MOOC Course- Regression Analysis and Forecasting - January 2017

Assignment 3

Note: All the questions in this assignment are based on the multiple linear regression model $y = X\beta + \epsilon$ where y is a $n \times 1$ vector of n observations on study variable, X is a $n \times k$ matrix of rank k with n observations on each of the k explanatory variables, β is a $k \times 1$ vector of associated regression coefficients and ϵ is a $n \times 1$ vector of random errors following $N(\mathbf{0}, \sigma^2\mathbf{I})$.

[1] Consider the multiple linear regression model $y = X\beta + \epsilon$ where y is a $n \times 1$ vector of observations on study variable, X is a $n \times k$ matrix of rank k with n observations on k explanatory variables, β is a $k \times 1$ vector of regression coefficients and ϵ is a $n \times 1$ vector of random errors with $E(\epsilon) = \mathbf{0}, V(\epsilon) = \sigma^2\mathbf{I}$. Suppose ϵ does not follow the normal distribution. Then which of the following statements are correct?

Statement 1 : The OLSE of β remains unbiased.

Statement 2 : The OLSE of β does not remain unbiased.

Statement 3 : The covariance matrix of OLSE of β remains as $\sigma^2(X'X)^{-1}$.

Statement 4 : The covariance matrix of OLSE of β does not remain as $\sigma^2(X'X)^{-1}$.

A. Statements 1 and 4 are correct.

B. Statements 2 and 4 are correct.

C. Statements 1 and 3 are correct.

D. Statements 2 and 3 are correct.

[2] Consider the multiple linear regression model $y = X\beta + \epsilon$ where y is a $n \times 1$ vector of observations on study variable, X is a nonstochastic $n \times k$ matrix of rank k having n observations on each of the k explanatory variables, β is a $k \times 1$ vector of regression coefficients and ϵ is a $n \times 1$ vector of random errors. Assume $E(\epsilon) = \mu \neq 0, V(\epsilon) = \sigma^2 \mathbf{I}$ and σ^2 is known. Then $E(\hat{\beta})$ where $\hat{\beta}$ is the ordinary least squares estimator of β is

A. β

B. $\beta - (X'X)^{-1}X'\mu$

C. $\beta + (X'X)^{-1}X'\mu$

D. $(X'X)^{-1}X'\mu$

[3] Consider the multiple linear regression model $y = X\beta + \epsilon$ where y is a $n \times 1$ vector of observations on study variable, X is a nonstochastic $n \times k$ matrix of rank k having n observations on each of the k explanatory variables, β is a $k \times 1$ vector of regression coefficients and ϵ is a $n \times 1$ vector of random errors. Assume $E(\epsilon) = \mathbf{0}, V(\epsilon) = \Omega \neq \mathbf{I}$ where Ω is a known positive definite matrix. The covariance matrix of ordinary least squares estimator of β is

A. $(X'X)^{-1}$

B. $(X'\Omega^{-1}X)^{-1}$

C. $(X'\Omega^{-1}X)X'X(X'\Omega^{-1}X)^{-1}$

D. $(X'X)^{-1}X'\Omega X(X'X)^{-1}$

[4] The covariance matrix of residuals based on the least squares estimation of β with n observations and k explanatory variables in a multiple linear regression model $y = X\beta + \epsilon$, $E(\epsilon) = \mathbf{0}, V(\epsilon) = \sigma^2\mathbf{I}$ under usual assumptions is

A. $\sigma^2\mathbf{I}$

B. $\sigma^2 X(X'X)^{-1}X'$

C. $\sigma^2[\mathbf{I} - X(X'X)^{-1}X']$

D. $n\sigma^2\mathbf{I}$

[5] Suppose there are 5 explanatory variables (including an intercept term) in the model $y = X\beta + \epsilon$ and $n = 26$. Further, the sum of squares due to total is obtained as 300 and sum of squares due to error is obtained as 100. The value of F-statistic in the context of analysis of variance is

A. 2

B. 2.4

C. 8

D. 10.5

[6] In the usual analysis of variance table, the estimate of σ^2 is obtained for $n = 25, k = 5$ as

A. $\frac{1}{20}\times$(Sum of squares due to regression).

B. $\frac{1}{20}\times$(Sum of squares due to residuals).

C. $\frac{1}{25}\times$(Sum of squares due to regression).

D. $\frac{1}{25}\times$(Sum of squares due to residuals).

[7] Consider the multiple linear regression model with $n = 20$ and $k = 4$ as $y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \epsilon_i$, $i = 1, 2, \ldots, 20$. The p-values for testing the null hypothesis $H_0 : \beta_1 = 0$, $H_0 : \beta_2 = 0$, $H_0 : \beta_3 = 0$ and $H_0 : \beta_4 = 0$ are obtained as 0.03, 0.02, 0.08 and 0.09 respectively at 5% level of significance. Then the conclusion about the presence of explanatory variables in the model is that

A. X_1, X_2, X_3 and X_4 remain in the model.

B. X_1 and X_2 remain in the model.

C. X_3 and X_4 remain in the model.

D. X_1, X_3 and X_4 remain in the model.

[8] Consider the multiple linear regression model $y = X\beta + \epsilon$ where $\epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ with $n = 41$ and $k = 6$ (including intercept term). If the coefficient of determination (R^2) is 0.98, then we can conclude at 5% level of significance about the null hypothesis $H_0 : \beta_1 = \beta_2 = \ldots = \beta_6 = 0$ based on the F-statistics in connection to analysis of variance that

A. H_0 is accepted.

B. H_0 is rejected.

C. Nothing can be concluded.

D. Given information is not sufficient to take any decision about acceptance or rejection of H_0.

[9] Choose the alternative that correctly matches the statements in the columns (1) and (2) in connection with $y = X\beta + \epsilon$, $\epsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I})$.

Column 1	**Column 2**
(I) Quantile - Quantile plot	(a) Checking of constancy of variance of ϵ_i's.
(II) Plots of residuals versus fitted values	(b) Checking of variance of ϵ_i's with time.
(III) Plots of residuals versus time	(c) Checking of normality of ϵ_i's.
(IV) Variance inflation factor	(d) Checking of independence of explanatory variables.

A. I - d , II - c, III - b, IV- a.

B. I - c , II - b, III - d, IV- a.

C. I - c , II - a, III - b, IV- d.

D. I - d , II - c, III - b, IV- a.

[10] Choose the alternative that correctly matches the statements in the columns (1) and (2) in connection with $y = X\beta + \epsilon$, $\epsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I})$.

Column 1	**Column 2**
(I) DW statistic	(a) Presence of first order autocorrelation.
(II) Cook's statistic	(b) Measure of performance of model for prediction.
(III) PRESS	(c) Measure of influence.

A. I - c , II - b, III - a.

B. I - a , II - b, III - c.

C. I - a , II - c, III - b.

D. I - b , II - c, III - a.

Solution to Assignment 3

Answer of Question 1 – C

Answer of Question 2 – C

Answer of Question 3 – D

Answer of Question 4 – C

Answer of Question 5 – D

Answer of Question 6 – B

Answer of Question 7 – B

Answer of Question 8 – B

Answer of Question 9 – C

Answer of Question 10 – C