Exam Computational Methods in Econometrics

Minor Applied Econometrics
Faculty of Economics and Business Administration
Wednesday, October 21, 2020

Exam:	Computational Methods in Econometrics
Code:	E_EOR3_CME
Coordinator:	M.H.C. Nientker
Date:	October 21, 2020
Time:	12:15
Duration:	2 hours
Number of questions:	8
Type of questions:	Multiple choice & open
Answer in:	English or Dutch
Grades:	Made public within 10 working days
Number of pages:	6, including front page

IMPORTANT

This exam has two different types of multiple choice questions. I expect answers to be as follows.

- "**show why it is the correct answer**": Report your choice and then fully derive why that answer is the correct one. Make sure to show all your steps.
- "**explain why the other answers perform worse**": Report your choice and then explain for each of the other alternatives why it is a worse idea than your choice.

Good luck!

Question 1. We have a nonparametric regression model with autoregressive errors

$$y_t = x_t'\beta + \varepsilon_t, \qquad \varepsilon_t = \rho\varepsilon_{t-1} + \nu_t, \qquad \nu_1, \ldots, \nu_n \sim \text{IID}(0, \sigma_\nu^2).$$

To test H_0: $\mathring{\rho} = 0$ versus $H_1 : \mathring{\rho} \neq 0$ we use the test statistic

$$T(\overrightharpoon{Y}) = \frac{\sum_{t=2}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=2}^n \hat{\varepsilon}_{t-1}^2},$$

where $\hat{\varepsilon} = (I_n - X(X'X)^{-1}X')y = (I_n - X(X'X)^{-1}X')\varepsilon$. Choose one of the following options and **show why it is the correct answer**.

A. The test statistic $T_n(\overrightharpoon{Y})$ is not a pivot under the null.

B. The test statistic $T_n(\overrightharpoon{Y})$ is a pivot under the null.

Solution. The correct answer is option A. The test statistic is not a pivot under the null because the statistical model is nonparametric. Any distribution with finite second moment is included, which means that the test statistic can't possibly have the same distribution under all the distributions under the null. Take for instance a two-point distribution for $\varepsilon_t = \nu_t$ versus a continues distribution such as a Gaussian one.
6 points for correct explanation.

Question 2. In finance people are very interested in calculating the risk of loss of investment portfolios, especially the maximum possible loss within a reasonable probability. To do so they have defined the Value at Risk (VaR) of a loss random variable Y with known cdf F as

$$\mathrm{VaR}_\alpha(Y) = c(F) = \inf\{x \in \mathbb{R} : \; F(x) \geq 1 - \alpha\}.$$

To estimate the characteristic $\vartheta = c(F)$ we simulate $y_1^*, \ldots, y_B^*$ from F to construct the Monte-Carlo distribution function $\hat{F}_B$. As an estimator we then use $\hat{\vartheta} = c(\hat{F}_B)$. Suppose that $n = 200$ and $\alpha = 0.05$. Choose one of the following options and **show why it is the correct answer**.

A. $\hat{\vartheta} = y^*_{(10)}$.

B. $\hat{\vartheta} = y^*_{(190)}$.

C. $\hat{\vartheta} = y^*_{(195)}$.

D. $\hat{\vartheta} = y^*_{(5)}$.

Solution. The correct answer is option B. The function $\hat{F}_B$ is a step function that jumps with size $\frac{1}{B}$ at each of the simulated observations. We then have

$$c(\hat{F}_B) = \inf\{x \in \mathbb{R} : \; \hat{F}_B(x) \geq 1 - \alpha\} = \inf\{x \in \mathbb{R} : \; \hat{F}_B(x) \geq 0.95\} = y^*_{(190)},$$

because that is the point where the function jumps to $0.95 = \frac{190}{200}$.
2 points for writing the estimator, 2 points for realizing $0.95 = \frac{190}{200}$, 2 points for correct explanation.

Question 3. Suppose that we have categorical observable variables $y_1, \ldots, y_n \in \{0, 1\}$ and explanatory variables $x^1, \ldots, x^k$ and that we would like to model the probability $P(Y_1 = 1)$. The LOGIT model specifies this probability as

$$P(Y_1 = 1 \mid x_t) = \frac{1}{1 + e^{-(x_t' \beta)}},$$

where $x_t = (x_t^1, \ldots, x_t^k)$ are treated as fixed and β is a parameter vector. Suppose that we split the vector $\beta' = (\alpha, \gamma)$ and want to test H_0: $\gamma_1 = \ldots = \gamma_m = 0$ versus H_1: $\gamma \neq 0$. We opt for a bootstrap procedure and let $\hat{\alpha}$ be an estimator for α under the null hypothesis. Then the question remains how we can use $\hat{\alpha}$ to simulate $y_1^*, \ldots, y_n^*$. Choose one of the following options and **show why it is the correct answer**.

A. We simulate $u_t \sim \text{Uniform}(0, 1)$ and set $y_t^* = \mathbb{1}\{\frac{1}{1+e^{-x_t'(\hat{\alpha}', 0)}} + u_t \geq 0\}$.

B. We simulate $u_t \sim \text{Uniform}(0, 1)$ and set $y_t^* = \mathbb{1}\{\frac{1}{1+e^{-x_t'(\hat{\alpha}', 0)}} - u_t \geq 0\}$.

C. We simulate $u_t \sim \text{N}(0, 1)$ and set $y_t^* = \mathbb{1}\{\frac{1}{1+e^{-x_t'(\hat{\alpha}', 0)}} - u_t \geq 0\}$.

D. We simulate $u_t \sim \text{N}(0, 1)$ and set $y_t^* = \mathbb{1}\{\frac{1}{1+e^{-x_t'(\hat{\alpha}', 0)}} + u_t \geq 0\}$.

Solution. The correct answer is option B. In that case we have that

$$\begin{aligned} P(Y_t^* = 1) &= P(\mathbb{1}\{\frac{1}{1 + e^{-x_t'(\hat{\alpha}', 0)}} - u_t \geq 0\} = 1) \\ &= P(\frac{1}{1 + e^{-x_t'(\hat{\alpha}', 0)}} - u_t \geq 0) \\ &= P(u_t \leq \frac{1}{1 + e^{-x_t'(\hat{\alpha}', 0)}}) \\ &= \frac{1}{1 + e^{-x_t'(\hat{\alpha}', 0)}}, \end{aligned}$$

where we used the fact that u_t is uniformly distributed for the last probability. Under the null we have γ equal to zero. Therefore as $\hat{\alpha} \to \mathring{\alpha}$ as $n \to \infty$ we see that this is a close approximation to the data generating process under the null.
2 points for correctly rewriting the probability as in the first two lines, 3 points for properly calculating the last probability and using the uniform distribution, 1 point for correct explanation why this indicates that it is a good approximation under the null.

Question 4. Pick one of the following statements and **explain why the other answers are incorrect**.

A. The power of a Monte-Carlo test can exceed that of the theoretical test if we let $B \to \infty$.

B. In a Monte-Carlo hypothesis testing approach we need to make B as large as possible to ensure that the size of the test is as close as possible to the chosen level α.

C. If a test statistic is a pivot it is best to use a bootstrap approximation to the unknown population.

D. In a bootstrap approach to hypothesis testing we prefer to directly derive p-values, but often use Monte-Carlo because the finite sample distributions cannot be derived.

SOLUTION. The correct answer is option D.

- A. is incorrect because the power of the Monte Carlo test converges to that of the theoretical test if we let $B \to \infty$. Therefore it doesn't exceed the power but equals it in the limit.
- B. is incorrect, because the test is of level α for any B needs that satisfies $\alpha(B+1) \in \mathbb{N}$. Therefore we don't have to make B as large as possible for the purpose of size, we do this to increase the power as much as possible.
- C. is incorrect, because we don't need a bootstrap approximation if a test statistic is a pivot. We can directly sample from it's distribution with the Monte-Carlo testing method and so the bootstrap approximation only introduces unnecessary error in the inference.

2 points for each correct explanation.

Question 5. We have a regression model

$$y_t = x_t'\beta + \varepsilon_t, \qquad \varepsilon_1, \ldots, \varepsilon_n \sim \text{NID}(0, \sigma^2).$$

We wish to test H_0: $\mathring{\beta}_1 = \beta_0$ at level α and opt for the t-statistic $T_n(\overrightarrow{Y}) = \frac{\hat{\beta}_{1,OLS} - \beta_0}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_{1,OLS})}}$. Let $\hat{\beta}_{OLS}$ be the OLS estimator for $\mathring{\beta}$ and let $\hat{\beta}$ be an estimator for $\mathring{\beta}$ under the null, i.e. $\hat{\beta}_1 = 0$. Choose one of the following test procedures and **explain why the other answers perform worse**.

A. Calculate the observed t-statistic $T_n(\vec{y})$ and reject H_0 if $|T_n(\vec{y})| > c$, where c is the $1 - \alpha/2$ quantile of the t_{n-1} distribution.

B. Use a pairs bootstrap to obtain a p-value and reject H_0 if this p-value is below α.

C. Calculate residuals $\hat{\varepsilon}_t = y_t - x_t'\hat{\beta}$ for $1 \leq t \leq n$ and simulate $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from their empirical distribution function. Then derive $y_t^* = x_t'\hat{\beta}_{OLS} + \varepsilon_t^*$ and $t^* = T_n(\vec{y}^*)$. Redo this B times and use Monte-Carlo testing to obtain a p-value and reject H_0 if this p-value is below α.

D. Calculate residuals $\hat{\varepsilon}_t = y_t - x_t'\hat{\beta}$ for $1 \leq t \leq n$ and simulate $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from their empirical distribution function. Then derive $y_t^* = x_t'\hat{\beta} + \varepsilon_t^*$ and $t^* = T_n(\vec{y}^*)$. Redo this B times and use Monte-Carlo testing to obtain a p-value and reject H_0 if this p-value is below α.

Solution. The correct answer is option A. Note that we have $\varepsilon_1, \ldots, \varepsilon_n \sim \text{NID}(0, \sigma^2)$. This distributional assumption makes the test statistic a pivot, which means that we can just apply the regular t-test. All the other methods perform worse, because they are approximations of the distribution of the statistic instead of the true t_{n-1} distribution.
6 points for correct explanation.

Question 6. We have an autoregressive model of order three:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t, \qquad \varepsilon_1, \ldots, \varepsilon_n \sim \text{IID}(0, \sigma_\varepsilon^2).$$

We wish to test for a unit root and thus rewrite the model as

$$\begin{aligned} \Delta y_t &= (\phi_1 + \phi_2 + \phi_3 - 1) y_{t-1} - (\phi_2 + \phi_3)\Delta y_{t-1} - \phi_3 \Delta y_{t-2} + \varepsilon_t \\ &= \beta y_{t-1} - (\phi_2 + \phi_3)\Delta y_{t-1} - \phi_3 \Delta y_{t-2} + \varepsilon_t, \end{aligned}$$

where $\Delta y_t = y_t - y_{t-1}$. To test H_0: $\mathring{\beta} = 0$ versus H_1: $\mathring{\beta} \neq 0$ we use the t-statistic

$$T_n(\vec{Y}) = \frac{\hat{\beta}_{OLS}}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_{OLS})}}.$$

Let $(\hat{\beta}_{OLS}, \hat{\phi}_{2,OLS}, \hat{\phi}_{3,OLS})$ be the OLS estimators of the rewritten model and let $(\hat{\phi}_2, \hat{\phi}_3)$ be the OLS estimators under the null hypothesis. Choose the best option out of the following simulation procedures to obtain a simulated test statistic and **explain why the other answers perform worse**.

A. Calculate residuals $\hat{\varepsilon}_t = \Delta y_t + (\hat{\phi}_2 + \hat{\phi}_3)\Delta y_{t-1} + \hat{\phi}_3 \Delta y_{t-2}$ for $3 \leq t \leq n$. Simulate $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from their empirical distribution function. Derive recursively $y_t^* = (1 - \hat{\phi}_2 - \hat{\phi}_3) y_{t-1}^* + \hat{\phi}_2 y_{t-2}^* + \hat{\phi}_3 y_{t-3}^* + \varepsilon_t^*$ and calculate $t^* = T_n(\vec{y}^*)$.

B. Calculate residuals $\hat{\varepsilon}_t = \Delta y_t + (\hat{\phi}_2 + \hat{\phi}_3)\Delta y_{t-1} + \hat{\phi}_3 \Delta y_{t-2}$ for $4 \leq t \leq n$. Simulate $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from their empirical distribution function. Derive recursively $y_t^* = (1 - \hat{\phi}_2 - \hat{\phi}_3) y_{t-1}^* + \hat{\phi}_2 y_{t-2}^* + \hat{\phi}_3 y_{t-3}^* + \varepsilon_t^*$ and calculate $t^* = T_n(\vec{y}^*)$.

C. Calculate residuals $\hat{\varepsilon}_t = \Delta y_t + (\hat{\phi}_2 + \hat{\phi}_3)\Delta y_{t-1} + \hat{\phi}_3 \Delta y_{t-2}$ for $3 \leq t \leq n$. Simulate $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from their empirical distribution function. Derive recursively $y_t^* = (1 - \hat{\phi}_{2,OLS} - \hat{\phi}_{3,OLS}) y_{t-1}^* + \hat{\phi}_{2,OLS} y_{t-2}^* + \hat{\phi}_{3,OLS} y_{t-3}^* + \varepsilon_t^*$ and calculate $t^* = T_n(\vec{y}^*)$.

D. Calculate residuals $\hat{\varepsilon}_t = \Delta y_t + (\hat{\phi}_2 + \hat{\phi}_3)\Delta y_{t-1} + \hat{\phi}_3 \Delta y_{t-2}$ for $4 \leq t \leq n$. Simulate $\varepsilon_1^*, \ldots, \varepsilon_n^*$ from their empirical distribution function. Derive recursively $y_t^* = (1 - \hat{\phi}_{2,OLS} - \hat{\phi}_{3,OLS}) y_{t-1}^* + \hat{\phi}_{2,OLS} y_{t-2}^* + \hat{\phi}_{3,OLS} y_{t-3}^* + \varepsilon_t^*$ and calculate $t^* = T_n(\vec{y}^*)$.

SOLUTION. The correct answer is option B. Option A doesn't work, because when we start at $t = 3$ we would have $\hat{\varepsilon}_3 = \Delta y_3 + (\hat{\phi}_2 + \hat{\phi}_3)\Delta y_2 + \hat{\phi}_3 \Delta y_1$. However Δy_1 contains y_0, which is not observed. Option D works less good, because it reconstructs the simulated y's with $\hat{\phi}_{OLS}$, which doesn't necessarily satisfy the null hypothesis. Therefore the generated data is not a good approximation of the true data generating process under the null. Option C has both of the problems detailed out above.

3 points for correct explanation $t = 3$, 3 points for correct explanation regarding simulating under the null.

Question 7. Pick one of the following statements and **explain why the other answers are incorrect**.

A. Any pivot is also an asymptotic pivot.

B. Parametric models are only useful if we don't know the population $\mathring{F}$. If $\mathring{F}$ is known then it's better to use a nonparametric model.

C. Nonparametric models always outperform parametric models, because it is more likely that the true data generating process is contained in the model.

D. Any asymptotic pivot is also a pivot.

SOLUTION. The correct answer is option A. Option B is incorrect because the purpose of statistical models is to define a set of possible distributions of the random variables behind observed data. If the population is known, then that model should only contain the population, which is not nonparametric. Option C is incorrect, because even though larger models have higher likeliness to contain the true DGP they also have more distributions that we then have to choose from. Ideally we want our model to be as small as possible, given that it still contains the true population. Option D is incorrect as an asymptotic pivot could only be a pivot in the limit, but not necessarily be one for finite n.
2 points for each correct explanation.

Question 8. Let $Y_1 \ldots, Y_n$ be a vector of random variables from an exponential statistical model $\{1 - e^{-\theta y} \mid \theta > 0\}$. It is well known that the maximum likelihood estimator for $\vartheta = c(F) = 1/\mathbb{E}(Y) = \theta_0$ is given by $T_n(\overrightarrow{Y}) = 1/\overline{Y}$. The distribution of $T_n(\overrightarrow{Y})$ depends on ϑ and so we are unable to determine the bias of $T_n(\overrightarrow{Y})$. As an approximation we decide to use the parametric bootstrap function $\hat{F}_n \sim \text{Exp}(1/\overline{y})$ and derive

$$\text{Bias}(T_n(\overrightarrow{Y}), \hat{F}_n) = \mathbb{E}(T_n(\overrightarrow{Y}) \mid \hat{F}_n) - c(\hat{F}_n).$$

Choose one of the following options and **show why it is the correct answer**. You are allowed to use that $\mathbb{E}(1/\overline{Y} \mid \mathring{F}) = \frac{n\theta_0}{n-1}$.

A. $\text{Bias}(T_n(\overrightarrow{Y}), E_n) = \frac{n\overline{y}}{n-1} - \frac{1}{\overline{y}}$.

B. $\text{Bias}(T_n(\overrightarrow{Y}), E_n) = \frac{n\overline{y}}{n-1} - \overline{y}$.

C. $\text{Bias}(T_n(\overrightarrow{Y}), E_n) = \frac{n/\overline{y}}{n-1} - \frac{1}{\overline{y}}$.

D. $\text{Bias}(T_n(\overrightarrow{Y}), E_n) = \frac{n/\overline{y}}{n-1} - \overline{y}$.

SOLUTION. The correct answer is option C. For a random variable $X \sim \hat{F}_n$ we have $\mathbb{E}(X) = \frac{1}{1/\overline{y}} = \overline{y}$. Therefore $c(\hat{F}_n) = 1/\mathbb{E}(X) = 1/\overline{y}$. For the other part we use the hint to quickly derive that

$$\mathbb{E}(T_n(\overrightarrow{Y}) \mid \hat{F}_n) = \mathbb{E}(1/\overline{Y} \mid \hat{F}_n) = \frac{n/\overline{y}}{n-1},$$

where we used that $\hat{F}_B$ is an exponential distribution with $\theta = \frac{1}{\overline{y}}$.
3 points for each correct derivation.