Project Prob & Stat Assignment 2

Q1)

a) $\sum x_i = 850$ $\quad \sum x_i^2 = 925000$ $\quad (\bar{x})$

$\sum y_i = 1737$ $\quad \sum y_i^2 = 3727\,17$

$\sum x_i y_i = 182560$

$$S_{xy} = \sum x_i y_i - n(\bar{x})(\bar{y})$$

$$= 182560 - 10 \times \frac{850}{10} \times \frac{1737}{10}$$

$$= 34915$$

$$S_{xx} = \sum x_i^2 - n(\bar{x})^2$$

$$= 925000 - \frac{850^2}{10}$$

$$= 852750$$

$$S_{yy} = \sum y_i^2 - n(\bar{y})^2$$

$$= 3727\,17 - \frac{(1737)^2}{10}$$

$$= 71000.1$$

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{34915}{852750} = 0.040944$$

$\therefore \bar{y} = \hat{\alpha} + \hat{\beta}\bar{x}$

$\therefore \hat{\alpha} = 170.21976$

b) .. As $\hat{\beta}$ is close to zero, hence, it is not significant for

the model.

Q2) $\Sigma x = 385.2$

i) $\Sigma x^2 = 12666.58$ $\quad n = 12$

$\Sigma y = 1162.5$

$\Sigma y^2 = 119026.9$

$\Sigma xy = 38191.41$

$\therefore S_{xy} = \Sigma xy - n(\bar{x})(\bar{y})$

$= 38191.41 - \frac{385.2 \times 12666.58}{12}$

$= 875.16$

$\therefore S_{xx} = \Sigma x^2 - n(\bar{x})^2$

$= 12666.58 - \frac{385.2^2}{12}$

$= 301.66$

$\therefore S_{yy} = \Sigma y^2 - n(\bar{y})^2$

$= 6409.7125$

$\hat{\beta} = \frac{875.16}{301.60} = 2.9011$

$\therefore \bar{y} = \hat{\alpha} + \hat{\beta}\bar{x}$

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

$= 3.7497$

$\therefore$ The model,

$y_i = 3.7497 + 2.9011x_i$

(ii)

$\hat{\beta} - \beta$

$\hat{\beta} \sim N(\beta, \sigma^2/S_{xx})$

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

$\therefore C.I \Rightarrow \hat{\beta} \pm t_{0.025,10} \sqrt{\hat{\sigma}^2/S_{xx}}$

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

$= \dfrac{1}{10}(\quad)$

$= 387.0745$

$C.I \Rightarrow \hat{\beta} \pm t_{\alpha/2,n-2} \sqrt{\hat{\sigma}^2/S_{xx}}$

$\therefore C.I \Rightarrow 2.9011 \pm 2.228 \sqrt{\dfrac{387.0745}{301.66}}$

$= 2.9011 \pm 2.5238$

$= (0.3773, 5.4249)$

(iii)

$y_i \sim N\left(\hat{\alpha} + \hat{\beta}x_0, \left(\dfrac{1}{n} + \dfrac{(x_0 - \bar{x})^2}{S_{xx}}\right)\sigma^2\right)$

$\therefore E(\mu_0) = \hat{\alpha} + \hat{\beta}x_0$

$= 119.7936$

$$\therefore \text{C.I} \Rightarrow \mu_0 \pm t_{\alpha/2, n-2} \sqrt{\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right) \hat{\sigma}^2}$$

$$= 119.7936 \pm 2.228 \sqrt{\left(\frac{1}{12} + \frac{(40 - 32.1)^2}{301.66}\right) 387.0745}$$

$$= 119.7936 \pm 22.1431$$

$$= (97.6504, 141.9366)$$

Q3)

i) The variance of all the cities appear to be mostly same. The centres of all distributions differ.

ii) Assumptions →
Population must be normal
Must have common variance.
Observations are independent

iii) H_0 ⇒ mean is same
H_1 ⇒ Mean is not same

$$SS_T = 1495 - \frac{103^2}{28} = 1116.11$$

$$SS_{reg} = \frac{1}{7}(64^2 + 44^2 + 8^2 + (-13)^2) - \frac{103^2}{28}$$

$$= 516.11$$

$$SS_{RES} = 1116.11 - 516.11$$

$$= 600$$

ANOVA Table

Source	df	SS	MS	F
Reg	3	516.11	172.04	6.88
Res	24	600	25	
Tot	27	1116.11		

$F = 6.88$

$F_{3,24} = 3.089$ at 5% LOS

$\therefore$ We reject H_0.

Q4)

b) $SS_{TOT} = \sum y_{ij}^2 - \frac{y^2}{n}$

$= 69930.06$

$SS_{REG} = \frac{\sum y_i^2}{n_i} - \frac{y^2}{n_i}$

$= 22325.69$

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

$= 47604.37$

ANOVA Table

Source	df	SS	MSS	F
Reg	4	22325.69	5581.42	3.28
Res	28	47604.37	1700.156	
Tot	32	69930.06		

$F = 3.28$ $\quad F_{4,28} = 2.714$ at 5% LOS

∴ We reject H_0

c)

	Salary			
Papers Cleared	45	20	6	5
	7	20	9	6
	5	8	15	12

$$E_i = \frac{\text{row tot} \times \text{col. tot}}{\text{Table tot.}}$$

∴ E_i =

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

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

$$= 49.919$$

$$\chi^2_{0.025,7} = 17.53$$

∴ We reject H_0.

∴ No. of papers cleared are dependent on salary.

Q6) $\Sigma x = 39$

$\Sigma x^2 = 207$

$\Sigma y = 562$

$\Sigma y^2 = 40,508$

$\Sigma xy = 2853$

$S_{xx} = 54.9$

$S_{yy} = 8923.6$

$S_{xy} = 661.2$

i) $\gamma = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = 0.944662$

ii) $\hat{B} = \frac{S_{xy}}{S_{xx}} = 12.0437$

$\therefore \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 9.22957$

$\therefore$ Eqn $= 9.22957 + 12.0437 x_i$

iii) $SS_{Tot} = SS_{reg} + SS_{res}$

$SS_{reg} = \frac{S^2_{xy}}{S_{xx}} = 7963.3$

$SS_{res} = S_{yy} - \frac{S^2_{xy}}{S_{xx}} = 960.295$

$\therefore SS_{Tot} = 8923.595$

iv) $R^2 = \frac{S^2_{xy}}{S_{xx} S_{yy}} = 0.8923$

Q7)

$\Sigma x = 174.13$

$\Sigma y = 197.84$

$\Sigma x^2 = 7545.9$

$\Sigma y^2 = 4747.45$

$S_{xy} = 2176.84$

$S_{xx} = 4789.4221$

$S_{yy} = 1189.2076$

i) $\hat{\beta} = \frac{S_{xy}}{S_{xx}} = 2.20017$

$\therefore \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = -16.8432$

$\therefore y_i = -16.8432 + 2.20017 x_i$

ii) $H_0 \Rightarrow \hat{\beta} = 0$

$\hat{\beta} \neq 0$

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

$= 199.8121$

$t_{cal} = \frac{\hat{\beta} - \beta}{\hat{\sigma}\sqrt{S_{xx}}} = 10.77176$

$\therefore t_{tab} = 2.776$

$\therefore t_{cal} > t_{tab}$

$\therefore$ We reject H_0.

iii) $\rho = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = 0.912128$

Q9)

i) Seems to be negatively correlated.

ii) $S_{xx} = 75$
$S_{yy} = 2482.4$
$S_{xy} = -296$

$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$
$= -0.6860002$
Hence (-)ve correlation.

iii) $\tanh^{-1} r \sim N(\tanh^{-1} \rho, 1/n-3)$

$H_0 : \rho = 0$
$H_1 : \rho_1 \neq 0$

$\frac{\tanh^{-1} r - \tanh^{-1} \rho}{\sqrt{1/n-3}} = Z_{cal}$

$\therefore Z_{cal} = -2.22338$

$\therefore Z_{tab} = Z_{2.5} = 1.96$
$Z_{0.5} = 2.5758$

at 5% LOS Reject H_0
at 1% LOS do not Reject H_0.