STAT 319: Probability & Statistics for Engineers & Scientists

Term 181, Quiz # 4

Name: ID #:

Q.No.1: - A multiple regression analysis yields the following partial results.

Source	DF	SS
Regression	4	750
Residual Error	35	500

(a) What is the total sample size? 40

(b) How many independent variables are being considered? 4

(c) Compute and interpret the coefficient of determination.

$$R^2 = \frac{SSR}{SST} = \frac{750}{750+500} = 0.6$$

(d) Estimate the error variance.

$$\hat{\sigma}^2 = \frac{SSE}{n-k-1} = \frac{500}{35} = 14.2857$$

(e) Test the hypothesis that at least one of the regression coefficients is not equal to zero. Let α = .05.

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

$H_1 :$ At least one variable is significant

$$F = \frac{SSR/k}{SSE/(n-k-1)} = \frac{750/4}{500/(35)} = 13.125$$

Reject H_0 if $F_0 > F_{\alpha, v1, v2}$ where $F_{0.05,4,35} \cong 2.65$

We reject H_0 and conclude that at least one variable is significant.

Q.No.2: - The following sample observations have been obtained by a chemical engineer investigating the relationship between the weight of final product and the volume of raw material:

Product	**Volume**	**Weight**	**$(\text{Volume})^2$**	**$(\text{Weight})^2$**	**(Volume)*(Weight)**
1	14	68	196	4624	952
2	23	105	529	11025	2415
3	9	40	81	1600	360
4	17	79	289	6241	1343
5	10	81	100	6561	810
6	22	95	484	9025	2090
7	5	31	25	961	155
8	12	72	144	5184	864
9	6	45	36	2025	270
10	16	93	256	8649	1488
Sum	**134**	**709**	**2140**	**55895**	**10747**

$Y \rightarrow$ Weight of the final product

$X \rightarrow$ Volume of the raw material

$\sum X = 134, \sum Y = 709, \sum X^2 = 2140, \sum Y^2 = 55895$ and $\sum XY = 10747$

(a) Calculate the least squares estimates of the simple linear regression equation for predicting weight of the final product based on volume of raw material.

$$S_{XX} = \sum X^2 - \frac{\left(\sum X\right)^2}{n} = 2140 - \frac{(134)^2}{10} = 344.4$$

$$S_{YY} = \sum Y^2 - \frac{\left(\sum Y\right)^2}{n} = 55895 - \frac{(709)^2}{10} = 5626.9$$

$$S_{XY} = \sum XY - \frac{\left(\sum X\right)\left(\sum Y\right)}{n} = 10747 - \frac{(134)(709)}{10} = 1246.4$$

$$\beta_1 = \frac{S_{XY}}{S_{XX}} = \frac{1246.4}{344.4} = 3.619$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} = \frac{709}{10} - 3.619\left(\frac{134}{10}\right) = 22.405$$

Hence the regression equation is $Y = 22.405 + 3.619X$

(b) Find the error in estimating the product weight if the materials volume is 16 gallons.

$x_0 = 16 \qquad y_0 = 3.619 + 22.405(16) = 80.309$

$e_0 = y_0 - y_0 = 93 - 80.309 = 12.691$

(c) Using a 95% confidence level, estimate an observation of the product weight if the materials volume is 16 gallons.

$x_0 = 16 \qquad y_0 = 3.619 + 22.405(16) = 80.309$

95% confidence interval for a product:

$\alpha = 0.05, \quad \frac{\alpha}{2} = 0.025, \quad v = n - 2 = 10 - 2 = 8$

$t_{\frac{\alpha}{2},v} = t_{0.025,8} = 2.306$

$$y_0 \pm t_{\frac{\alpha}{2},n-2}\sqrt{\sigma^2\left(1 + \frac{1}{n} + \frac{\left(x_0 - \bar{x}\right)^2}{S_{XX}}\right)}$$

$$80.309 \pm 2.306\sqrt{139.522\left(1 + \frac{1}{10} + \frac{(16 - 13.4)^2}{344.4}\right)}$$

$[51.487, 109.131]$

(d) At 5% level of significance, test the hypothesis that the regression line passes through origin.

$H_0: \beta_0 = 0$; $H_1: \beta_0 \neq 0$

Test Statistic:

$$T = \frac{\hat{\beta}_0 - \beta_{00}}{\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\overline{X}^2}{S_{XX}}\right]}} = \frac{22.405 - 0}{\sqrt{139.522\left(\frac{1}{10} + \frac{13.4^2}{344.4}\right)}} = 2.406$$

Decision Rule: Reject H_0 if $|T_0| > T_{\frac{\alpha}{2}, n-2}$

Critical Value: $T_{\frac{\alpha}{2}, n-2} = T_{0.025,8} = 2.306$

Decision: As $|T_0| > 2.306$ so we reject H_0.

Conclusion: The regression line does not pass through the origin.