Assignment - 2

Page No.:
Date:

que1 >

a) $\overline{X} = 85$ $\overline{Y} = 173.7$

$S_{xx} = 20250$ $S_{yy} = 71000.1$

$S_{xy} = \Sigma xy - n\overline{x}\overline{y}$

$= 182560 - 10 \times \frac{173.57}{10} \times 85$

$S_{xy} = 34915$

$\beta = \frac{S_{xy}}{S_{xx}} = \frac{34915}{20250}$ $\beta = 1.7242$

$\alpha = \overline{Y} - \beta\overline{X}$

$\alpha = 27.143$

$y = \alpha + \beta x$

$y = 27.143 + 1.7242x$

b) Gradient of Regression line $= r\left(\frac{S_y}{S_x}\right)$

$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = 0.003456$

$= 0.003456 \left(\frac{88.81948}{47.43416}\right)$

Gradient of regression line $= 0.006413$

Que 2) ① $\beta = \frac{S_{xy}}{S_{xx}}$ $\alpha = \bar{y}\ \beta\bar{x}$

$\bar{X} = \frac{\Sigma X}{n}$ $\bar{y} = \frac{\Sigma y}{n}$

$\bar{x} = 32.1$ $\bar{y} = 96.875$

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

$S_{xy} = 875.16$

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

$S_{xy} = 301.66$

$\beta = \frac{S_{xy}}{S_{xx}} = \frac{875.16}{301.66}$

$\beta = 2.90115$

$\alpha = 96.875 - 2.90115 \times 32.1$

$\alpha = 3.748$

Equation : $y = \alpha + \beta x$

$y = 3.748 + 2.90115x$

② Slope coefficient of model $= \beta = 2.90115$

95% C.I. $= (\bar{x} - \bar{y}) \pm t_{10}\sqrt{S_p\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

$S_p^2 = 3\cancel{0}5.5203$

$$CI = (96.875 - 32.1) \pm t_{10, 2.5\%} \sqrt{305.5703\left(\frac{1}{12} + \right.}$$

$$C.I = (48.8764, 80.674)$$

iii) Mean IBM $= \mu_0$

$$Se(\hat{\mu}_0) = \sqrt{\left(\frac{1}{n} + \left(\frac{(\bar{x} - x_0)^2}{S_{xx}}\right) \hat{\sigma}^2\right)}$$

$$= 10.59844$$

$$95\%\ C.I = \mu_0 \pm t_{10, 2.5\%}\ Se(\hat{\mu}_0)$$

$$= 115.5405 \pm 23.61443$$

$$= (36.17966, 143.4852)$$

Q3) i) Assumption.

Plot must be normal
population have common values.
the observation are independent.

iii) H_0: mean of difference is same
H_1: The mean is not same.

$$SS_{TOT} = \text{Sum}$$
$$= \sum y^2 - \left(\frac{\sum y}{n}\right)^{*n}$$

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

$$SS_{TOT} = 1116.11$$

$$SS_{Reg} = \left(64^2 + 44^2 + 8^2 + (-13)^2\right) - \frac{103^2}{28}$$

$$= 516.11$$

$$SS_{Res} = SS_{Tot} - SS_{Reg} = 600$$

iii> Anova

$$f = \frac{M_{SS\,Reg}}{M_{SS\,Res}} = \frac{172.04}{25} = 6.88$$

Test Statistics $= 6.88$

$f_{3,\,24,\,5\%} = 3.009$

Hence, we have sufficient evidance to Reject H_0 at 5% LOS.

iv> $\bar{y}_1 = 9.14$

$\bar{y}_2 = 6.299$ $\qquad \bar{y}_{1x} > \bar{y}_{2x} > \bar{y}_{3x} > \bar{y}_{4x}$

$\bar{y}_3 = 1.14$

$\bar{y}_4 = -1.86$

$$\hat{\sigma}^2 = \frac{SS_{Res}}{n} = 25$$

least significant diff. between ~~en~~ of mean is

$$f_{24,\,0.025} \times \hat{\sigma}\sqrt{\frac{1}{7} + \frac{1}{7}} = 2.064 \times \sqrt{25}\sqrt{\frac{1}{7}+\frac{1}{7}}$$

$$= 5.52$$

$\bar{Y}_1 - \bar{Y}_2 = 2.85$
$\bar{Y}_2 - \bar{Y}_3 = 5.15$
$\bar{Y}_3 - \bar{Y}_1 = 3$

$\bar{Y}_1 > \bar{Y}_2 > \bar{Y}_3 > \bar{Y}_4$

Hence proved.

Que 4) i) $\mu_{ij} = \mu + \tau_i + e_{ji}$ $\quad i = 1, 2, 3, \dots k$
$j = 1, 2, 3, \dots, k_i$

$k =$ number of treatment
$k_i =$ number of Responses.
$Y_i =$ Responses from i treatment.

Assumption e_{ij} is iid $(N(0, \sigma^2))$

ii) $n_1 = 7$
$n_2 = 8$
$n_3 = 6$
$n_4 = 7$
$n_5 = 5$
$n_6 = 33$

H_0: Mean claim amount is equal
H_1: Mean claim amount is not equal.

$\sum_{l=1}^{5} \sum j = y^2_i = 1856$

$\sum_{l=1} \sum_{J=1} Y^2_{ij} = 174310$

$C.F = (\sum\sum y_{ij})^2 / n = 104385.94$

$SS_{Tot} = 69930$

$SS_{Res} = 227325.0$

$SS_{Reg} = 64604.37$

$$T.S = \frac{MSS_{Reg}}{MSS_{Res}} = 3.283$$

$$F_{9,\,25,\,0.05} = 2.714$$

We have sufficient evidence to reject H_0

iii) O_i

	3-5	5-8	8-10	10-12	
0-3	47	20	6	6	7[illegible]
4-6	7	20	3	5	4[illegible]
7-9	5	8	15	12	4[illegible]
	54	48	30	23	1[illegible]

E_i

	3-5	5-8	8-10	10-12
0-3	27.41	23.08	14.43	11.07
4-6	15.15	12.75	7.57	6.11
7-9	14.43	12.15	7.54	5.88

Contingency Table:

H_0: Sales are independent on actual Diaper

H_1: —"— dependent on —"—

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

$$T.S = 49.91148$$

$D.O.F = (r-1)(c-1) 2 \times 3 = 6$

$$\chi^2_{0.5} = 12.59$$

we have sufficient evidence to reject H_0

ques) i) The Sample variance are very different from each other.

ii) Mean at $\mu = \frac{\sqrt{24} + \sqrt{13} + \sqrt{25}}{3} = 4.5264333$

var at $m = \frac{\Sigma (x_i - \bar{y})^2}{m} = 0.12124$

iii) $y_{ij} = \mu_i + \tau_i + e_{ij}$

for Anova

$H_0 \quad \tau_i = 0$
$H_1 \quad \tau_i \neq 0$

Rate	mean	variance	y_i^2
1	2.992	0.163	80.532
2	4.827	0.121	209.675
3	5.716	0.186	294.053
4	6.575	0.2	389.05
Total	20.110	0.618	973.373

$$y_i^2 = \left(\sum_{i=1}^{3} y_{ij} \right)^2 = \Sigma n y^2)$$

$$SS_{res} = \frac{\Sigma y_i^2}{3} - y_2^2 = \frac{973.373}{3} - \left(\frac{2r}{\cancel{b}}\right)$$

$$SS_{res} = 21.153$$

Source of variance	D.F	SS	MSS
Rate	3	21.153	0 v
Residual	8	1.226	0.15
Total	11	22.32	

$T.S = 45.630$

$F_{3,8,\alpha} = 7.951.$

Que 6)

x	5	9	2	3	4	1	6	5	4
y	13	120	34	46	24	26	93	45	66

$n = 10.$

$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{661.2}{\sqrt{54.9 \times 8923.6}}$

$r = 0.94466$

fitted Regression eqn

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

$\alpha = 9.2245$

$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = 12.0434$

$\hat{y} = 9.2245 + 12.0434\bar{x}$

$SS_{Tot} = S_{yy} = 8923.6$

$$SS_{Reg} = \frac{S^2_{xy}}{S_{xy}} = 7463.30492$$

$$SS_{Res} = 960.248$$

$$R^2 = 89.23807 \quad \left(\frac{S^2_{xy}}{S_{xy}S_{yy}}\right)$$

Model is good fit as it contain almost 89 % variability in the output.

Q7) i) $S_{xy} = 47.85$
$S_{yy} = 2176.4221$
$S_{xy} = 1189209.$

$$\hat{\beta} = 0.45454$$
$$\hat{\alpha} = 10.78086$$

$$\hat{y} = 10.78086 + 0.45454x$$

ii) $H_0 : \beta = 0$
$H_1 : \beta \neq 0$

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

$$\hat{\sigma}^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S^2_{xy}}{S_{xx}}\right) = 22.20136$$

$$T.S = \frac{\hat{\beta} - 0}{\sqrt{\sigma^2 / S_{xx}}} = 6.67568$$

we Have sufficient evidance to reject H_0

Assumption =

population have common variance
population follows normal distribution

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = 0.91213$$

iv) For mean response.

$$x_0 = 25$$
$$y_0 = \hat{\alpha} + \hat{\beta} x_0 = 22.1534$$

$$S_y(\hat{y}_0) = \sqrt{\left(\frac{1}{n} + \frac{(\bar{x} - x_0)^2}{S_{xx}}\right) \hat{\sigma}^2} = 1.85781$$

$$\frac{\hat{\mu}_0 - \mu_0}{se(\mu)} \sim t_9$$

95% C.I for $\hat{\mu}_0$

$$= \mu_0 \pm t_{9, 2.5\%} . S_{\hat{\mu}_0}$$

$$= (18.84313, 25.4634)$$

Individual Responses

$$= (10.932, 33.37462)$$

v) There is Clearly Some Relationship Between the residuals and the explanatory variable values Hence our assumption of normality might be ...

Ques 8) i) Principal Component   Diag   P.V.

Principal Component	Diag	P.V.
PC_1	0.436	65%
PC_2	0.937	19.5%
PC_3	0.08	11.4%
PC_4	0.0165	2.4%
PC_5	0.072	1.7%
		100%

(ii) As the 3rd principle componant explain 95% C.I. total variance can be futher uesed for binary clasifiel ~~repres~~ regression model

Ques 9) i) $S_{xx} = 75$

$S_{xy} = 2482.64$

$S_{yy} = -296$

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

$r = -0.686$

Hence may know m correlation coffieciant

ii) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$

$\tan^{-1}(r)$ $\tanh^{-1}(\rho) \sim N\left(0, \frac{1}{n}\right)$

$Z_{cal} = -2.07893$

$Z_{tab} = -1.6449$

$Z_{cal} > Z_{tab}$

Hence we have sufficient evidance to reject H_0

iii) $\hat{\beta} = \frac{S_{xy}}{S_{xy}} = -3.94667$

$\hat{\alpha} = 82.10$

$\hat{y} = 82.10 \mp 3.94467\hat{x}$

iv) $R^2 = 47.0598$

Hence model explain 47 % variability

$\therefore$ model is not good fit

v) $x = 1$

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

$\hat{y}_0 = 78.2183$

Que 10

x	0.27	0.36	0.3
y	0.05	0.1	0.08

$S_{xy} = 0.0042$

$S_{yy} = 0.00120$

$S_{xy} = 0.0022$

$SS_{Tot} = 0.00126$

$SS_{Reg} = 0.00125$

$SS_{Res} = 0.00011$

Anova table

Reg	1	0.00115
Res	~~1~~ 2	0.00011
Tot	3	0.00126

$H_0: \beta = 0$
$H_1: \beta \neq 0$

$$T.S = \frac{MSSR}{SSE} = 10.245$$

$$= 161.4$$

Hence we have insufficient evidence to reject H_0