Exam Computational Methods in Econometrics

Minor Applied Econometrics
Faculty of Economics and Business Administration
Thursday, October 24, 2019

Exam:	Computational Methods in Econometrics
Code:	E_EOR3_CME
Coordinator:	M.H.C. Nientker
Date:	October 24, 2019
Time:	8:45
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:	4, 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 <u>fully derive</u> 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 regression model with autoregressive errors

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

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

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

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 $T_n(\vec{Y})$ is not an asymptotic pivot under the null.

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

C. The test statistic $T_n(\vec{Y})$ is an asymptotic pivot under the null, but not a pivot under the null.

Question 2. Let Y be a random variable with known cumulative distribution function F. We want to estimate the characteristic $\vartheta = \mathbb{E}(Y) = c(F)$ and thus simulate $y_1^*, \dots, y_B^*$ from F to construct the Monte-Carlo distribution function $\hat{F}_B$. As an estimator we then use $\hat{\vartheta} = c(\hat{F}_B)$. Choose one of the following options and **show why it is the correct answer**.

A. $\hat{\vartheta} = \frac{1}{B}\sum_{i=1}^B y_i^*$.

B. $\hat{\vartheta} = y^*_{((B+1)/2)}$.

C. $\hat{\vartheta} = \frac{1}{B}\sum_{i=1}^B (y_i^*)^2 - \left(\frac{1}{B}\sum_{i=1}^B y_i^*\right)^2$.

D. $\hat{\vartheta} = \min\{y_1^*, \dots, y_B^*\}$.

Question 3. Suppose that we have a pivotal test statistic $T_n(\vec{Y})$ with one sided rejection region $R_T = (c, \infty)$ for a test at level α. We then perform a Monte Carlo test with B simulations to obtain $t_1^*, \dots, t_B^*$ which together with $T_n(\vec{Y})$ form an ordered row $t^*_{(1)}, \dots, t^*_{(B+1)}$. Let $k \in \mathbb{N}$ be the largest integer such that $\frac{k}{B+1} \leq \alpha$, then the Monte-Carlo approximated rejection region is $\hat{R}_T = (\hat{c}, \infty)$ where $\hat{c} = t^*_{(B+1-k)}$. Choose one of the following options and **show why it is the correct answer**. Hint: derive the probability $P(T_n \in \hat{R}_T \mid H_0)$.

A. We prefer to choose B such that $\alpha(B+1) \in \mathbb{N}$.

B. We prefer to choose B such that $\frac{B+1}{\alpha/2} \in \mathbb{N}$.

C. We prefer to choose B such that $\frac{B+1}{\alpha} \in \mathbb{N}$.

D. We prefer to choose B such that $(\alpha/2)(B+1) \in \mathbb{N}$.

Question 4. We have a regression model $y_t = x_t'\beta + \varepsilon_t$, where $\varepsilon_1, \dots, \varepsilon_n$ are all independent and conditional moments are given by

$$\mathbb{E}(\varepsilon_t \mid x_t) = 0 \qquad \text{and} \qquad \mathbb{V}\text{ar}(\varepsilon_t \mid x_t) = 1 + x_t^2 \qquad \forall t \in \{1, \dots, n\}.$$

To test H_0: $\mathring{\beta}_1 = \beta_0$ at level α we use the t-statistic $T_n(\overrightarrow{Y}) = \frac{\hat{\beta}_1 - \beta_0}{\sqrt{\widehat{\mathbb{V}\text{ar}}(\hat{\beta}_1)}}$. 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 $1 - \alpha$ quantile of the t_{n-1} distribution.

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

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

D. Use a parametric residual bootstrap, by simulating $\varepsilon_1^*, \dots, \varepsilon_n^* \sim N(0, 1)$, to obtain a p-value and reject H_0 if this p-value is below α.

Question 5. We have a regression model

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

We wish to test H_0: $\mathring{\beta}_1 = \beta_0$ at level α and opt for the t-statistic $T_n(\overrightarrow{Y}) = \frac{\hat{\beta}_1 - \beta_0}{\sqrt{\widehat{\mathbb{V}\text{ar}}(\hat{\beta}_1)}}$. To perform the test we use a nonparametric residual bootstrap approach. 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 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 $\hat{\varepsilon}_t = y_t - x_t'\hat{\beta}_{OLS}$ for $1 \le t \le n$ and simulate $\varepsilon_1^*, \dots, \varepsilon_n^*$ from their empirical distribution function. Then derive $y_t^* = x_t'\hat{\beta} + \varepsilon_t^*$ and $t^* = T_n(\vec{y}^*)$.

B. Calculate residuals $\hat{\varepsilon}_t = y_t - x_t'\hat{\beta}$ for $1 \le t \le n$ and simulate $\varepsilon_1^*, \dots, \varepsilon_n^*$ from their empirical distribution function. Then derive $y_t^* = x_t'\hat{\beta}_{OLS} + \varepsilon_t^*$ and $t^* = T_n(\vec{y}^*)$.

C. Calculate residuals $\hat{\varepsilon}_t = y_t - x_t'\hat{\beta}$ for $1 \le t \le n$ and simulate $\varepsilon_1^*, \dots, \varepsilon_n^*$ from their empirical distribution function. Then derive $y_t^* = x_t'\hat{\beta} + \varepsilon_t^*$ and $t^* = T_n(\vec{y}^*)$.

D. Calculate residuals $\hat{\varepsilon}_t = y_t - x_t'\hat{\beta}_{OLS}$ for $1 \le t \le n$ and simulate $\varepsilon_1^*, \dots, \varepsilon_n^*$ from their empirical distribution function. Then derive $y_t^* = x_t'\hat{\beta}_{OLS} + \varepsilon_t^*$ and $t^* = T_n(\vec{y}^*)$.

Question 6. We have a regression model $y_t = x_t'\beta + \varepsilon_t$, where $\mathbb{E}(\varepsilon_t \mid x_t) = 0$ and $\mathbb{V}\text{ar}(\varepsilon_t \mid x_t) = 1 + \varepsilon_{t-1}^2$ for all $1 \le t \le n$. To test H_0: $\mathring{\beta}_1 = \beta_0$ at level α we use the t-statistic $T_n(\vec{Y}) = \frac{\hat{\beta}_1 - \beta_0}{\sqrt{\widehat{\mathbb{V}\text{ar}}(\hat{\beta}_1)}}$. 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 $1 - \alpha$ 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. Use a sieve bootstrap to obtain a p-value and reject H_0 if this p-value is below α.

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

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

A. A test statistic cannot be a pivot under the null if the full statistical model is nonparametric.

B. Nuisance parameters in a model are parameters that are unobserved, but do not influence the distribution of T_n under the null.

C. To perform a Monte-Carlo testing procedure we need our test statistic to be an asymptotic pivot.

D. A pivot is a test statistic whose distribution is the same for all possible data generating processes in the model.

Question 8. Suppose that we have data $Y_1 \ldots, Y_n$ from an unknown population $\mathring{F}$ and that we have estimated some characteristic $\vartheta = c(\mathring{F})$ with some estimator $\hat{\vartheta} = T_n(\vec{Y})$. We expect that our estimator might be biased and so we would like to apply a correction using the empirical distribution function $\hat{F}_n$. Let $\vartheta_1^*, \ldots, \vartheta_B^*$ be test statistics obtained by simulating $y_1^*, \ldots, y_n^*$ from $\hat{F}_n$ and setting $\vartheta^* = T_n(\vec{y}^*)$. Choose one of the following options and **show why it is the correct answer**.

A. The bias correction is given by $\hat{\vartheta}_{BR} = \hat{\vartheta} - 2\frac{1}{B}\sum_{i=1}^B \vartheta_i^*$.

B. The bias correction is given by $\hat{\vartheta}_{BR} = \hat{\vartheta} - \frac{1}{B}\sum_{i=1}^B \vartheta_i^*$.

C. The bias correction is given by $\hat{\vartheta}_{BR} = 2\hat{\vartheta} - 2\frac{1}{B}\sum_{i=1}^B \vartheta_i^*$.

D. The bias correction is given by $\hat{\vartheta}_{BR} = 2\hat{\vartheta} - \frac{1}{B}\sum_{i=1}^B \vartheta_i^*$.