

1. State any two assumptions under simple linear regression [2 marks]
 2. A departmental store gives trainings to its salesmen followed by a test to consider whether it should terminate the service of any of the salesman who does not qualify in the test. The data below gives the test scores and sales made by 9 salesmen during certain period.
- | | | | | | | | | | |
|--------------------|----|----|----|----|----|----|----|----|----|
| Test Scores(X) | 14 | 19 | 24 | 21 | 28 | 22 | 15 | 20 | 19 |
| Sales(Y) | 31 | 36 | 48 | 37 | 50 | 45 | 33 | 41 | 39 |
- (i) Test the following hypothesis at $\alpha = 0.05$. $H_0 : \beta_0 = 0$ Vs $H_1 : \beta_0 > 0$ and $H_0 : \beta_1 = 0$ Vs $H_1 : \beta_1 < 0$ [4 marks]
 - (ii) Determine the ANOVA table for the straight line regression of the sales on the test scores at $\alpha = 0.05$ [4 marks]
 - (iii) Determine the R^2 and interpret it [2 mark]
 - (iv) Suggest whether the termination of low score is justified? [1 mark]
3. The particular situation matrix X given as $X' = \begin{bmatrix} 3 & 9 & 7 & 5 & 6 \\ 20 & 33 & 25 & 30 & 24 \end{bmatrix}$ and Y is given as $Y' = \begin{bmatrix} 10 & 20 & 17 & 12 & 11 \end{bmatrix}$. Obtain $(X'X)^{-1}$ and $X'Y$ hence give the least square estimates of β_i and its covariance matrix [7 marks]
 4. State the property of kernel density [3 marks]
 5. Statistician is interested in regressing the returns (Y) of business on the number of branches (X_1) and the age of business X_2 . The statistician collected the relevant data from 10 big companies as shown in the table below

Y	61	47	75	63	79	75	67	47	71	84
X_1	19	19	10	17	19	12	18	10	15	15
X_2	11	4	8	12	5	12	6	14	13	11

Write a well commented R programme that does the following

- (i) Reads in data [3 marks]
- (ii) Estimates the parameters β_1 and β_2 in the regression model $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + e$ where $e \sim N(0, \sigma^2)$ [3 marks]
- (iii) Estimates the parameters a , b and c in the regression model $Y = a + b * x_1 + \exp(c * x_2) + e$ where $e \sim N(0, \sigma^2)$ [3 marks]
- (iv) Determine the p-values of the estimated parameters in the part (ii) and (iii) above [3 marks]