Probability & Stats.

Vishwaa

Sol^1

(1.) $\Sigma x^2 = 92500$, $\bar{x} = 85$
$\Sigma y^2 = 372717$; $\bar{y} = 173.7$
$\Sigma xy = 182560$

$S_{xx} = \Sigma x_i^2 - n(\bar{x})^2 = 20550$
$S_{yy} = \Sigma y^2 - n(\bar{y})^2 = 71000.1$
$S_{xy} = \Sigma xy - n(\bar{x}\,\bar{y}) = 34915$

a) $\beta = \dfrac{S_{xy}}{S_{xx}} = 1.7242$

$\alpha = \bar{y} - \beta\bar{x} \Rightarrow 27.143$

$\hat{y} = 27.143 + 1.7242x$.

b) gradient is no. of hours he spent in ATM upon amount withdrawn.

Q2. $\Sigma x = 385.2$, $\Sigma x^2 = 12666.58$
(i) $\Sigma y = 1162.5$, $\Sigma y^2 = 119026.9$
$\Sigma xy = 38191.41$

$S_{xx} = 301.66$ $S_{yy} = 6409.7125$
$S_{xy} = 875.16$
$\hat{\beta} = 2.901147$ $\hat{\alpha} = 3.74818$
$y = 3.74818 + 2.901147x$

(ii) $\hat{\sigma}^2 = \dfrac{1}{10}\left[S_{yy} - \dfrac{S_{xy}^2}{S_{xx}}\right]$
$= 387.07 45$

$$-2.228 \le \frac{\hat{\beta} - \beta}{\hat{\sigma}/\sqrt{S_{xx}}} \le 2.228$$

$$\hat{\beta} - 2.228 \frac{\hat{\sigma}}{\sqrt{S_{xx}}} \le \beta \le \hat{\beta} + 2.288 \frac{\hat{\sigma}}{\sqrt{S_{xx}}}$$

$$(0.377, 5.425)$$

(iii) $\hat{\mu}_0 = 119.794$

$S_e(\hat{\mu}_0) = 10.5989$

95% confidence Interval of μ_0

$119.794 \pm 2.228(10.598)$

119.794 ± 23.6144

$(\cancel{9.6}\ 96.17956, 143.4084)$

(3.)

(i) Comments on the plot :-

- The centre of distribution differs for all four cities, Thus there's a Prime facie case for suggesting that underlying means are different

- The difference between mean time taken to communicate to Office in peak hours & non peak hours are in order city A > city D [highest, lowest].

<u>assumptions</u> :-

- observations are independent
- The population must have common variance.
- population must be normal.

DATE
PAGE NO.

(iii) H_0 : The mean diff is same for each city
H_a: The mean diff are not same.

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

$$SS_B = \frac{1}{7}\left[64^2 + 44^2 + 8^2 + (-13)^2\right) - \frac{103^2}{28}$$
$$= 516.11$$

$$SS_R = SS_T - SS_B = 600.00.$$

ANOVA

Source	df	SS	MS	F
Treatments	3	516.11	172.04	6.88
residual	24	600.00	25.00	
Total	27	111.611		

Under H_0 = $f = \frac{172.04}{25} = 6.88$

The 5% critical point is 3.009, so we have sufficient evidence to reject H_0.

(iv) Since : $\overline{y}_1 = 9.14$ $\overline{y}_2 = 6.29$ $\overline{y}_3 = 1.14$
$\overline{y}_4 = -1.86$

$\overline{y}_1 > \overline{y}_2 > \overline{y}_3 > \overline{y}_4$; $\hat{\sigma}^2 = \frac{SS_R}{n-k} = 25.$

$t_{24,\ 0.025} * \hat{\sigma}\cdot\sqrt{\left(\frac{1}{7}+\frac{1}{7}\right)} = 2.064 * \sqrt{25}$
$\sqrt{\frac{1}{7}+\frac{1}{7}}$ = 5.52.

$\bar{y}_1 - \bar{y}_2 = 2.85$, $\bar{y}_2 - \bar{y}_3 = 5.15$,
$\bar{y}_3 - \bar{y}_4 = 3$

observing that are 3 difference are less than 5.52, we can say they have no significant difference.

examining $\bar{y}_1 - \bar{y}_3 = 8$

Hence there is a significant diff b/w means 1 & 3.

(4) (i) Already in sol (3), (iii)

(ii) $H_0 : \mu_1 = \mu_2 \cdots \mu_5$
H_a : atleast one of the mean isn't same.

$$\frac{T^2}{N} = \frac{(\Sigma y_1 + \Sigma y_2 + \Sigma y_3 + \Sigma y_4 + \Sigma y_5)^2}{N} =$$

$$104385.9394$$

$$SS_{Reg} = \frac{(354)^2}{7} + \frac{(386)^2}{8} + \frac{(87)^2}{6} + \frac{(645)^2}{7} + \frac{(384)^2}{5} - \frac{T^2}{N}$$

$$= 22325.68917.$$

$SS_{TOT} = 69930.0606$

$SS_{Res} = SS_{TOT} - SS_{Reg}$

$= 47604.37143$

ANOVA.

Sources of variation	df	SS	MS	F
Reg	4	22325.69	5581.4225	3.28
Res	48	47604.3134	1700.154	
Total	33	69930.0606		

So, $F_{4,28}$ Tabulated is : 2.21
and our calculated value is 3.28

So, we have sufficient evidences to reject Ho.

(C.) Ho : They are independent of each other.
Ha: They are dependent.

Paper cleared	3-5	5-8	8-10	10-12
0-3	27.42	23.089	14.43	11.06
4-6	15.15	12.76	7.97	6.14
7-9	14.63	12.15	7.59	5.82

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

$\Rightarrow$ 11.27 + 0.4133 + 4.925 + 3.3204 + 4.384 + 4.1079 + 0.133 + 0.00319 + 6.162 + 1.4175 + 7.234 + 6.56

$\Rightarrow \boxed{39.929}$

DATE

PAGE NO.

$\chi^2_6 = 12.592$

Since $\chi^2_{cal} > \chi^2_{Tab}$; hence we have sufficient evidence to reject H_0.