Test Exercise 2.
Tamara Grigorian

(a) Prove that $E(b_R) = \beta_1 + P\beta_2$.

$$b_R = \left(X_1^{'}X_1\right)^{-1} X_1^{'}y = \left(X_1^{'}X_1\right)^{-1} X_1^{'}\left(X_1\beta_1 + X_2\beta_2 + \varepsilon\right) = \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_1\beta_1 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2\beta_2 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon$$

$$\left(X_1^{'}X_1\right)^{-1} X_1^{'}X_1 = I,\ \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2 = P, \qquad b_R = \beta_1 + P\beta_2 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon.$$

$$E(b_R) = E\left(\beta_1 + P\beta_2 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon\right) = E(\beta_1) + PE(\beta_2) + \left(X_1^{'}X_1\right)^{-1} X_1^{'}E(\varepsilon) = \beta_1 + P\beta_2.$$

Use that β_1, β_2 are constants and $E(\varepsilon) = 0$, according to assumptions A1-A6.

(b) Prove that $\operatorname{var}(b_R) = \sigma^2\left(X_1^{'}X_1\right)^{-1}$.

$$\operatorname{var}(b_R) = E\left(\left(b_R - E(b_R)\right)\left(b_R - E(b_R)\right)^{'}\right).$$

Compute $b_R - E(b_R)$.

$$b_R - E(b_R) = \left(X_1^{'}X_1\right)^{-1} X_1^{'}y - (\beta_1 + P\beta_2) = \left(X_1^{'}X_1\right)^{-1} X_1^{'}y - \beta_1 - \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2\beta_2.$$

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

$$b_R - E(b_R) = \left(X_1^{'}X_1\right)^{-1} X_1^{'}\left(X_1\beta_1 + X_2\beta_2 + \varepsilon\right) - \beta_1 - \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2\beta_2 =$$
$$= \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_1\beta_1 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2\beta_2 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon - \beta_1 - \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2\beta_2 = \beta_1 - \beta_1 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon.$$

So, $b_R - E(b_R) = \left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon$.

$$\operatorname{var}(b_R) = E\left(\left(\left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon\right)\left(\left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon\right)^{'}\right) = E\left(\left(X_1^{'}X_1\right)^{-1} X_1^{'}\varepsilon\varepsilon' X_1\left(X_1^{'}X_1\right)^{-1}\right) =$$
$$= \left(X_1^{'}X_1\right)^{-1} X_1^{'}E(\varepsilon\varepsilon')X_1\left(X_1^{'}X_1\right)^{-1} = \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_1\left(X_1^{'}X_1\right)^{-1} E(\varepsilon\varepsilon') = \left(X_1^{'}X_1\right)^{-1}\sigma^2.$$

Used that $X_1^{'}X_1\left(X_1^{'}X_1\right)^{-1} = I, E(\varepsilon\varepsilon') = \sigma^2$.

(c) Prove that $b_R = b_1 + Pb_2$.

$$b_R = \left(X_1^{'}X_1\right)^{-1} X_1^{'}y = \left(X_1^{'}X_1\right)^{-1} X_1^{'}\left(X_1b_1 + X_2b_2 + e\right) = \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_1b_1 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2b_2 + \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_1e.$$

$$\left(X_1^{'}X_1\right)^{-1} X_1^{'}X_1 = I,\ X_1^{'}e = 0$$

since, according to OLS, $e \perp X_1$, $e \perp X_2$. $\dfrac{\partial S}{\partial b_{1i}} = 0, \dfrac{\partial S}{\partial b_{2i}} = 0 \Rightarrow \sum e_i X_{1i} = 0, \sum e_i X_{2i} = 0$.

(d) $P = \left(X_1^{'}X_1\right)^{-1} X_1^{'}X_2 = X_1^{-1}\left(X_1^{'}\right)^{-1} X_1^{'}X_2 = X_1^{-1}X_2$, we obtain $P = X_1^{-1}X_2$, $X_1P = X_2$. This equality shows a connection of the variables ‘Age’, ‘Educ’, ‘Parttime’ with a constant term and the variable ‘Female’.

(e) Denote rows $x_1 = 1, x_2 = Female, x_3 = Age, x_4 = Educ, x_5 = Parttime$.

(f)

Then $X_1 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} = \begin{pmatrix} 1 & x_2 \end{pmatrix}, \mathrm{X}_2 = \begin{pmatrix} x_3 & x_4 & x_5 \end{pmatrix}$. From Lecture 2.1 we have: $\begin{matrix} x_3 = 40,05 \cdot 1 - 0,11 x_2 \\ x_4 = 2,66 \cdot 1 - 0,49 x_2 \\ x_5 = 0,2 x_1 + +0,25 x_2 \end{matrix}$.

We can right it using matrices $\begin{pmatrix} x_3 & x_4 & x_5 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 40,05 & 2,26 & 0,2 \\ -0,11 & -0,49 & 0,25 \end{pmatrix}$,

$$X_2 = X_1 \begin{pmatrix} 40,05 & 2,26 & 0,2 \\ -0,11 & -0,49 & 0,25 \end{pmatrix}.$$

$$P = \left(X_1^{'} X_1 \right)^{-1} X_1^{'} X_2 = \left(X_1^{'} X_1 \right)^{-1} X_1^{'} X_1 \begin{pmatrix} 40,05 & 2,26 & 0,2 \\ -0,11 & -0,49 & 0,25 \end{pmatrix} = \begin{pmatrix} 40,05 & 2,26 & 0,2 \\ -0,11 & -0,49 & 0,25 \end{pmatrix}.$$

(g) Using results of the Lection 2.1 we can determine the values of b_R as a coefficients of the log(Wage) simple regression: $b_R = \begin{pmatrix} 4,73 \\ -0,25 \end{pmatrix}$.

Using that $\log\left(Wage\right)_i = 3,05 - 0,04 Female_i + 0,03 Age_i + 0,23 Educ_i - 0,37 Parttime_i + e_i$, we can determine the values of b_1 and b_2: $b_1 = \begin{pmatrix} 3,05 \\ -0,04 \end{pmatrix}, b_2 = \begin{pmatrix} 0,03 \\ 0,23 \\ -0,37 \end{pmatrix}$.

$$b_1 + P b_2 = \begin{pmatrix} 3,05 \\ -0,04 \end{pmatrix} + \begin{pmatrix} 40,05 & 2,26 & 0,2 \\ -0,11 & -0,49 & 0,25 \end{pmatrix} \begin{pmatrix} 0,03 \\ 0,23 \\ -0,37 \end{pmatrix} = \begin{pmatrix} 4,6273 \\ -0,2782 \end{pmatrix} \approx \begin{pmatrix} 4,73 \\ -0,25 \end{pmatrix} = b_R.$$