PS2

March 25, 2024

1. (a) *Proof.* Let $e = \beta_2 x_2 + u$, then the new regression is $y = \beta_0 + \beta_1 x_1 + e$.
   And $\mathbb{E}[e|x_1] = \beta_2 \mathbb{E}[x_2|x_1] + \mathbb{E}[u|x_1] = 0 + 0 = 0$, so it's actually a simple linear regression model.
   And the OLS estimator of β_1 is $\tilde{\beta_1}$, which means $\mathbb{E}[\tilde{\beta_1}] = \beta_1$. □

   (b) *Proof.* Using the conclusion in OLS, we have $\mathbb{V}[\tilde{\beta_1}|x_1] = \frac{\mathbb{V}[e|x_1]}{\sum_{i=1}^{n}(x_{i1}-\overline{x_1})^2}$.
   Then, we only need to proof: $\mathbb{V}[e|x_1] = \sigma_u^2 + \beta_2^2\sigma_2^2$.

$$\begin{aligned}\mathbb{V}[e|x_1] &= \mathbb{V}[u + \beta_2 x_2|x_1] \\ &= \mathbb{V}[u|x_1] + \beta_2^2 \mathbb{V}[x_2|x_1] \\ &= \sigma_u^2 + \beta_2^2\sigma_2^2\end{aligned}$$

   So, $\mathbb{V}[\tilde{\beta_1}|x_1] = \frac{\sigma_u^2+\beta_2^2\sigma_2^2}{\sum_{i=1}^{n}(x_{i1}-\overline{x_1})^2}$. □

   (c) $R_1^2 \to 0$, since the more samples we have, the more accurate the estimator is, and the variance of the estimator is smaller.

   (d) $\widehat{\beta_1}$, since it has a smaller variance when n is large.

2. (a) $\mathbb{V}[\widehat{\beta}_1 - \widehat{\beta}_2] = \mathbb{V}[\widehat{\beta}_1] + \mathbb{V}[\widehat{\beta}_2] - 2Cov[\widehat{\beta}_1, \widehat{\beta}_2]$.

   (b) t-stat:

$$\begin{aligned}t &= \frac{\widehat{\beta}_1 - \widehat{\beta}_2 - 1}{\mathrm{se}(\widehat{\beta}_1 - \widehat{\beta}_2)} \\ &= \frac{\widehat{\beta}_1 - \widehat{\beta}_2 - 1}{\sqrt{\mathrm{se}(\widehat{\beta}_1)^2 + \mathrm{se}(\widehat{\beta}_2)^2 - 2Cov(\widehat{\beta}_1, \widehat{\beta}_2)}}\end{aligned}$$

   (c)

$$\begin{aligned}y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u \\ &= \beta_0 + (\beta_2 + \theta_1)x_1 + \beta_2 x_2 + \beta_3 x_3 + u \\ &= \beta_0 + \theta_1 x_1 + \beta_2(x_1 + x_2) + \beta_3 x_3 + u \\ &= \beta_0 + \theta_1 x_1 + \beta_2 z + \beta_3 x_3 + u \quad (z = x_1 + x_2)\end{aligned}$$

3. (a) See the log file.

   (b) Overestimate. Because the higher IQ and longer education time, the higher wage.(By intuition)

   (c) See the log file.

   (d) The longer the education time, the higher the wage.

   (e) Conclusion: 0.0118¿1% , so the null hypothesis can not be rejected.

   (f) $\beta_2 = \beta_1$.

   (g) Conclusion: the confidence interval of θ doesn't contain 0, so the null hypothesis is rejected.