MEI MT Problem Set 5[1]

Timothee Carayol

January 15, 2010

[1] Available on http://personal.lse.ac.uk/carayolt/ec402.htm

Question 1

Question

$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$; $A1, A2, A3F, A4GM, A5normal$ hold, with $i = 1, .., 11$. Further assume: $x_1' x_1 = 2$, $x_2' x_2 = 2$, $x_1' x_2 = 1$, $x_1' y = 1$, $x_2' y = 1$, $y' y = \frac{4}{3}$. Finally assume that for two hypothetical observations we have

$$\begin{pmatrix} x_{1,12} & x_{2,12} \\ x_{1,13} & x_{2,13} \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 3 & -7 \end{pmatrix}.$$

Answer

- Some preliminary algebra:
  $$X'X = \begin{pmatrix} x_1' x_1 & x_1' x_2 \\ x_2' x_1 & x_2' x_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$
- Therefore $(X'X)^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$.
- $\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = (X'X)^{-1} X' y = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$.

Question 1 (cont)

Answer

- $$s^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{11-2} = \frac{y'M_X y}{9} = \frac{y'y - y'X(X'X)^{-1}X'y}{9}$$
$$= \frac{\frac{4}{3} - \begin{pmatrix} 1 & 1 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{9}$$
$$= \frac{\frac{4-2}{3}}{9} = \frac{2}{27}$$
- We'll start with (b), and then do (a), because (a) builds on (b).

Question 1 (b)

Answer

- (b) Want to find 80% *prediction* intervals for the expected value of the dependent variable y for observation 12 and 13.
- Note that we talk about *prediction* rather than *confidence* intervals here, because we want to find an interval around a predicted value than around a population parameter. But the method is the same: find a pivotal quantity whose distribution is known, and which allows us to build an interval which has a known probability to include our quantity of interest.
-

$$\begin{aligned} E(y_{12}) = x_{12}'\beta &\Rightarrow E(y_{12}) - x_{12}'\hat{\beta} = x_{12}'\beta - x_{12}'\hat{\beta} \\ &\Rightarrow E(y_{12}) - x_{12}'\hat{\beta} = x_{12}'(\beta - \hat{\beta}) \end{aligned}$$

Question 1 (b) (cont)

Answer

- $\beta - \hat{\beta} \sim \mathcal{N}\left(0, \sigma_\epsilon^2 (X'X)^{-1}\right) \Rightarrow x'_{12}(\beta - \hat{\beta}) \sim \mathcal{N}\left(0, \sigma_\epsilon^2 x'_{12}(X'X)^{-1}x_{12}\right)$
- $$\begin{aligned} E(y_{12}) - x'_{12}\hat{\beta} &\sim \mathcal{N}\left(0, \sigma_\epsilon^2 x'_{12}(X'X)^{-1}x_{12}\right) \\ &\sim \mathcal{N}\left(0, \sigma_\epsilon^2 \begin{pmatrix} 5 & -2 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \end{pmatrix}\right) \\ &\sim \mathcal{N}\left(0, \sigma_\epsilon^2(26)\right) \\ &\sim \mathcal{N}\left(0, 26\sigma_\epsilon^2\right) \end{aligned}$$

Question 1 (b) (cont)

Answer

- But we do not know σ_ε, so we use instead
  $T = \frac{E(y_{12}) - x_{12}'\hat{\beta}}{\sqrt{26s^2}} = \frac{E(y_{12}) - x_{12}'\hat{\beta}}{\sqrt{1.93}} \sim t(N-K) = t(9)$.
- Our prediction interval for $E(y_{12})$ is, then,
  $[x_{12}'\hat{\beta} - t_{9,10\%}\sqrt{1.93}, x_{12}'\hat{\beta} + t_{9,10\%}\sqrt{1.93}] =$
  $[1 - 1.383\sqrt{1.93}, 1 + 1.383\sqrt{1.93}]$.
- Likewise, we could find an 80% prediction interval for $E(y_{13})$:
  $[-\frac{4}{3} - 1.383\sqrt{3.901}, -\frac{4}{3} + 1.383\sqrt{3.901}]$.

Question 1 (a)

Answer

- (a) Want to find 80% *prediction* intervals for the dependent variable y for observation 12 and 13.

$$
\begin{aligned}
y_{12} = x_{12}'\beta + \varepsilon_{12} &\Rightarrow y_{12} - x_{12}'\hat{\beta} = x_{12}'\beta - x_{12}'\hat{\beta} + \varepsilon_{12} \\
&\Rightarrow y_{12} - x_{12}'\hat{\beta} = x_{12}'(\beta - \hat{\beta}) + \varepsilon_{12}
\end{aligned}
$$

- Note that the only difference from (b) is this extra randomness from the error term, which will increase the variance, and hence lead to a wider prediction interval compared to (b).

$$
\begin{aligned}
y_{12} - x_{12}'\hat{\beta} &\sim \mathcal{N}\left(0, \sigma_\epsilon^2 x_{12}'(X'X)^{-1}x_{12} + \sigma_\varepsilon^2\right) \\
&\sim \mathcal{N}\left(0, \sigma_\epsilon^2(26+1)\right) \\
&\sim \mathcal{N}\left(0, 27\sigma_\epsilon^2\right)
\end{aligned}
$$

Question 1 (a) (cont)

Answer

- Otherwise, same thing as (b): substitute s^2 for σ_ε^2, etc.
- Our prediction interval for y_{12} is, then, $[x_{12}'\hat{\beta} - t_{9,10\%}\sqrt{2}, x_{12}'\hat{\beta} + t_{9,10\%}\sqrt{2}] =$ $[1 - 1.383\sqrt{2}, 1 + 1.383\sqrt{2}]$.
- Likewise, we could find an 80% prediction interval for y_{13}: $[-\frac{4}{3} - 1.383\sqrt{3.975}, -\frac{4}{3} + 1.383\sqrt{3.975}]$.

Question 1 (c)

Answer

- Our answers differ because our prediction for y has to take into account the added randomness of the error term, which our prediction for $E(y)$ ignores. Hence our prediction interval for y is wider.

Question 2

Question

Consider the model $y = X\beta + \varepsilon$, $X : T \times K$, with (say) $A1, A2, A3Rmi$.

Answer

- First note that you can see both of the transformations suggested as a multiplication of the matrix of regressors X on the right by some $K \times K$ non-singular matrix, say A.
- To see this, note that for any matrix A, each column of XA is a linear combination of the columns of X. Conversely, any transformation of X whereby each regressor is transformed into a linear combination of the columns of X can be written as $Z = XA$ for some square matrix A; if the transformation is such that no information is lost[2], then A is non-singular.

[2] Loosely speaking. Another way of saying this is that each column of X is also a linear combination of the columns of Z.

Question 2

Answer

- Through such a transformation, the regression of y on Z is essentially the same as y on X. The coefficients may change, but not the residuals, the predictions, nor the goodness-of-fit. To see this, note that $P_Z = (XA)([XA]'XA)^{-1}(XA)' = (XA)(A'X'XA)^{-1}A'X' = XAA^{-1}(X'X)^{-1}A'^{-1}A'X' = X(X'X)^{-1}X' = P_X$, hence $\hat{y}_Z = P_Z y = P_X y = \hat{y}_X$, and also $\hat{\varepsilon}_Z = M_Z y = M_X y = \hat{\varepsilon}_X$.
- But, *in general*, the estimated coefficients will change: $\hat{\beta}_Z = ([XA]'XA)^{-1}(XA)'y = A^{-1}(X'X)^{-1}A'^{-1}A'X'y = A^{-1}\hat{\beta}_X$.

Question 2 (a)

Question

(a) Consider the transformation whereby one regressor is multiplied by $\lambda \neq 0$. Use this to explain whether and how linear regressions are sensitive to units of measurement.

Answer

- The transformation leaves all columns of X unchanged except one which is multiplied by λ, meaning that A will be the $K \times K$ identity matrix with one 1 substituted with λ.
- The inverse of A is then the identity with the same 1 replaced with $\frac{1}{\lambda}$. This implies, from the previous discussion, that the coefficient on that regressor is multiplied by $\frac{1}{\lambda}$, while all the others are unchanged.
- Note: Easy to see from $\beta_k.x_k = \frac{\beta_k}{\lambda}.(\lambda x_k)$.

Question 2 (b)

Question

(b) Consider the transformation whereby we add a constant λ to one regressor. Use this to explain whether and how linear regressions are sensitive to units of measurement when the corresponding regressor is measured in logs.

Answer

- Again, the transformation leaves all columns unchanged except one to which λ is added. A will be the $K \times K$ identity matrix, with a λ at the first row of the column corresponding to the modified regressor.
- The inverse of A is then the identity with a $-\lambda$ instead of the λ in A (see board). Hence the coefficient on the regressors are unchanged, except for the constant, to which we substract λ.
- Note: Easy to see from $\beta_0 + \beta_k . x_k = (\beta_0 - \lambda\beta_k) + \beta_k.(\lambda + x_k)$.