BT307 Quiz

Total Marks: 10 Time: 45 min

Name: **Roll No.:**

Q1. A simple linear regression model is usually written as $\hat{y}_i = bx_i + a$. Using the matrix notation, I have written the same model as $\hat{\mathbf{y}} = \mathbf{Z}\mathbf{y}$. Here, **y** is the vector for the dependent variable data. $\hat{\mathbf{y}}$ is the vector for the fitted or regressed values of **y**. The data matrix for this regression is **X**. Express **Z** in terms of **X**. You **MUST** show a few lines of your derivation.

Ans 1: In Linear regression, $\hat{\vec{y}} = X\vec{B}$. By the least square method, $\vec{B} = (X^T X)^{-1} X^T \vec{y}$.

$\therefore \hat{\vec{y}} = X\vec{B} = X(X^T X)^{-1} X^T \vec{y} = Z\vec{y}$

$\therefore \underline{Z = X(X^T X)^{-1} X^T}$ Ans

Q2. We have performed linear regression, considering *x* as a predictor for the following data. The outcome or dependent variable is *y*. Calculate the TSS of this regression. **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0.

x	1	2	3	5
y	2	6	8	**M**

Ans 2:

Let, $M = 0$.

$TSS = \sum (y_i - \bar{y})^2$; For $M = 0$, $\bar{y} = 4$

$\therefore \underline{TSS = 40}$ Ans

Final answer will depend on Roll No.

Q3. You fitted $y = b_1 x + b_2 x^2 + a$ to a data set. The population regression coefficients corresponding to b_1, b_2, and a are β_1, β_2 and α, respectively. What is the Null hypothesis for the F-test for this regression?

Ans 3:

The Null Hypothesis for F-test is,

$H_0: \beta_1 = \beta_2 = 0$ Ans

Q4. In the following table, P and Q are two random variables. Calculate the sum of the diagonal elements of the covariance matrix for these two variables. **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0.

P	Q
1	6
2	M
3	4

Ans 4: Covariance matrix

$$S = \begin{vmatrix} var(P) & Cov(P,Q) \\ Cov(P,Q) & var(Q) \end{vmatrix}$$

For $M = 0$, $var(P) = \frac{1}{(3-1)} \Sigma (P_i - \bar{P})^2 = 1$

$var(Q) = \frac{1}{(3-1)} \Sigma (Q_i - \bar{Q})^2 = 9.333$

$\therefore$ Sum of diagonal elements of the covariance matrix $= var(P) + var(Q) = 1 + 9.333 = 10.333$

* Use Bessel's correction
* Answer will depend upon roll no.

•

BT307 Quiz

Total Marks: 10 Time: 45 min

Roll No. **Name:**

Q1. We performed a multiple linear regression to fit the following model to our data. The number of data points is (**M**+10). **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0. The adjusted R-squared value for this regression is 0.5. Calculate the R-squared or the coefficient of determination for this regression.

$$y = a + b_1 x_1 + b_2 x_2$$

Ans 1: Adjusted R^2, $\bar{R}^2 = 1 - \frac{RSS}{TSS} \cdot \frac{(n-1)}{(n-k-1)} = 1 - (1-R^2)\frac{(n-1)}{(n-k-1)}$

$\therefore R^2 = 1 - (1-\bar{R}^2)\frac{(n-k-1)}{(n-1)}$

For, M = 0; number of data points, $n = M + 10 = 0 + 10 = 10$.

Number of predictors, $k = 2$.

$\therefore R^2 = 1 - (1 - 0.5)\frac{(10-2-1)}{(10-1)} = 1 - 0.5 \times \frac{7}{9}$

$= 0.61$ Ans

⊗ Answer depends on Roll No.

Q2. In the following table, *x* is the predictor or independent variable, and *y* is the outcome or dependent variable. We have performed linear regression to fit the following linear model $\hat{y}_i = bx_i + a$ to this data set. Calculate the mean of the predicted or regressed values of *y* (i.e. mean of $\hat{y}_i$ s). **M** is the last digit of your roll number. For example, if your roll number is 190106050, then **M** = 0.

X	1	2	3	4
Y	3	7	12	**M**

Ans 2:

We know,

Mean of dependent variable, $\bar{y}$ = Mean of regressed values of y_i, $\bar{\hat{y}}_i$

$\therefore$ Mean of $\hat{y}_i$, $\bar{\hat{y}}_i = \bar{y} = \frac{3 + 7 + 12 + 0}{4}$

[considering M = 0]

$= 5.5$ Ans

⊗ Answer depends on Roll No.

Q3. We want to detect the multi-collinearity issue in a regression problem. For that, we are performing auxiliary regression on each predictor. The R-squared of the auxiliary regression for the predictor x_5 is **AB**/100. Here, AB represents the last two digits of your roll number. For example, if your roll number is 210106090, then A is 9 and B is 0; so AB = 90. Calculate the Variance Inflation Factor for this predictor.

Ans 3:

Variance Inflation factor,

$$VIF = \frac{1}{1 - R_j^2}$$, where R_j^2 is the R-squared for the auxillary regression for x_j.

Consider, AB = 90

$R_5^2 = \frac{90}{100}$; $VIF = \frac{1}{1 - 0.9} = 10$ <u>Ans</u>

(*) Answer will depend on Roll No.

Q4. In the following table, y is the dependent variable, and x is the predictor. We fit a quadratic polynomial to this data using regression. For regression, we are using the linear algebra-based method. In this method, we use the relationship $\mathbf{B} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$. Here, **B** is the vector with unknown parameters/coefficients of the polynomial. **X** is a matrix, and **Y** is a vector. What is the sum of all elements in the first row of the **X** matrix? **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0. Very briefly show your calculation (only the essential steps).

x	**M**	3	7	9	13	15	19
y	12	36	79	92	103	121	137

Ans 4:

$$X = \begin{vmatrix} x_1 & x_1^2 & 1 \\ x_2 & x_2^2 & 1 \\ \vdots & & \\ x_7 & x_7^2 & 1 \end{vmatrix} = \begin{vmatrix} M & M^2 & 1 \\ 3 & 9 & 1 \\ \vdots & & \\ 19 & 19^2 & 1 \end{vmatrix}$$

Sum of first row of X

$= M + M^2 + 1 = 0 + 0^2 + 1 = 1$ <u>Ans</u>

[considering M = 0]

(*) Answer will depend on Roll No.

BT307 Quiz

Total Marks: 10 | Time: 45 min

Name: | **Roll No.:**

Q1. You fitted $y = b_1x + b_2x^2 + a$ to a data set. The population regression coefficients corresponding to b_1, b_2, and a are β_1, β_2 and α, respectively. What is the Null hypothesis for the F-test for this regression?

Ans 1:

Check the answer given for similar question.

Q2. We want to detect the multi-collinearity issue in a regression problem. For that, we are performing auxiliary regression on each predictor. The R-squared of the auxiliary regression for the predictor x_5 is **AB**/100. Here, AB represents the last two digits of your roll number. For example, if your roll number is 210106090, then A is 9 and B is 0; so AB = 90. Calculate the Variance Inflation Factor for this predictor.

Ans 2:

Check answer given for similar question.

Q3. A simple linear regression model is usually written as $\hat{y}_i = bx_i + a$. Using the matrix notation, I have written the same model as $\hat{\mathbf{y}} = \mathbf{Z}\mathbf{y}$. Here, **y** is the vector for the dependent variable data. $\hat{\mathbf{y}}$ is the vector for the fitted or regressed values of **y**. The data matrix for this regression is **X**. Express **Z** in terms of **X**. You **MUST** show a few lines of your derivation.

Ans 3:

Check answer for similar question.

Q4. We have performed linear regression, considering *x* as a predictor for the following data. The outcome or dependent variable is *y*. Calculate the TSS of this regression. **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0.

x	1	2	3	5
y	2	4	8	**M**

Ans 4:

Check answer for similar question.

BT307 Quiz

Time: 45 min

Total Marks: 10

Roll No.

Name:

Q1. In the following table, y is the dependent variable, and x is the predictor. We fit a cubic polynomial to this data using regression. For regression, we are using the linear algebra-based method. In this method, we use the relationship $\mathbf{B} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$. Here, **B** is the vector with unknown parameters/coefficients of the polynomial. **X** is a matrix, and **Y** is a vector. What is the sum of all elements in the first row of the **X** matrix? **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0.

x	**M**	13	17	37	45	65	79
y	21	36	42	52	65	79	91

Ans 1:

$$X = \begin{vmatrix} x_1 & x_1^2 & x_1^3 & 1 \\ x_2 & x_2^2 & x_2^3 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_7 & x_7^2 & x_7^3 & 1 \end{vmatrix} = \begin{vmatrix} M & M^2 & M^3 & 1 \\ 13 & 13^2 & 13^3 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 79 & 79^2 & 79^3 & 1 \end{vmatrix}$$

$\therefore$ Sum of first row elements of X

$= M + M^2 + M^3 + 1 = 0 + 0 + 0 + 1 = 1$ Ans. [considering M = 0]

(*) Answer will depend on Roll number

Q2. In the following table, P and Q are two random variables. Calculate the sum of the diagonal elements of the covariance matrix for these two variables. **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0.

P	Q
1	6
2	**M**
3	5

Ans 2:

Check answer to similar question

Q3. We performed a multiple linear regression to fit the following model to our data. The number of data points is (**M**+10). **M** is the last digit of your roll number. For example, if your roll number is 210106090, then **M** = 0. The adjusted R-squared value for this regression is 0.6. Calculate the R-squared or the coefficient of determination for this regression.

$$y = a + b_1x_1 + b_2x_2$$

Ans 3:

Check answer to similar question

Q4. In the following table, x is the predictor or independent variable, and y is the outcome or dependent variable. We have performed linear regression to fit the following linear model $\hat{y}_i = bx_i + a$ to this data set. Calculate the mean of the predicted or regressed values of y (i.e. mean of $\hat{y}_i$s). **M** is the last digit of your roll number. For example, if your roll number is 190106050, then **M** = 0.

X	1	2	3	4
Y	3	7	9	**M**

Ans 4:

Check answer to similar question.