}{}} = 174316 - \frac{845362}{n}$

$= 69930.06$

$SS_R = \frac{\sum y_i^2}{k} - \frac{y^2}{n}$

$= 22325.7$

$S_{RES} = SS_T - SS_R$

$= 47604.4$

c) Using contingency table

	Salary			row Total
45	20	6	5	76
7	20	9	6	42
5	8	15	12	40
57	48	30	23	158.

$= \frac{\text{row total} \times \text{column total}}{\text{Table total}}$

$\Rightarrow \cancel{9\,76 \times 57 / 158}$ . . . . .

. . . . . .

= 27.4	23.8	14.4	11.06
15.4	12.7	8.19	6.09
14.43	12.6	7.8	5.8

$$\chi^2 = \frac{\Sigma(O_i - E_i)^2}{E_i}$$

Ans 5 i) The samples ~~is~~ taken from normal distribution
- The observations are iid
- The variance of different groups is same

Ans 6) $r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}} = \frac{\cancel{66.12}\ 661.2}{\sqrt{54.9 \times 8923.6}} = 0.94466$

$$S_{xy} = \Sigma XY - n\bar{X}\bar{Y}$$
$$= 2853 - 10\left(\frac{39}{10}\right)\left(\frac{562}{10}\right)$$
$$= 661.2$$

$$S_{xx} = \Sigma X^2 - n\bar{X}^2$$
$$= 207 - 10(3.9)^2$$
$$= 54.9$$

$$S_{yy} = \Sigma Y^2 - n\bar{Y}^2$$
$$= 40{,}508 - (10)(56.2)^2$$
$$= 8923.6.$$

ii) $B = \frac{S_{xy}}{S_{xx}} = \frac{661.2}{54.9} = 12.043$

$\alpha = \bar{Y} - \hat{B}\bar{X} = 9.2323$

$y = 9.2323 + 12.043X$

$= SS_T = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$

$= 8923.6$

$SS_{REG} = \frac{S^2_{xy}}{S_{xx}}$

$= 7963.305$

$SS_{RES} = SS_T - SS_{REG}$

$= 960.295$

iv) $R^2 = \frac{SS_{REG}}{SS_{TOT}} = 0.8923$

Ans 7) i) $S_{xy} = \Sigma xy - n\bar{x}\bar{y}$

$= \Sigma(x-\bar{x})(y-\bar{y})$

$= 2176.84$

$S_{xx} = \Sigma x_i^2 - \frac{(\Sigma x)^2}{n}$

$= 4789.4221$

$\beta = \frac{S_{xy}}{S_{xx}} = 0.4545$

$\alpha = \bar{y} - \beta\bar{x}$

$= \left(\frac{197.84}{11}\right) - (0.4545)\left(\frac{174.13}{11}\right)$

$= 17.9854 - 7.1947$

$= 10.7907$

$y = 10.7907 + 0.4545x$

ii) $\dfrac{\hat{\beta} - \beta}{\sqrt{\sigma^2 / S_{xx}}} \sim t_9$

$$\sigma^2 = \frac{1}{9}\left(S_{yy} - \frac{S^2_{xy}}{S_{xx}} \right)$$

$$= \frac{1}{9}(1189.207 - 989.395)$$

$$= 22.201$$

iii) $$r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}}$$

$$= \frac{2176.84}{\sqrt{4789.4221 \times 1189.207}}$$

$$= 0.9121289159$$

Ans 8) i) It is used to reduce observed, correlated variables to lower ~~no~~ number of variables/dimensions known as factors.

ii)

Entry	Percentage
0.456	65%.
0.137	19.5%.
0.080	11.4%.
0.0165	2.35%.
0.012	1.7%.
0.7015	

Ans 9) i) $r = \dfrac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}}$

$$S_{xy} = \Sigma xy - n\bar{x}\bar{y}$$
$$= 2602 - 10\left(\frac{45}{10}\right)\left(\frac{644}{10}\right)$$
$$= -296$$

$$S_{xx} = \Sigma x_i^2 - n\bar{x}^2$$
$$= 277.5 - 10(4.5)^2$$
$$= 75$$

$$S_{yy} = \Sigma y_i^2 - n\bar{y}^2$$
$$= 43956 - 10(64.4)^2$$
$$= 2482.4$$

$$r = \frac{-296}{\sqrt{75 \times 2482.4}} = -0.686002$$

iv) $\beta = \dfrac{S_{xy}}{S_{xx}} = \dfrac{-296}{75} = -3.94667$

$$\alpha = \bar{y} - \hat{\beta}\bar{x}$$
$$= 64.4 + (3.94667)(4.5)$$
$$= 82.16$$

$$y = 82.16 - 3.94667x$$

v) $R^2 = \dfrac{S_{xy}^2}{S_{xx} \cdot S_{yy}} = 0.47059$

classmate

Date

Page

vi) $x = 1$

$y = \hat{\alpha} - \hat{\beta} x$

$= 82.16 - 3.94667(1)$

$= 78.213$

Ans 10) $S_{xx} = 0.0042$

$S_{yy} = 0.00126$

$S_{xy} = 0.0022$

$SS_T = 0.00126$

$SS_R = \dfrac{S^2_{xy}}{S_{xx}} = 0.0115$

$SS_{Res} = 0.0011$