STAT5044 IN CLASS QUESTIONS

Zhengzhi Lin

September 19, 2019

1. Question 1:

   Suppose the true model is M2: $y_i = \beta_0 + x_i\beta_1 + \epsilon_i$ and we use model M1: $y_i = \beta_0 + \epsilon_i$
   Then our estimator for β will be
   $$\hat{\beta}_0 = \overline{y}. \tag{1}$$
   The expectation of $\hat{\beta}_0$ is:
   $$\begin{aligned} E(\hat{\beta}_0) &= E(\overline{y}) \\ &= \beta_0 + \overline{x}\beta_1 \neq \beta_0 \end{aligned} \tag{2}$$
   The variance of $\hat{\beta}_0$ is :
   $$\begin{aligned} \mathrm{Var}[\hat{\beta}_0] &= \mathrm{Var}[\overline{y}] \\ &= \frac{\sigma^2}{n} < (\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}})\sigma^2 = \mathrm{Var}[\widehat{\beta_{0,true}}] \end{aligned} \tag{3}$$
   Therefore the estimator is biased but has smaller variance than unbiased estimator for true model.

2. Question 2

   Suppose M1 is true.
   $$\mathrm{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} \tag{4}$$
   We have OLS estimator for β:
   $$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (X'X)^{-1}X'Y \tag{5}$$
   Then
   $$\begin{aligned} \mathrm{E}[\hat{\beta}] &= (X'X)^{-1}X'\mathrm{E}[Y] \\ &= \begin{bmatrix} \frac{1}{n} + \frac{\overline{x}^2}{S_{xx}} & -\frac{\overline{x}}{S_{xx}} \\ -\frac{\overline{x}}{S_{xx}} & \frac{1}{S_{xx}} \end{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_0 \\ \dots \\ \beta_0 \end{bmatrix} \\ &= \begin{bmatrix} \beta_0 \\ 0 \end{bmatrix} \end{aligned} \tag{6}$$

Thus the estimator is unbiased.

$$\begin{aligned}\mathrm{Var}[\hat{\beta}] &= \mathrm{Var}[(X'X)^{-1}X'Y] \\ &= \sigma^2(X'X)^{-1}\end{aligned} \tag{7}$$

Therefore we get

$$\mathrm{Var}[\hat{\beta}_0] = \sigma^2(\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}) > \mathrm{Var}[\widehat{\beta_{0,true}}] = \frac{\sigma^2}{n} \tag{8}$$

The variance is larger.