Econometrics 710
Answers to Midterm Exam
March 6, 2001

1. We know that both $\hat{\beta}$ and $\tilde{\beta}$ are unbiased for β, $\hat{\beta}-\beta=(X'X)^{-1}(X'e)$ and $\tilde{\beta}-\beta=(X'D^{-1}X)^{-1}(X'D^{-1}e)$

   (a) Note that

$$\begin{aligned}\hat{\beta}-\tilde{\beta} &= \left(\hat{\beta}-\beta\right)-\left(\tilde{\beta}-\beta\right)\\ &= (X'X)^{-1}(X'e)-\left(X'D^{-1}X\right)^{-1}\left(X'D^{-1}e\right)\\ &= \left[(X'X)^{-1}X'-\left(X'D^{-1}X\right)^{-1}X'D^{-1}\right]e\end{aligned}$$

   and $E\left(\hat{\beta}-\tilde{\beta}\mid X\right)=0$. Thus

$$\begin{aligned}&Cov\left(\hat{\beta}-\tilde{\beta},\tilde{\beta}\mid X\right)\\ =\ & E\left[\left(\hat{\beta}-\tilde{\beta}\right)\left(\tilde{\beta}-\beta\right)'\mid X\right]\\ =\ & E\left\{\left[(X'X)^{-1}X'-\left(X'D^{-1}X\right)^{-1}X'D^{-1}\right]e\left(e'D^{-1}X\right)\left(X'D^{-1}X\right)^{-1}\mid X\right\}\\ =\ & \left[(X'X)^{-1}X'-\left(X'D^{-1}X\right)^{-1}X'D^{-1}\right]E\left(ee'\mid X\right)D^{-1}X\left(X'D^{-1}X\right)^{-1}\\ =\ & \left[(X'X)^{-1}X'-\left(X'D^{-1}X\right)^{-1}X'D^{-1}\right]DD^{-1}X\left(X'D^{-1}X\right)^{-1}\\ =\ & \left[(X'X)^{-1}X'-\left(X'D^{-1}X\right)^{-1}X'D^{-1}\right]X\left(X'D^{-1}X\right)^{-1}\\ =\ & \left[(X'X)^{-1}X'X-\left(X'D^{-1}X\right)^{-1}X'D^{-1}X\right]\left(X'D^{-1}X\right)^{-1}\\ =\ & \left[I_k-I_k\right]\left(X'D^{-1}X\right)^{-1}\\ =\ & 0\end{aligned}$$

   .

   (b) Since

$$\begin{aligned}0 &= Cov\left(\hat{\beta}-\tilde{\beta},\tilde{\beta}\mid X\right)\\ &= Cov\left(\hat{\beta},\tilde{\beta}\mid X\right)-Cov\left(\tilde{\beta},\tilde{\beta}\mid X\right)\\ &= Cov\left(\hat{\beta},\tilde{\beta}\mid X\right)-Var\left(\tilde{\beta}\mid X\right),\end{aligned}$$

   it follows from our result in part (a) that $Cov\left(\hat{\beta},\tilde{\beta}\mid X\right)=Var\left(\tilde{\beta}\mid X\right)$.

(c) Using part (b),

$$\begin{aligned} Var\left(\hat{\beta}-\tilde{\beta} \mid X\right) &= Var\left(\hat{\beta} \mid X\right)+Var\left(\tilde{\beta} \mid X\right)+Cov\left(\hat{\beta},\tilde{\beta} \mid X\right)+Cov\left(\hat{\beta},\tilde{\beta} \mid X\right)' \\ &= Var\left(\hat{\beta} \mid X\right)+Var\left(\tilde{\beta} \mid X\right)-Var\left(\tilde{\beta} \mid X\right)-Var\left(\tilde{\beta} \mid X\right) \\ &= Var\left(\hat{\beta} \mid X\right)-Var\left(\tilde{\beta} \mid X\right) \end{aligned}$$

(d) Using part (b) and the known variances of $\hat{\beta}$ and $\tilde{\beta}$,

$$\begin{aligned} V_n &= Var\left(\hat{\beta}-\tilde{\beta} \mid X\right) \\ &= Var\left(\hat{\beta} \mid X\right)-Var\left(\tilde{\beta} \mid X\right) \\ &= \left(X'X\right)^{-1}\left(X'DX\right)\left(X'X\right)^{-1}-\left(X'D^{-1}X\right)^{-1}. \end{aligned}$$

Note 1: Perhaps an easier method of proof would be to first prove part (b), and use this to prove part (a). The first argument works as follows:

$$\begin{aligned} Cov\left(\hat{\beta},\tilde{\beta} \mid X\right) &= E\left[\left(\hat{\beta}-\beta\right)\left(\tilde{\beta}-\beta\right)' \mid X\right] \\ &= E\left[\left(X'X\right)^{-1}\left(X'e\right)\left(e'D^{-1}X\right)\left(X'D^{-1}X\right)^{-1} \mid X\right] \\ &= \left(X'X\right)^{-1}X'E\left[ee' \mid X\right]D^{-1}X\left(X'D^{-1}X\right)^{-1} \\ &= \left(X'X\right)^{-1}X'DD^{-1}X\left(X'D^{-1}X\right)^{-1} \\ &= \left(X'X\right)^{-1}X'X\left(X'D^{-1}X\right)^{-1} \\ &= \left(X'D^{-1}X\right)^{-1} \\ &= Var(\tilde{\beta} \mid X). \end{aligned}$$

Note 2: This result holds quite generally when one estimator is efficient. In general, for any unbiased estimators $\hat{\beta}$ and $\tilde{\beta}$, where $\tilde{\beta}$ is efficient, then $Var\left(\hat{\beta}-\tilde{\beta} \mid X\right) = Var\left(\hat{\beta} \mid X\right)-Var\left(\tilde{\beta} \mid X\right)$.

2. To solve this question, you need to recognize the following. First, you can write the estimator as

$$\tilde{\beta}-\beta=\left(\sum_{i=1}^{n} x_i x_i' 1\left(|x_i| \leq c\right)\right)^{-1}\left(\sum_{i=1}^{n} x_i e_i 1\left(|x_i| \leq c\right)\right).$$

Second, $x_i x_i' 1\left(|x_i| \leq c\right)$ is an iid random variable with mean

$$E\left(x_i x_i' 1\left(|x_i| \leq c\right)\right) \equiv Q_c.$$

Third, $x_i e_i 1(|x_i| \le c)$ is an iid random variable with mean zero:

$$E(x_i e_i 1(|x_i| \le c)) = E(x_i 1(|x_i| \le c) E(e_i \mid x_i)) = 0$$

(using the law of iterated expectations), and has variance

$$E\left[(x_i e_i 1(|x_i| \le c))(x_i e_i 1(|x_i| \le c))'\right] = E\left(x_i x_i' e_i^2 1(|x_i| \le c)\right) \equiv \Omega_c.$$

Note that $Q_c \ne Q = E(x_i x_i')$ and $\Omega_c \ne \Omega = E(x_i x_i' e_i^2)$.

(a) By the WLLN,

$$\frac{1}{n}\sum_{i=1}^{n} x_i x_i' 1(|x_i| \le c) \to_p E(x_i x_i' 1(|x_i| \le c)) = Q_c$$

and

$$\frac{1}{n}\sum_{i=1}^{n} x_i e 1(|x_i| \le c) \to_p E(x_i e 1(|x_i| \le c)) = 0.$$

Hence, if $Q_c > 0$, then by the CMT,

$$\tilde{\beta} - \beta = \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i' 1(|x_i| \le c)\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i e_i 1(|x_i| \le c)\right) \to_p Q_c^{-1} \cdot 0 = 0.$$

(b) By the CLT,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i e 1(|x_i| \le c) \to_d N(0, \Omega_c).$$

Hence

$$\begin{aligned} \sqrt{n}\left(\tilde{\beta} - \beta\right) &= \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i' 1(|x_i| \le c)\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i e_i 1(|x_i| \le c)\right) \\ &\to_d Q_c^{-1} N(0, \Omega_c) = N\left(0, Q_c^{-1}\Omega_c Q_c^{-1}\right). \end{aligned}$$

(c) Bonus Question: The condition $E(x_i e_i) = 0$ is *not* sufficient for

$$E(x_i e 1(|x_i| \le c)) = 0.$$

Thus $\tilde{\beta}$ will not necessarily be consistent for β. In fact,

$$\tilde{\beta} \to_p \beta + Q_c^{-1} E(x_i e 1(|x_i| \le c)).$$

3.

(a) These results follow from the fact that the sample mean is an unbiased estimator of the population mean, regardless of the distribution the data. Observe that

$$E\left(\overline{y}\right) = Ey_i = \mu.$$

Hence

$$\tau_n = E\left(\overline{y} - \mu\right) = \mu - \mu = 0.$$

Let y_i^* be a random variable with distribution F_n, and $\overline{y}^*$ the sample mean of a random sample $\{y_1^*, ..., y_n^*\}$. By linearity,

$$\begin{aligned} E\overline{y}^* &= Ey_i^* \\ &= \sum_{j=1}^{n} y_j P\left(y_i^* = y_j\right) \\ &= \frac{1}{n}\sum_{j=1}^{n} y_j = \overline{y} \end{aligned}$$

The bootstrap estimate of bias treats $\hat{\mu} = \overline{y}$ as the true value. Hence

$$\tau_n^* = E\overline{y}^* - \hat{\mu} = \overline{y} - \overline{y} = 0.$$

(b) The bias is $E\hat{\mu}^2 - \mu^2$. This is the variance of $\hat{\mu} = \overline{y}$, which we know is $n^{-1}\sigma^2$. In detail,

$$\begin{aligned} \tau_n &= E\hat{\mu}^2 - \mu^2 \\ &= E\left(\hat{\mu} - \mu\right)^2 \\ &= E\left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \mu\right)\right)^2 \\ &= \frac{1}{n^2}\sum_{i=1}^{n} E\left(y_i - \mu\right)^2 \\ &= \frac{\sigma^2}{n}. \end{aligned}$$

(c) Bonus Question: The variance of the EDF is $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2$. Thus the bootstrap estimate τ_n^* of τ_n is

$$\tau_n^* = \frac{\hat{\sigma}^2}{n}.$$

In this case, you do not need to do a simulation to calculate τ_n^*.