Answers for Homework #3

Zheng Tian

5.5 A regression of *TestScore* on *SmallClass* yields

$$\widehat{TestScore} = \underset{(1.6)}{918.0} + \underset{(2.5)}{13.9} \times SmallClass,\ R^2 = 0.01,\ SER = 74.6.$$

- **a.** The estimated gain from being in a small class is 13.9 points. This is equal to approximately 1/5 of the standard deviation in test scores, a moderate increase.
- **b.** The t-statistic is $t^{act} = 13.9/2.5 = 5.56$, which has a p-value of 0.00. Thus the null hypothesis is rejected at the 5% and 1% levels.
- **c** $13.9 \pm 2.58 \times 2.5 = 13.9 \pm 6.45$.

5.6 **a.** The question asks whether the variability in test scores in large classes is the same as the variability in small classes. It is hard to say. On the one hand, teachers in small classes might able so spend more time bringing all of the students along, reducing the poor performance of particularly unprepared students. On the other hand, most of the variability in test scores might be beyond the control of the teacher.

- **b.** The formula in Equation (5.3) is valid heteroskesdasticity or homoskedasticity; thus inferences are valid in either case.

5.8 **a.** Since $u_i \sim N(0, \sigma_u^2)$ and the sample size is small, we use the Student-t distribution for the t-statistic. The 5% critical value for a two-sided test from a Student-t distribution with the degrees of freedom of 28 is 2.05. Therefore, the 95% confidence interval for β_0 is $43.2 \pm 2.05 \times 10.2 = 43.2 \pm 20.91$.

- **b.** The t-statistic is $t^{act} = (61.5 - 55)/7.4 = 0.88$, which is less (in absolute value) than the critical value of 20.5. Thus, the null hypothesis is not rejected at the 5% level.
- **c.** The one sided 5% critical value is 1.70; t^{act} is less than this critical value, so that the null hypothesis is not rejected at the 5% level.

5.10 Let n_0 denote the number of observations with $X = 0$ and n_0 denote the number of observations with $X = 1$. So the total number of observations $n = n_1 + n_2$. Then define the proportion in all observations of $X = 1$ as $\alpha = \frac{n_1}{n}$ and the rest is $1 - \alpha = \frac{n_0}{n}$.

1. Calculate $\bar{X}$ and $\bar{Y}$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\left(\sum_{i:X_i=1}^{n_1} X_i + \sum_{i:X_i=0}^{n_0} X_i\right) = \frac{n_1}{n} = \alpha$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\left(\sum_{i:X_i=1}^{n_1} Y_i + \sum_{i:X_i=0}^{n_0} Y_i\right) = \frac{1}{n}(n_1\bar{Y}_1 + n_0\bar{Y}_0) = \alpha\bar{Y}_1 + (1-\alpha)\bar{Y}_0$$

2. Show $\hat{\beta}_1 = \bar{Y}_1 - \bar{Y}_0$

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2} \\
&= \frac{\sum_i X_i(Y_i - \bar{Y}) - \sum_i \bar{X}(Y_i - \bar{Y})}{\sum_i X_i(X_i - \bar{X}) - \sum_i \bar{X}(X_i - \bar{X})} \\
&= \frac{\sum_i X_i(Y_i - \bar{Y})}{\sum_i X_i(X_i - \bar{X})} \\
&= \frac{\sum_{i:X_i=1}(Y_i - \bar{Y})}{\sum_{i:X_i=1}(X_i - \bar{X})} \\
&= \frac{n_1\bar{Y}_1 - n_1\bar{Y}}{n_1 - n_1\bar{X}} \\
&= \frac{\bar{Y}_1 - \bar{Y}}{1 - \bar{X}}
\end{aligned}$$

Then, we have

$$\begin{aligned}
(1-\bar{X})\hat{\beta}_1 &= \bar{Y}_1 - \bar{Y} \\
(1-\alpha)\hat{\beta}_1 &= \bar{Y}_1 - \alpha\bar{Y}_1 - (1-\alpha)\bar{Y}_0 \\
\hat{\beta}_1 &= \bar{Y}_1 - \bar{Y}_0
\end{aligned}$$

3. Show that $\hat{\beta}_0 = \bar{Y}_0$ and $\hat{\beta}_0 + \hat{\beta}_1 = \bar{Y}_1$

   Since $\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1\bar{X}$, we have

$$\begin{aligned}
\alpha\bar{Y}_1 + (1-\alpha)\bar{Y}_0 &= \hat{\beta}_0 + \alpha\hat{\beta}_1 \\
\alpha\bar{Y}_1 + (1-\alpha)\bar{Y}_0 &= \hat{\beta}_0 + \alpha(\bar{Y}_1 - \bar{Y}_0) \\
\hat{\beta}_0 &= \bar{Y}_0 \\
\hat{\beta}_0 + \hat{\beta}_1 &= \bar{Y}_1
\end{aligned}$$

5.14

a. The least squares estimator of the model of $Y_i = \beta X_i + u_i$ is the solution to the minimization problem of

$$\min_b \sum_i (Y_i - bX_i)^2$$

The first order condition is

$$-2\sum_i X_i(Y_i - bX_i) = 0$$

From this equation, we get the OLS estimator of β as

$$\hat{\beta} = \frac{\sum_i X_i Y_i}{\sum_i X_i^2}$$

$\hat{\beta}_1$ is a linear function of $Y_1, \ldots, Y_n$ since we can write it as

$$\hat{\beta} = \sum_i \alpha_i Y_i, \quad \text{where } \alpha_i = \frac{X_i}{\sum_i X_i^2}$$

b. From one of the Gauss-Markov conditions, $E(u_i|X) = 0$ where $X = (X_1, \ldots, X_n)$, we can derive the unbiasedness of $\hat{\beta}$ as follows.

$$\begin{aligned} \hat{\beta} &= \frac{\sum_i X_i(\beta X_i + u_i)}{\sum_i X_i^2} = \beta + \frac{\sum_i X_i u_i}{\sum_i X_i^2} \\ E[(\hat{\beta} - \beta)|X] &= \frac{\sum_i X_i E(u_i|X)}{\sum_i X_i^2} = 0 \\ E(\hat{\beta}) &= \beta \end{aligned}$$

c. From the Gauss-Markov condition, $\text{var}(u_i|X) = \sigma_u^2$ and $E(u_i u_j|X) = 0$, we can derive the conditional variance of $\hat{\beta}$ as follows.

$$\begin{aligned} \text{var}(\hat{\beta}|X) &= \text{var}((\hat{\beta} - \beta)|X) = E((\hat{\beta} - \beta)^2|X) \\ &= E\left[\frac{(\sum_i X_i u_i)^2}{(\sum_i X_i^2)^2}|X\right] = \frac{E(\sum_i \sum_j u_i u_j X_i X_j|X)}{(\sum_i X_i^2)^2} \\ &= \frac{\sum_i E(u_i^2 X_i^2|X)}{(\sum_i X_i^2)^2} = \frac{\sigma_u^2 \sum_i X_i^2}{(\sum_i X_i^2)^2} \\ &= \frac{\sigma_u^2}{\sum_i X_i^2} \end{aligned}$$

d. Let $\tilde{\beta} = \sum_i a_i Y_i$ be any unbiased linear estimator of β. Then

$$\tilde{\beta} = \sum_i a_i(\beta X_i + u_i) = (\sum_i a_i X_i)\beta + \sum_i a_i u_i$$

For $\tilde{\beta}$ being unbiased, we must have $\sum_i a_i X_i = 1$. Since the OLS estimator $\hat{\beta}$ is also an unbiased linear estimator, it must satisfy $\sum_i \alpha_i X_i = 1$. And we can also write $a_i = \alpha_i + d_i$ where d_i can be any number, reflecting the difference between a_i and α_i. To show that $\hat{\beta}$ is BLUE, that is, it has the smallest conditional variance, we need to derive the conditional variance of $\tilde{\beta}$.

$$
\begin{aligned}
\operatorname{var}(\tilde{\beta}|X) &= E\left[(\sum_i a_i u_i)^2 |X\right] = E\left[\sum_i a_i^2 u_i^2\right] = \sigma_u^2(\sum_i a_i^2) \\
&= \sigma_u^2(\sum_i (\alpha_i + d_i)^2) = \sigma_u^2(\sum_i \alpha_i^2 + 2\sum_i \alpha_i d_i + \sum_i d_i^2) \\
&= \operatorname{var}(\hat{\beta}|X) + 2\sigma_u^2 \sum_i \alpha_i d_i + \sigma_u^2 \sum_i d_i^2
\end{aligned}
$$

where

$$
\sum_i \alpha_i d_i = \frac{\sum_i X_i d_i}{\sum_i X_i^2} = \frac{\sum_i X_i(a_i - \alpha_i)}{\sum_i X_i^2} = 0
$$

Therefore

$$
\operatorname{var}(\tilde{\beta}|X) - \operatorname{var}(\hat{\beta}|X) = \sigma_u^2 \sum_i d_i^2 \geq 0
$$

Finally, we conclude that $\operatorname{var}(\tilde{\beta}|X) \geq \operatorname{var}(\hat{\beta}|X)$ and the equality holds only when $\tilde{\beta} = \hat{\beta}$.