Quiz 2- Linear Regression Analysis (Based on Lectures 15-31)

Time: 1 Hour

1. The random errors ε in multiple linear regression model $y = X\beta + \varepsilon$ are assumed to be identically and independently distributed following the normal distribution with zero mean and constant variance. Here y is a $n \times 1$ vector of observations on response variable, X is a $n \times K$ matrix of 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. The residuals $\hat{\varepsilon} = y - \hat{y}$ based on the ordinary least squares estimator of β have, in general,

(A) zero mean, constant variance and are independent

(B) zero mean, constant variance and are not independent

(C) zero mean, non constant variance and are not independent

(D) non zero mean, non constant variance and are not independent

Answer: (C)

<u>Solution:</u> The residual is

$$
\begin{aligned}
\hat{\varepsilon} &= y - \hat{y} \\
&= y - Xb \text{ where } b = (X'X)^{-1}X'y \\
&= (I - H)y \text{ where } H = X(X'X)^{-1}X' \\
&= (I - H)\varepsilon \\
E(\hat{\varepsilon}) &= 0 \\
V(\hat{\varepsilon}) &= \sigma^2 (I - H)
\end{aligned}
$$

- Since $E(\hat{\varepsilon}) = 0$, so $\hat{\varepsilon}_i$'s have zero mean.
- Since $I - H$ is not generally a diagonal matrix, so $\hat{\varepsilon}_i$'s do not have necessarily the same variances.
- The off-diagonal elements in $(I - H)$ are not zero, in general. So $\hat{\varepsilon}_i$'s are not independent.

2. Consider the multiple linear regression model $y = X\beta + \varepsilon, E(\varepsilon) = 0,$ $V(\varepsilon) = diag(\sigma_1^2, \sigma_2^2, ..., \sigma_n^2)$ where y is a $n \times 1$ vector of observations on response variable, X is a $n \times K$ matrix of 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. Let h_{ii} be the i^{th} diagonal element of matrix $H = X(X'X)^{-1}$. The variance of the i^{th} PRESS residual is

(A) $\sigma_i^2(1 - h_{ii})$

(B) $\sigma_i^2(1 - h_{ii})^2$

(C) $\dfrac{\sigma_i^2}{1 - h_{ii}}$

(D) $\dfrac{\sigma_i^2}{(1 - h_{ii})^2}$

Answer: (C)

Solution: Let $e_i = y_i - \hat{y}_i$ be the i^{th} ordinary residual based on the ordinary least squares estimator of β. The i^{th} PRESS residual $e_{(\text{i})}$ is

$$e_{(i)} = \frac{e_i}{1 - h_{ii}}.$$

The variance of $e_{(i)}$ is obtained as follows:

$$Var(e_{(i)}) = \frac{Var(e_i)}{(1 - h_{ii})^2}$$

and

$$V(e) = (I - H)V(\varepsilon)$$
$$\Rightarrow Var(e_i) = \sigma_i^2(1 - h_{ii}).$$

Thus

$$Var(e_{(i)}) = \frac{\sigma_i^2}{1 - h_{ii}}.$$

3. Consider the linear model $y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon, E(\varepsilon) = 0, V(\varepsilon) = I$ where the study variable y and the explanatory variables X_1 and X_2 are scaled to length unity and the correlation coefficient between X_1 and X_2 is 0.5. Let b_1 and b_2 be the ordinary least squares estimators of β_1 and β_2 respectively. The covariance between b_1 and b_2 is

(A) 2/3

(B) -2/3

(C) -0.5

(D) 1/3

Answer: (B)

Solution: If r is the correlation coefficient between X_1 and X_2, then the ordinary least squares estimators b_1 and b_2 of β_1 and β_2 are the solutions of

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = (X'X)^{-1}X'y \quad \text{where} \quad X'X = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}.$$

Then

$$(X'X)^{-1} = \frac{1}{1-r^2}\begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix}$$

$$V(b) = (X'X)^{-1} = \frac{1}{1-r^2}\begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix}$$

$$Cov(b_1, b_2) = -\frac{r}{1-r^2}.$$

When $r = 0.5, \quad \text{cov}(b_1, b_2) = -\frac{2}{3}.$

4. Under the multicolinearity problem in the data in the model $y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon$, where the study variable y and the explanatory variables X_1 and X_2 are scaled to length unity and the random error ε is normally distributed with $E(\varepsilon) = 0, V(\varepsilon) = \sigma^2 I$ where σ^2 is unknown. What do you conclude about the null hypothesis $H_{01}: \beta_1 = 0$ and $H_{02}: \beta_2 = 0$ out of the following when sample size is small.

(A) Both H_{01} and H_{02} are more often accepted.

(B) Both H_{01} and H_{02} are more often rejected.

(C) H_{01} is accepted more often and H_{02} is rejected more often.

(D) H_{01} is rejected more often and H_{02} is accepted more often.

Answer: (A)

<u>Solution:</u> If r is the correlation coefficient between X_1 and X_2, then the ordinary least squares estimators b_1 and b_2 are the solutions of

$$b = (X'X)^{-1} X'y \quad \text{where} \quad X'X = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}.$$

Then

$$E(b) = \beta$$

$$V(b) = \sigma^2 (X'X)^{-1} = \frac{\sigma^2}{1-r^2} \begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix}.$$

As r becomes higher, the variance of b_1 and b_2 become larger and so the test statistic

$$t_i = \frac{b_i}{\sqrt{\widehat{\text{var}}(b_i)}}, i = 1, 2$$

for testing $H_{0i}: \beta_i = 0$ becomes very small and so H_{0i} is more often accepted.

5. Consider the setup of multiple linear regression model $y = X\beta + \varepsilon$ where y is a $n \times 1$ vector of observations on response variable, X is a $n \times K$ matrix of 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 following $N(0, \sigma^2 I)$ with all usual assumptions. Let $H = X(X'X)^{-1}$ and h_{ii} be the i^{th} diagonal element of H. A data point is a leverage point if h_{ii} is

(A) greater than the value of average size of hat diagonal.

(B) less than the value of average size of hat diagonal.

(C) greater or less than the value of average size of hat diagonal.

(D) none of the above.

Answer: (A)

Solution: The i^{th} diagonal element of H is $h_{ii} = \underline{x}_i^1 (X^1 X)^{-1} \underline{x}_i$ where $\underline{x}_i$ is the i^{th} row of X matrix. The h_{ii} is a standardized measure of distance of i^{th} observation from the center (or centroid) of the x-space.

Average size of hat diagonal $\left(\bar{h}\right) = \dfrac{\sum_{1=i}^{n} h_{ii}}{n} = \dfrac{rank(H)}{n} = \dfrac{trH}{n} = \dfrac{K}{n}$.

If $h_{ii} > 2\bar{h}$ then the point is remote enough to be considered as a leverage point.