

Question one (6 Marks)

Consider the following regression model

$$Y = 1 + \beta_0 e^{\beta_1 x} \varepsilon$$

1. Write the above model in linearized form (2 Marks).

2. Consider the multiple linear regression model; (4 Marks)

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i, \quad i = 1, 2, \dots, n.$$

a. Prove that the regression coefficient **vector b** is unbiased estimator of **B**.

b. Derive variance covariance matrix of regression coefficients.

Question two (11 Marks)

In a study of the effect of a dietary component (**X**) on plasma lipid composition (**Y**), the following data were obtained on a sample of 12 experimental animals

| | | | | | | | | | | | | |
|---|----|----|----|----|----|----|----|----|----|----|----|----|
| X | 16 | 18 | 24 | 35 | 45 | 32 | 40 | 42 | 28 | 42 | 20 | 20 |
| Y | 37 | 39 | 46 | 54 | 66 | 52 | 58 | 64 | 46 | 62 | 39 | 38 |

1) Fit the linear regression model and obtain the estimated regression function? (1 Mark)

2) Obtain the residual for the plasma lipid composition when the dietary component **X** = 35. (1 Mark)

3) Obtain a 99% prediction interval for a expected plasma lipid composition when the dietary component **X** = 35 .(1 Mark)

4) Perform the F test for the significant of the regression model. Use $\alpha = 0.01$. (2 Marks)

H_0 :

VS

H_1 :

The test-statistic is:

Decision:

Conclusion:

5) Test at 0.05 level of significance whether the simple linear regression model suffers from lack of fit . (2 Marks)

H_0 :

VS

H_1 :

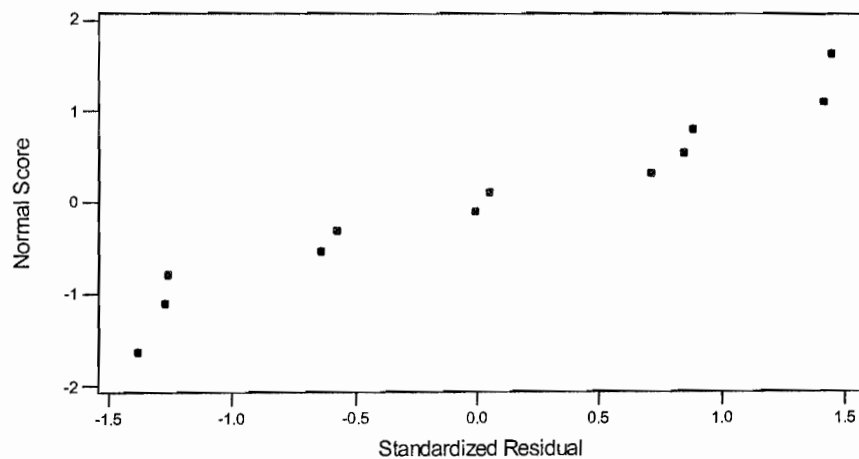
The test-statistic is:

Decision:

Conclusion:

6) Compute r and **comment** on its value . (2 Marks)

7) Some residual model diagnostics were obtained for this fit. Comment on each chart separately stating what does each plot indicates. (2 Marks)



$H_0 :$

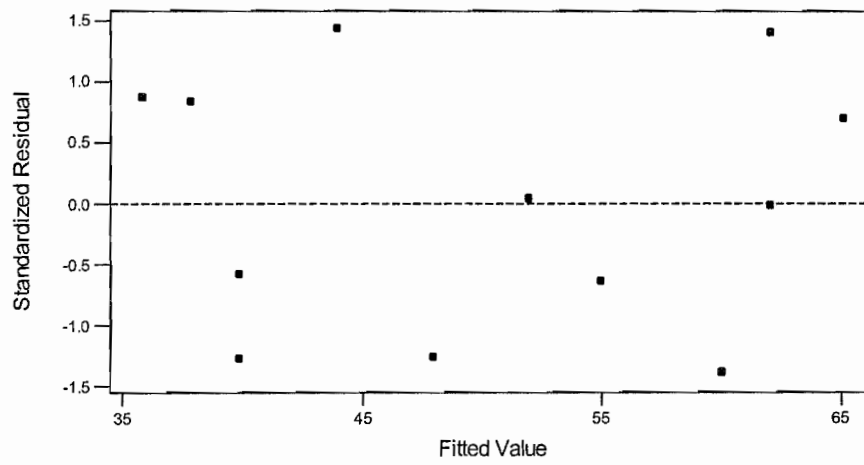
VS

$H_1 :$

Decision:

Conclusion:

b)



$H_0 :$

VS

$H_1 :$

Decision:

Conclusion:

Question Three (16 Marks)

A business administrator wished to study the relation between **employees satisfaction (Y)** and **employee's age (X_1 , in years)**, **level of responsibility (X_2 , an index)**, and **employee's ambition level (X_3 , an index)**. The administrator randomly selected **46** employees and collected the data presented in satisfaction.mpj.

1. Fit the multiple regression model containing all three predictors variables.(1 mark).
2. Determine R^2 and **explain** what does this value indicate .(2 mark).
3. Obtain the analysis of variance table that decomposes the regression sum of squares associated with X_2 ; with X_1 given X_2 ; and X_3 given X_1 and X_2 . (3 Marks)

| Source | Sum of Squares | d. f. | MS |
|------------|----------------|-------|----|
| Regression | | | |
| | | | |
| Residual | | | |
| Total | | | |

7. Obtain a **98%** confidence interval for expected satisfaction for a 45 years old employee with degree of responsibility = **48**, ambition index = 3 . (1 mark).

8. For the same values of regressors which is wider :the **98%** prediction interval for a employees satisfaction or the **98%** confidence interval of the mean response? Why ? (2 Marks)

Question four (7 Marks)

Given the following information for $n=15$ observations, use the stepwise regression method to find the best subset of predictor variables to predict the response value. In each step, find the F-statistic, decision and conclusion (use $F_{in}=4$, $F_{out}=3.5$).

| Regressors in Model | b_0 | b_1 | b_2 | b_3 | SSE |
|--|--------|-------|-------|--------|--------|
| Non | 58.67 | | | | 4505.3 |
| X₁ | 124 | -1.65 | | | 1883.3 |
| X₂ | 207 | | -2.95 | | 1773.1 |
| X₃ | 63 | | | -45.3 | 1776.1 |
| X₁, X₂ | 196.1 | -1.02 | -1.93 | | 1092.8 |
| X₁, X₃ | 158.82 | -.92 | | -27.79 | 1373.6 |
| X₂, X₃ | 200 | | -1.66 | -25.4 | 1436.9 |
| X₁, X₂, X₃ | 195 | -.89 | -1.60 | -9.2 | 1060.4 |

