CHAPTER-2: MULTIPLE LINEAR REGRESSION

Multiple Linear Regression:-

In the Chocolate company problem some other regression variable could be number of sales person, this is a motivating example.

– More than one regression variables, say $k-1$ regression variables.

Model: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{k-1} x_{i,k-1} + \epsilon_i$; $i = 1(1)n$

This model is linear of unknown parameters $\beta_0, \beta_1, \ldots, \beta_{k-1}$.

Assumption:- $\epsilon_i \overset{ind}{\sim} N(0, \sigma^2)$; Vector notations;

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{k-1} \end{pmatrix} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

vector of obsn. — Vector of parameters — vector of errors

$n \times k$ matrix →
$$\underset{n\times k}{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k-1} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k-1} \\ \vdots & & & & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk-1} \end{pmatrix}$$

observations (1^{st}) on regression $(1, \ldots, k-1)$ is known.

Since, data with us is ; $(Y_i, X_{i1}, X_{i2}, \ldots, X_{ik-1})$ $i = 1(1)n$,

Using matrix notation, we write

Model:- $\underset{n\times 1}{Y} = \underset{n\times k}{X}\underset{k\times 1}{\beta} + \epsilon_{n\times 1}$

Estimation of Model parameters:

LSM (Least square method) determines the parameters by minimizing

$$SS_{Res} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \qquad \hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_{k-1} X_{k-1}$$

$$\boxed{SS_{Res} = \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_{k-1} x_{ik-1})^2}$$

Alt. Method:

In matrix form: $\underset{\sim}{e} = (\underset{\sim}{Y} - \hat{\underset{\sim}{Y}})$; where $\underset{\sim}{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$, $\underset{\sim}{Y}, \hat{\underset{\sim}{Y}}$.

$$SS_{Res} = \sum_{i=1}^{n} e_i^2 = e'e = (Y - \hat{Y})'(Y - \hat{Y}) \qquad [\text{Since } \hat{Y} = X\hat{\beta}]$$

$$= (Y - X\hat{\beta})'(Y - X\hat{\beta}) = Y'Y - \underbrace{Y'X\hat{\beta} - \hat{\beta}'X'Y}_{(1\times 1)\text{ scalar quantity and equal}} + \hat{\beta}'X'X\hat{\beta}$$

$$\boxed{SS_{Res} = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta}}$$

Normal equations:-

$$\frac{\partial SS_{Res}}{\partial \hat{\beta}_0} = 0 \Rightarrow \sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \cdots - \hat{\beta}_{k-1} x_{ik-1}) = 0$$

$$\Rightarrow \sum_{i=1}^{n} e_i = 0$$

$$\frac{\partial SS_{Res}}{\partial \hat{\beta}_1} = 0 \Rightarrow \sum e_i x_{i1} = 0$$

$$\vdots$$

$$\frac{\partial SS_{Res}}{\partial \hat{\beta}_{k-1}} = 0 \Rightarrow \sum e_i x_{ik-1} = 0$$

k constraints

k normal equations (independent)

Solving these k normal equations, we can have k unknown parameters.

Alt. Method: (In Matrix form):

$$\frac{\partial SS_{Res}}{\partial \hat{\beta}} = 0$$

$\Rightarrow -2X'Y + 2X'X\hat{\beta} = 0$

$\Rightarrow$

$\Rightarrow \hat{\underset{\sim}{\beta}} = (X'X)^{-1}(X'Y)$ = Least square estimation of k unknown parameters for multiple linear regression model.

Statistical properties of LSE:-

- $\hat{\beta}$ is an UE of β.

$\rightarrow$ $\hat{\beta} = (X'X)^{-1}(X'Y)$ $\quad \left[\because Y = (X\beta + \epsilon)\right]$

$$E(\hat{\beta}) = E\left[(X'X)^{-1}X'Y\right]$$
$$= E\left[(X'X)^{-1}X'(X\beta + \epsilon)\right]$$
$$= E\left[(X'X)^{-1}(X'X)\beta\right] + E\left[(X'X)^{-1}X'\epsilon\right]$$
$$= \beta + 0 = \beta; \quad E(\epsilon) = 0$$

[Note: $(X'X)$ is known as variance-covariance matrix which is symmetric]

$\rightarrow$ $V(\hat{\beta}) = V\left((X'X)^{-1}X'Y\right) = (X'X)^{-1}X'I\sigma^2X(X'X)^{-1}$

$= \sigma^2(X'X)^{-1}$ is known as variance-covariance matrix of $\hat{\beta}$.

Estimation of σ^2:

$$SS_{Res} = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \quad ; \quad \hat{\beta} = (X'X)^{-1}X'Y$$
$$= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'Y$$
$$= Y'Y - \hat{\beta}'X'Y$$
$$= \sum_{i=1}^{n} e_i \quad ; \quad e_i \sim N(0, \sigma^2). \Rightarrow \frac{e_i}{\sigma} \sim N(0,1)$$

$\Leftrightarrow$ Matrix representation $\equiv$ standard representation

You can choose only $(n-K)$ e_i's independently. Since K constraints are there which are satisfied by e_i's (in normal equations).

SS_{Res} has $(n-K)$ df. $\quad \frac{e_i^2}{\sigma^2} \sim \chi_1^2$

(In SLR, it was $(n-2)$ d.f. for SS_{Res})

$$\boxed{MS_{Res} = \frac{SS_{Res}}{n-K}} \qquad \boxed{\frac{SS_{Res}}{\sigma^2} = \frac{\sum e_i^2}{\sigma^2} \sim \chi^2_{n-k}}$$

$E(MS_{Res}) = \sigma^2$ $\Rightarrow$ MS_{Res} is an UE for σ^2 (similar to simple linear regression (SLR).

$$SS_T = \sum_{1}^{n}(Y_i - \bar{Y})^2 = \sum Y_i^2 - n\bar{Y}^2$$

Data: $(Y_i, X_{i1}, \ldots, X_{ik-1})$

SS_T has df $n-1$ since $\sum(Y_i - \bar{Y}) = 0$

$$SS_{Reg} = SS_T - SS_{Res} = \hat{\beta}'X'Y - n\bar{Y}^2$$

SS_{Reg} has $(K-1)$ DF. $\Rightarrow$ $\because$ DF for SS_{Res} is $(n-K)$.

$$\left[\begin{aligned} SS_T - SS_{Res} &= \\ &= \sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2 - (Y'Y - \hat{\beta}'X'Y) \\ &= \cancel{Y'Y} - n\bar{Y}^2 - \cancel{Y'Y} + \hat{\beta}'X'Y \\ &= \hat{\beta}'X'Y - n\bar{Y}^2 \end{aligned}\right]$$

[df = DF = degree of freedom]

$SS_T = SS_{Reg} + SS_{Res}$

where SS_T: Total variability in the response variable

SS_{Reg}: The amount of variability explained by Reg. model

SS_{Res}: Unexplained variability

Target:- We want to maximize SS_{Reg} and reduce SS_{Res} using our reg. model.

Test for significance of Regression model:- (Global test) [F-test to check linearity]

If there is linear relationship between the response and any one of the regression variable $x_1, x_2, \dots, x_{k-1}$.

$H_0: \beta_1 = \beta_2 = \dots = \beta_{k-1} = 0$

Vs. $H_1: \beta_j \neq 0$ for at least one j.

$$SS_T = SS_{Reg} + SS_{Res}$$

DF: $(n-1) = (n-k) + (k-1)$

$$\left.\begin{array}{l} \dfrac{SS_{Reg}}{\sigma^2} \sim \chi^2_{k-1} \\ \dfrac{SS_{Res}}{\sigma^2} \sim \chi^2_{n-k} \end{array}\right\} \text{independently}$$

By the definition of F-statistic,

$$F = \frac{SS_{Reg}/k-1}{SS_{Res}/n-k} \sim F_{k-1,n-k}$$

$$F = \frac{MS_{Reg}}{MS_{Res}}$$

[Higher value of observed F suggests that at least one $\beta_j \neq 0$]

$$MS_{Res} = \frac{SS_{Res}}{n-k}$$

$$E(MS_{Res}) = \sigma^2$$

$$E(MS_{Reg}) = \sigma^2 + \frac{\beta^{*\prime} X_c' X_c \beta^*}{(k-1)\sigma^2}$$

where

$$\beta^* = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{k-1} \end{pmatrix}$$

$$X_c = \begin{pmatrix} x_{11} - \bar{x} & \dots & x_{1k-1} - \bar{x} \\ \vdots & & \\ x_{n1} - \bar{x} & \dots & x_{nk-1} - \bar{x} \end{pmatrix}$$

We reject $H_0: \beta_1 = \beta_2 = \dots = \beta_{k-1} = 0$ if $F > F_{\alpha, k-1, n-k}$.

ANOVA table easily one can do. (See next Page)

Test on individual regression coefficient (Partial/Marginal test):-

Test the significance of x_j in the presence of other regressors in the model.

$H_0: \beta_j = 0$ Vs $H_1: \beta_j \neq 0$

$$\hat{\beta} = (X'X)^{-1} X'Y$$

$$\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$$

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2 (X'X)^{-1}_{jj}}} \sim N(0,1)$$

Test statistic:-

$$t = \frac{\hat{\beta}_j}{\sqrt{MS_{Res} (X'X)^{-1}_{jj}}} \sim t_{n-k} \text{ under } H_0.$$

[Under $H_0 (\beta_j = 0)$]

$H_0: \beta_j = 0$ is rejected if $|t| > t_{\alpha/2, n-k}$.

Partial test is performed once from the test of significance of Regression model (above one) (global test) it is confirmed that there is at least one β_j non-zero

[P.T.O. – Part 2]

(Continued: Global Test) ANOVA TABLE

Source of Variation	DF	SS	MS	F
Reg	$k-1$	SS_{Reg}	$MS_{Reg} = \frac{SS_{Reg}}{k-1}$	$F = \frac{MS_{Reg}}{MS_{Res}}$
Res	$n-k$	SS_{Res}	$MS_{Res} = \frac{SS_{Res}}{n-k}$	
Total	$n-1$	SS_T		

$F \sim F_{k-1, n-k}$

We reject H_0 if $F_{calculated} > F_{\alpha; k-1, n-k}$ (tabulated value from statistical table)

Once we determine that there is linear relationship between response and regression variable from ANOVA table (i.e., H_0 is rejected) then there is at least one regression variable which has significant contribution to the response variable.

Next question is: Which regression variable has significant contribution?

[P.T.O. – Part-2]

Confidence interval on Regression Coefficients:

$$\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{k-1} \end{pmatrix}$$

UE (point) of β is $\hat{\beta} = (X'X)^{-1}X'Y$

$E(\hat{\beta}) = \beta$ and $V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$.

$$\hat{\beta}_i \sim N\left(\beta_i, \sigma^2 (X'X)^{-1}_{ii}\right)$$

$$\frac{\hat{\beta}_i - \beta_i}{\sqrt{MS_{Res}(X'X)^{-1}_{ii}}} \sim t_{n-k} \; ; \; P\left\{\left|\frac{\hat{\beta}_i - \beta_i}{\sqrt{MS_{Res}(X'X)^{-1}_{ii}}}\right| \le t_{\frac{\alpha}{2}, n-k}\right\} = 1-\alpha.$$

Thus, $100(1-\alpha)\%$ CI for the parameter β_i is

$$\hat{\beta}_i - t_{\frac{\alpha}{2}, n-k}\sqrt{MS_{Res}(X'X)^{-1}_{ii}} \le \beta_i \le \hat{\beta}_i + t_{\alpha/2, n-k}\sqrt{MS_{Res}(X'X)^{-1}_{ii}}.$$

Confidence Interval on Mean response at a particular point, say, $x_0 = (1, x_{01}, x_{02}, \ldots, x_{0k-1})$:-

$$E(Y|x_0) = \overset{1\times k}{x_0}\ \overset{k\times 1}{\beta}$$

An unbiased estimator of $E(Y|x_0)$ is $\hat{y}_0 = x_0\hat{\beta}$; since

$$E(x_0\hat{\beta}) = x_0 E(\hat{\beta}) = x_0\beta.$$

$$V(\hat{y}_0) = V(x_0\hat{\beta}) = x_0 V(\hat{\beta}) x_0'$$
$$= x_0 \sigma^2 (X'X)^{-1} x_0'$$

$$\frac{\hat{y}_0 - x_0\beta}{\sqrt{x_0\, MS_{Res}(X'X)^{-1}x_0'}} \sim t_{n-k}$$

$\therefore$ $100(1-\alpha)\%$ confidence interval on mean response at the point x_0 is

$$x_0\hat{\beta} + t_{\alpha/2,\, n-k}\sqrt{MS_{Res}\, x_0 (X'X)^{-1} x_0'}$$

Prediction of New Observation:

$$x_0 = (1, x_{01}, \ldots, x_{0k-1})$$

A point estimator of the future observation y_0 at the point x_0 is

$$\hat{y}_0 = x_0\hat{\beta}.$$

New random variable, $\psi = \hat{y}_0 - y_0$, $E(\psi) = 0$

$$V(\psi) = V(\hat{y}_0 - y_0)$$
$$= \sigma^2\left(1 + x_0 (X'X)^{-1} x_0'\right)$$

$$\frac{\hat{y}_0 - y_0}{\sqrt{MS_{Res}\left(1 + x_0 (X'X)^{-1}x_0'\right)}} \sim t_{n-k}.$$

$\left[\because V(y_0) = \sigma^2\right.$ and $\hat{y}_0$ and y_0 are independent. $\left.\right]$

Thus, $100(1-\alpha)\%$ PI for y_0 is

$$x_0\hat{\beta} \pm t_{\alpha/2,\, n-k}\sqrt{MS_{Res}\left(1 + x_0 (X'X)^{-1} x_0\right.}$$

Some useful results:

If a is a $k\times 1$ vector of constants, A is a $k\times k$ matrix of constants and y is a $k\times 1$ random vector with mean μ and non-singular variance-covariance matrix V. Then (i) $E(a'y) = a'\mu$; (ii) $E(Ay) = A\mu$; (iii) $Var(a'y) = a'Va$; (iv) $Var(Ay) = AVA'$; (v) $E(y'Ay) = tr(AV) + \mu' A\mu$.

Properties of Least-square estimates – Multiple variable case

DEFINITION: An estimator $\hat{\theta}$ is called the Best Linear Unbiased Estimator (BLUE) for θ if

1. $\hat{\theta}$ is a linear combination of sample observations.
2. $Var(\hat{\theta}) \leq Var(\hat{\theta}')$, where $\hat{\theta}'$ is any other estimator which is unbiased and $\hat{\theta}$ is also an UE.

We simply restrict our set of estimators to those which are linear in data and unbiased. Among all these estimators, we pick the one with minimum variance and we call it BLUE.

Theorem (Gauss-Markov theorem): The Gauss-Markov theorem states that the ordinary least square estimator of β is the Best Linear Unbiased Estimator (BLUE).

Proof: From OLS estimation, we have $\hat{\beta} = (X'X)^{-1}X'Y$

$$E(\hat{\beta}) = \beta \text{ and } Var(\hat{\beta}) = \sigma^2 (X'X)^{-1}.$$

[**We need to show:** For MLR model, $Y = X\beta + \epsilon$ with $E(\epsilon) = 0$ and $V(\epsilon) = \sigma^2 I$, the LSEs are unbiased and minimum variance when compared to all other unbiased estimators that are linear combinations of y_is. Thus, LSEs are BLUE.]

- We consider that $\hat{\beta}$ to be the best which minimizes the variance for any linear combination of the estimated coefficients, $l'\hat{\beta}$. [i.e., $l_0\hat{\beta}_0 + l_1\hat{\beta}_1 + l_2\hat{\beta}_2 + \cdots + l_{p-1}\hat{\beta}_{p-1}$]

Now, $Var(l'\hat{\beta}) = \sigma^2 l'(X'X)^{-1} l$ = scalar.

Let $\tilde{\beta}$ be another unbiased estimator of β which is a linear combination of data. Our goal, then is to show that

$$Var(l'\tilde{\beta}) \geq \sigma^2 l'(X'X)^{-1} l.$$

We use the fact that any other estimator of β can be written as $\tilde{\beta} = \left[(X'X)^{-1}X' + \underset{p\times n}{B}\right]Y + b_{0\,p\times 1}$

$E(\tilde{\beta}) = \beta + BX\beta + \beta_0$. However, $\tilde{\beta}$ is assumed to be biased, hence $b_0 = 0$ and $BX = 0$. Similarly, it can be shown that

$$Var(\tilde{\beta}) = \sigma^2\left[(X'X)^{-1}X' + BB'\right].$$

Since BB' is positive semi-definite, we can see that $l'BB'l = (B'l)'(B'l) \geq 0$.

Hence $\hat{\beta}$ is the BLUE.

NOTE: MLEs for the model parameters are the LSEs of the LR (MLR) model. (PROOF!)

Example of MULTIPLE LINEAR REGRESSION

Example: Consider the following data in the Table.

X_1	X_2	Y
1	8	6
4	2	8
9	-8	1
11	-10	0
3	6	5
3	-6	3
5	0	2
10	-12	-4
2	4	10
7	-2	-3
6	-4	5

(a) Using LSM, estimates the β's in the MLR model.

(b) Write down the ANOVA table. Using $\alpha = 0.05$, test to determine if overall regression is statistically significant (GLOBAL TEST).

(c) Calculate R^2.

(d) Calculate the estimated variance of $\hat{\beta}$.

(e) What does X_2 contribute, given that X_1 is already in the regression?

(f) How useful is the regression using X_1 alone?

(g) Find the variance of the predicted value of Y for the point $x_1 = 3$ and $x_2 = 5$.

Solution: (a) $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \underset{11\times1}{\epsilon}$; where ϵ is a random error vector and

$$\underset{11\times1}{Y} = \begin{pmatrix} 6\\8\\1\\0\\5\\3\\2\\-4\\10\\-3\\5 \end{pmatrix}_{11\times1} \quad \text{and} \quad X_{11\times3} = \begin{array}{c} \begin{array}{ccc} & X_1 & X_2 \end{array} \\ \begin{pmatrix} 1 & 1 & 6\\ 1 & 4 & 8\\ 1 & 9 & 1\\ 1 & 11 & 0\\ 1 & 3 & 5\\ 1 & 3 & 3\\ 1 & 5 & 2\\ 1 & 10 & -4\\ 1 & 2 & 10\\ 1 & 7 & -3\\ 1 & 6 & 5 \end{pmatrix}_{11\times3} \end{array} \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0\\ \beta_1\\ \beta_2 \end{pmatrix}_{3\times1}$$

and $Y = X\beta + \epsilon$

LS estimates of β_0, β_1 and β_2 are $\hat{\beta} = (\underset{3\times11}{X'}\underset{11\times3}{X})^{-1}\underset{3\times11}{X'}\underset{11\times1}{Y}$.

$$\hat{\beta} = (X'X)^{-1}X'Y = \begin{pmatrix} 11 & 66 & -22\\ & 506 & -346\\ & & 484 \end{pmatrix}^{-1} \begin{pmatrix} 33\\85\\142 \end{pmatrix} = \begin{pmatrix} 4.3705 & -0.849 & -0.4086\\ & 0.169 & 0.8222\\ & & 0.0422 \end{pmatrix} \times \begin{pmatrix} 33\\85\\142 \end{pmatrix}$$

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} 14\\ -2\\ -\frac{1}{2} \end{pmatrix}.$$

$\Rightarrow \hat{\beta}_0 = 14, \hat{\beta}_1 = -2, \hat{\beta}_2 = -0.5$

Fitted regression model is: $\hat{Y} = 14 - 2X_1 - 0.5X_2$.

(b) Table of fitted values and residuals: Calculate: SS_{Res}, SS_T, SS_{Reg}.

X_1	X_2	Y	$\hat{Y} = 14 - 2X_1 - 0.5X_2$	$e = Y - \hat{Y}$
1	8	6	8	-2
4	2	8	5	3
⋮	⋮	⋮	⋮	⋮

$\bar{Y} = 3$

$n = 11$

$\sum_1^{11} Y_i^2 = 289$

Total (DF) = 10 since $\sum_1^{11}(Y_i - \bar{Y}) = 0$

Reg (DF) = 2

Res (DF) = 8 = 11 − 3 since $\begin{cases} \sum e_i = 0 \\ \sum e_i x_1 = 0 \\ \sum e_i x_2 = 0 \end{cases}$

$$SS_{Res} = \sum_1^{11} e_i^2 = 68.$$

$$SS_{Total} = \sum_1^{11} (y_i - \bar{y})^2 = \sum Y_i^2 - n\bar{Y}^2 = 289 - 11 \times 9 = 190.$$

$$SS_{Reg} = SS_{Total} - SS_{Res} = 190 - 68 = 122.$$

ANOVA Table:

Source of Variation	DF	SS	$MS = \frac{SS}{DF}$	$F = \frac{MS_{Reg}}{MS_{Res}}$
Regression	2	122	61	7.17
Residual	8	68	8.5	
Total	10	190		

Now, we test the following: $H_0: \beta_1 = \beta_2 = 0$ (no linear relationship)

vs. $H_1: \beta_i \neq 0$ for at least one i (linear relationship)

Here $F_{observed} = 7.17 > F_{0.05, 2, 8} = 4.46$ (from statistical table)

So, we reject $H_0: \beta_1 = \beta_2 = 0$ and we use the fitted equation: $\hat{Y} = 14 - 2X_1 - 0.5X_2$ (there is linear relationship)

Thus, the fitted equation is statistically significant.

(c) $R^2 = \frac{SS_{Reg}}{SS_T} = \frac{122}{190} = 64.21\%$.

Thus, the regression model (fitted) explains 64.21% of the total variability in the target variable (Y_i) using two causal variables (X_1 and X_2).

(d) $E(\hat{\beta}) = \beta$, $V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$; we know $E(MS_{Res}) = \sigma^2$.

Estimated variance of $\hat{\beta}$ $= \hat{V}(\hat{\beta}) = MS_{Res}(X'X)^{-1}$

$$= 8.5 \times \begin{pmatrix} 4.3705 & -0.849 & -0.4086 \\ & 0.169 & 0.8222 \\ & & 0.0422 \end{pmatrix}$$

$V(\hat{\beta}_i) = \sigma^2 (X'X)^{-1}_{ii}$

Estimated variance of $\hat{\beta}_1 = MS_{Res}(X'X)^{-1}_{11}$
$= 8.5 \times 0.169 = 1.43$

Estimated variance of $\hat{\beta}_2 = 8.5 \times (X'X)^{-1}_{22} = 8.5 \times 0.0422 = 0.35$

(e) $H_0: \beta_2 = 0$ Vs. $H_1: \beta_2 \neq 0$ (Partial test)

$$t = \frac{\hat{\beta}_2}{\sqrt{MS_{Res}(X'X)^{-1}_{22}}} = \frac{-0.5}{\sqrt{0.35}} = -0.83$$

Here, $H_0: \beta_2 = 0$ is accepted since $|t| \not> t_{0.025,8} = 2.306$.

So, X_2 does not have any contribution in the regression model in the presence of X_1.

Let us also test: How much X_1 contribute given that X_2 is already in the regression?

$\rightarrow$ $H_0: \beta_1 = 0$ Vs. $\beta_1 \neq 0$

$$t = \frac{\hat{\beta}_1}{\sqrt{MS_{Res}(X'X)^{-1}_{11}}} = \frac{-2}{\sqrt{1.4365}} = 1.668$$

$$|t| \not> t_{0.025,8} = 2.306.$$

We accept $H_0: \beta_1 = 0$. Then, X_1 is not significant in the model in the presence of X_2.

This is an example of : GLOBAL test says X_1, X_2 are significant.

PARTIAL test says neither $X_1 (X_2)$ is significant in the presence of $X_2 (X_1)$.

This example explains the problem of MULTICOLLINEARITY.

(f)

X_1	Y	$\hat{Y}$	e
1	6	8.135	−2.135
4	8	5.054	2.946
⋮	⋮	⋮	⋮

This is the problem of SLR. We get using OLS estimates:

$\hat{Y} = 9.162 - 1.027 X_1$
$= \hat{\beta}_0 + \hat{\beta}_1 X_1$.

$SS_{Res} = \sum_1^{11} e_i^2 = 74,$

$SS_{Total} = \sum_1^{11} (y_i - \bar{y})^2 = 196.$

ANOVA Table

Source of variation	DF	SS	MS	F
Regression	1	116	116	14.15
Residual	9	74	8.2	
Total	10			

$F_{0.05,1,9} = 5.12$ (from statistical table)

So, the fitted model is significant.

$R^2 = \frac{116}{190} = 61.05\%$

This suggests X_1 is more capable to explain the variability in Y in compared to X_2.

(g) $x_0 = (1, x_{01}, x_{02}) = (1, 3, 5)$

Point estimator of Y: $\hat{y}_0 = x_0 \hat{\beta}$

$$\begin{aligned} Var(\hat{y}_0) &= x_0 V(\hat{\beta}) x_0' \\ &= x_0 \sigma^2 (X'X)^{-1} x_0' \\ &= x_0 MS_{Res} (X'X)^{-1} x_0' \\ &= \underset{1\times3}{(1, 3, 5)} \left[8.5 \begin{pmatrix} 4.37 & -0.849 & -.4086 \\ & 0.169 & .0822 \\ & & .0422 \end{pmatrix} \right]_{3\times3} \underset{3\times1}{\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}} \end{aligned}$$

Estimated variance of predicted value of y at $x_1 = 3$ and $x_2 = 5$. $= 1.95$.