Test Exercise 2: Answers to the Questions

This test exercise is of a theoretical nature. In our discussion of the F-test, the total set of explanatory factors was split in two parts. The factors in X_1 are always included in the model, whereas those in X_2 are possibly removed. In questions (a), (b), and (c) you derive relations between the two **OLS** estimates of the effects of X_1 on y, one in the large model and the other in the small model. In parts (d), (e), and (f), you check the relation of question (c) numerically for the wage data of our lectures.

We use the notation of Lecture 2.4.2 and assume that the standard regression assumptions $A1 - A6$ are satisfied for the unrestricted model. The restricted model is obtained by deleting the set of g explanatory factors collected in the last g columns X_2 of X. We wrote the model with $X = (X_1 X_2)$ and corresponding partitioning of the **OLS** estimator b in b_1 and b_2 as $y = X_1\beta_1 + X_2\beta_2 + \epsilon = X_1 b_1 + X_2 b_2 + e$. We denote by b_R the **OLS** estimator of β_1 obtained by regressing y on X_1, so that $b_R = (X_1'X_1)^{-1}X_1'y$. Further, let $P = (X_1'X_1)^{-1}X_1'X_2$.

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

$$
\begin{aligned}
&E(b_R)\\
&= E((X_1'X_1)^{-1}X_1'y)\\
&= E((X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + \epsilon))\\
&= E((X_1'X_1)^{-1}(X_1'X_1)\beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\epsilon)\\
&= E(I.\beta_1 + P.\beta_2 + (X_1'X_1)^{-1}X_1'.\epsilon)\\
&= E(\beta_1) + P.E(\beta_2) + (X_1'X_1)^{-1}X_1'.E(\epsilon) && \text{by linearity of expectation and A2}\\
&= \beta_1 + P\beta_2 + (X_1'X_1)^{-1}X_1'.0 && \text{by A3}\\
&= \beta_1 + P\beta_2
\end{aligned}
$$

Six DGP assumptions

- A1 Linear model: $y = X\beta + \varepsilon$.
- A2 Fixed regressors: X non-random.
- A3 Random error terms with mean zero: $E(\varepsilon) = 0$.
- A4 Homoskedastic error terms: $E(\varepsilon_i^2) = \sigma^2$ for all $i = 1, \ldots, n$.
- A5 Uncorrelated error terms: $E(\varepsilon_i\varepsilon_j) = 0$ for all $i \neq j$.
- A6 Parameters β and σ^2 are fixed and unknown.

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

$$
\begin{aligned}
&var(b_R)\\
&= E((b_R - E(b_R))(b_R - E(b_R))') && \text{by A1 A2 A3 A6}\\
&= E((b_R - \beta_1 - P\beta_2)(b_R - \beta_1 - P\beta_2)')\\
&= E(((X_1'X_1)^{-1}X_1'y - \beta_1 - (X_1'X_1)^{-1}X_1'X_2\beta_2)((X_1'X_1)^{-1}X_1'y - \beta_1 - (X_1'X_1)^{-1}X_1'X_2\beta_2)')\\
&= E((((X_1'X_1)^{-1}X_1').(y - X_2\beta_2) - \beta_1)(((X_1'X_1)^{-1}X_1').(y - X_2\beta_2) - \beta_1)')\\
&= E((((X_1'X_1)^{-1}X_1').(X_1\beta_1 + \epsilon)) - \beta_1)(((X_1'X_1)^{-1}X_1').(X_1\beta_1 + \epsilon)) - \beta_1)')\\
&= E(((X_1'X_1)^{-1}.(X_1'X_1).\beta_1 + (X_1'X_1)^{-1}X_1'\epsilon - \beta_1).((X_1'X_1)^{-1}.(X_1'X_1).\beta_1 + (X_1'X_1)^{-1}X_1'\epsilon - \beta_1)')\\
&= E((I.\beta_1 + (X_1'X_1)^{-1}X_1'\epsilon - \beta_1).(I.\beta_1 + (X_1'X_1)^{-1}X_1'\epsilon - \beta_1)')\\
&= E((\beta_1 - \beta_1 + (X_1'X_1)^{-1}X_1'\epsilon).(\beta_1 - \beta_1 + (X_1'X_1)^{-1}X_1'\epsilon)')\\
&= E(((X_1'X_1)^{-1}X_1'\epsilon).((X_1'X_1)^{-1}X_1'\epsilon)')\\
&= E((X_1'X_1)^{-1}X_1'\epsilon\epsilon'X_1(X_1'X_1)^{-1})\\
&= (X_1'X_1)^{-1}X_1'E(\epsilon\epsilon')X_1(X_1'X_1)^{-1} && \text{by A2}\\
&= (X_1'X_1)^{-1}X_1'\sigma^2X_1(X_1'X_1)^{-1} && \text{by A4}\\
&= \sigma^2(X_1'X_1)^{-1}X_1'X_1(X_1'X_1)^{-1}\\
&= \sigma^2(X_1'X_1)^{-1}
\end{aligned}
$$

- © Prove that $b_R = b_1 + Pb_2$.
  - Proof:

$$
\begin{aligned}
&S(b_1, b_2) = e'e = (y - X_1b_1 - X_2b_2)'(y - X_1b_1 - X_2b_2)\\
&= y'y - y'(X_1b_1 + X_2b_2) - b_1'X_1'y + b_1'X_1'X_1b_1 + b_1'X_1'X_2b_2 - b_2'X_2'y + b_2'X_2'X_1b_1 + b_2'X_2'X_2b_2\\
&= y'y - (y'X_1b_1 + b_1'X_1'y) - (y'X_2b_2 + b_2'X_2'y) + b_1'X_1'X_1b_1 + b_1'X_1'X_2b_2 + b_2'X_2'X_1b_1 + b_2'X_2'X_2b_2\\
&= y'y - 2b_1'X_1'y - 2b_2'X_2'y + b_1'X_1'X_1b_1 + b_1'X_1'X_2b_2 + b_2'X_2'X_1b_1 + b_2'X_2'X_2b_2\\
&\Rightarrow \frac{\delta S}{\delta b1} = -2X_1'y + 2(X_1'X_1b_1) + X_1'X_2b_2 + b_2'X_2'X_1 = -2X_1'y + 2(X_1'X_1b_1) + 2(X_1'X_2b_2) = 0 && \text{by OLS criterion, minimize}\\
&\Rightarrow X_1'y = X_1'X_1b_1 + X_1'X_2b_2 && \text{by First order conditions}\\
&\Rightarrow (X_1'X_1)^{-1}X_1'y = b_1 + (X_1'X_1)^{-1}X_1'X_2b_2\\
&\Rightarrow b_R = b_1 + Pb_2
\end{aligned}
$$

- Now consider the wage data of Lectures 2.1 and 2.5. Let y be log-wage (500×1) vector, and let X_1 be the (500×2) matrix for the constant term and the variable 'Female'. Further let X_2 be the (500×3) matrix with observations of the variables 'Age', 'Educ' and 'Parttime'. The values of b_R were given in Lecture 2.1, and those of b in Lecture 2.5.

- (d) Argue that the columns of the (2×3) matrix P are obtained by regressing each of the variables 'Age', 'Educ', and 'Parttime' on a constant term and the variable 'Female'.

  - Let $X_2(i)$ (which is a (500×1) vector) denote the i^{th} column of the matrix X_2. Then, from lecture 2.1 , we have,

$$X_2(1) = Age = 40.05 - 0.11Female + e = X_1.(40.05 \ \ -0.11)^T + e$$
$$X_2(2) = Educ = 2.26 - 0.49Female + e = X_1.(2.26 \ \ -0.49)^T + e$$
$$X_2(3) = Parttime = 0.20 + 0.25Female + e = X_1.(0.20 \ 0.25)^T + e$$

$$\Rightarrow X_2 = (X_2(1) \ X_2(2) \ X_2(3)) = X_1.\begin{pmatrix} 40.05 & 2.26 & 0.20 \\ -0.11 & -0.49 & 0.25 \end{pmatrix} + e = X_1.P + e, \text{ since } P = (X_1'X_1)^{-1}.(X_1'X_2).$$

- (e) Determine the values of P from the results in Lecture 2.1.

  - From (d), we have, $P = \begin{pmatrix} 40.05 & 2.26 & 0.20 \\ -0.11 & -0.49 & 0.25 \end{pmatrix}$.

- (f) Check the numerical validity of the result in part (c). Note: This equation will not hold exactly because the coefficients have been rounded to two or three decimals; preciser results would have been obtained for higher precision coefficients.

  - From lecture 2.5 again we have,

$$log(Wage) = 3.05 - 0.04Female + 0.03Age + 0.23Educ - 0.37Parttime + e \Rightarrow b_1 = \begin{pmatrix} 3.05 \\ -0.04 \end{pmatrix}, \ b_2 = \begin{pmatrix} 0.03 \\ 0.23 \\ -0.37 \end{pmatrix}$$

$$\Rightarrow P.b_2 = \begin{pmatrix} 40.05 & 2.26 & 0.20 \\ -0.11 & -0.49 & 0.25 \end{pmatrix}.\begin{pmatrix} 0.03 \\ 0.23 \\ -0.37 \end{pmatrix} = \begin{pmatrix} 1.6473 \\ -0.2085 \end{pmatrix} \Rightarrow b_1 + Pb_2 = \begin{pmatrix} 4.6973 \\ -0.2485 \end{pmatrix}.$$

But, from lecture 2.1 we have, $log(Wage) = 4.73 - 0.25Female + e \Rightarrow b_R = \begin{pmatrix} 4.73 \\ -0.25 \end{pmatrix}$. Hence, we can see $b_R \approx b_1 + Pb_2$ (as per note, some loss of precision happens because of rounding).

```
## [1] "P.b2"
```

```
##         [,1]
## [1,]  1.6473
## [2,] -0.2085
```

```
## [1] "b1+P.b2"
```

```
##         [,1]
## [1,]  4.6973
## [2,] -0.2485
```