Departments of Economics and of Agricultural and Applied Economics

Ph.D. Qualifying Exam June 2023

PART III:

Econometrics

June 15, 2023

Question 1 (30 minutes)

(a) Consider the simple Normal model:

$$X_t \backsim \ \mathsf{NIID}(\mu, \sigma^2),\ t{=}1, 2, ..., n, ...,$$

with σ^2 is known; 'NIID' stands for 'Normal, Independent and Identically Distributed' with mean μ and variance σ^2.

(1) The sampling distribution of $\overline{X}_n{=}\frac{1}{n}\sum_{t=1}^{n} X_t$ is traditionally stated as:

$$\overline{X}_n \backsim \ \mathsf{N}(\mu, \tfrac{\sigma^2}{n}).$$

Explain why this claim makes no sense as it stands since it's not obvious why $E(\overline{X}_n){=}\mu$ and $Var(\overline{X}_n){=}\frac{\sigma^2}{n}$.

(2) Using your answer in (i) state the sampling distributions of the test statistic $d(\mathbf{X}){=}\frac{\sqrt{n}(\overline{X}_n-\mu_0)}{\sigma}$ under both the hypotheses:

$$H_0{:}\ \mu \leq \mu_0, \text{ vs.} \quad H_1{:}\ \mu > \mu_0. \tag{1}$$

(b)-(1) Using your answer in (a) (i)-(ii) state the sampling distribution of the pivot $d(\mathbf{X};\mu){=}\frac{\sqrt{n}(\overline{X}_n-\mu^*)}{\sigma}$, where μ^* denotes the 'true' value of μ, whatever that happens to be.

(2) Using your answer to (b)(i) construct an $(1{-}\alpha)$ Confidence Interval (CI) for μ and discuss its optimality.

(c)-(1) Using your answer in (a)-(ii) state the optimal Neyman-Pearson (N-P) test in terms of $d(\mathbf{X})$ and explain what 'optimal' means in this case.

(2) Define and explain briefly the concepts of (i) type I error probability, (ii) power of the test, (iii) the p-value, and (iv) compare and contrast (i) with (iii).

(d) State the fallacies of acceptance and rejection and explain why the accept/reject rules and the p-value are vulnerable to these fallacies when they are interpreted as providing evidence for or against H_0.and H_1.

Question 2: Regression with individual-specific intercept (20 points)

Consider a sample of $i = 1 \dots N$ individuals, where each individual contributes $t = 1 \dots T$ observations. Each individual's set of T observations is called a *panel*. Assume the data are stacked by panels, such that the first T observations come from individual 1, the next T observations come from individual 2, and so on. Let the full sample size be labeled as n, with $n = N * T$.

At the single-observation level, the *true model* is given as:

$$\begin{aligned} &y_{it} = c_i + \mathbf{x}_{it}'\boldsymbol{\beta} + \epsilon_{it}, \quad \text{with} \\ &\epsilon_{it} \sim n\left(0, \sigma^2\right) \end{aligned} \tag{1}$$

As is clear from (1), each individual has his/her own intercept term c_i, which is fixed but unknown.

For the following, let $\mathbf{i}_T$ be a vector of ones of length T, $\mathbf{i}_n$ be a vector of ones of length n, $\mathbf{I}_T$ the identity matrix of dimension T, $\mathbf{I}_n$ the identity matrix of dimension n, $\boldsymbol{\epsilon}_i = \begin{bmatrix} \epsilon_{i1} & \epsilon_{i2} & \dots & \epsilon_{iT} \end{bmatrix}'$ the panel-specific vector of error terms, and $\boldsymbol{\epsilon} = \begin{bmatrix} \boldsymbol{\epsilon}_1' & \boldsymbol{\epsilon}_2' & \dots & \boldsymbol{\epsilon}_N' \end{bmatrix}'$ the full-sample vector of error terms.

Part (a), 9 points

1. Assume the analyst erroneously estimates a generic model with a typical single-valued intercept term, i.e.
$$y_{it} = \alpha + \mathbf{x}_{it}'\boldsymbol{\beta} + \omega_{it} \tag{2}$$
Show the explicit form of ω_{it} and derive it's expectation and variance.

2. Let the model at the panel level for person i be written as:
$$\begin{aligned} &\mathbf{y}_i = \mathbf{i}_T * \alpha + \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\omega}_i \quad \text{with} \\ &\mathbf{y}_i = \begin{bmatrix} y_{i1} & y_{i2} & \dots & y_{iT} \end{bmatrix}', \\ &\mathbf{X}_i = \begin{bmatrix} \mathbf{x}_{i1}' \\ \mathbf{x}_{i2}' \\ \vdots \\ \mathbf{x}_{iT}' \end{bmatrix} \end{aligned} \tag{3}$$
Show the explicit form of $\boldsymbol{\omega}_i$ and derive it's expectation and variance.

3. Let the model at the full-sample level be written as:
$$\begin{aligned} &\mathbf{y} = \mathbf{i}_n * \alpha + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\omega} \quad \text{with} \\ &\mathbf{y} = \begin{bmatrix} \mathbf{y}_1' & \mathbf{y}_2' & \dots & \mathbf{y}_T' \end{bmatrix}', \\ &\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_T \end{bmatrix} \end{aligned} \tag{4}$$
Show the explicit form of $\boldsymbol{\omega}$ and derive it's expectation and variance.

Part (b), 6 points
Using partitioned regression based on the wrong / naive model in (2), show that the OLS estimate of $\boldsymbol{\beta}$ - call it $\mathbf{b}$ - is biased.

Part (c), 5 points
Now assume the *true model* is given as:

$$\begin{aligned} & y_{it} = c_i + \mathbf{x}_{it}'\boldsymbol{\beta} + \epsilon_{it}, \quad \text{with} \\ & \epsilon_{it} \sim n\left(0, \sigma^2\right) \quad \text{and} \\ & c_i \sim n\left(0, \sigma_c^2\right) \end{aligned} \tag{5}$$

In words, the individual-specific intercepts are now coming from a common normal distribution with mean zero and variance σ_c^2, where $\sigma_c^2 \neq \sigma^2$. You can assume that the two stochastic components (c_i and ϵ_{it}) are uncorrelated across all observations.

1. Assume again the analyst uses the flawed model in (2). Expressing the model again at the observation, panel, and full-sample level as in equations (2) through (4), show the explicit forms of ω_{it}, $\boldsymbol{\omega}_i$, and $\boldsymbol{\omega}$, as well as their expectation and variance.
2. In addition to biased coefficients, which other problem presents itself in this case compared to a well-behaved CLRM?

Question 3: Bayesian problem (20 points)

Consider the Exponential model for a random variate y with parameter λ (also called the "hazard rate"), given as

$$\begin{aligned} &p(y|\lambda) = \lambda\, exp(-\lambda y) \quad \text{with} \\ &E(y|\lambda) = \lambda^{-1}, \quad V(y|\lambda) = \lambda^{-2}, \quad \lambda > 0 \end{aligned} \tag{1}$$

Part (a), 4 points
Now consider a sample of n observations from this distribution, with each observation generically labeled $y_i, i = 1 \ldots n$.

Suppose you stipulate a *gamma* prior density for λ with shape parameter a and inverse scale ("rate") parameter b, given as

$$\begin{aligned} &p(\lambda) = g(a,b) = \frac{b^a}{\Gamma(a)}\lambda^{(a-1)} exp(-b\lambda), \quad \text{with} \\ &E(\lambda) = \frac{a}{b},\ V(\lambda) = \frac{a}{b^2}, \quad \lambda, a, b > 0, \end{aligned} \tag{2}$$

1. Write down the joint distribution for the sample data (in *un*-logged form). Call it $p(\mathbf{y}|\lambda)$.
2. Show that the posterior distribution of λ, given your collected data from the Exponential, is also a gamma. Show the form of the posterior shape and rate parameters (you can call them a^* and b^*).

Part (b), 10 points

1. Derive the posterior expectation of λ.
2. Derive the MLE estimate for λ, call it $\hat{\lambda}$.
3. Show that the posterior expectation can be written as a weighted average of the prior expectation and the MLE estimate. What happens to this posterior expectation as $n \to \infty$?

Part (c), 6 points
Assume you study the waiting time by motorists at an urban toll booth (where you pay a fee to use the highway). Let y_i be the waiting time by vehicle i, in minutes. In terms of prior information, assume you know from other nearby toll booths at similar highways in terms of traffic volume that the average waiting time over the last 12 months was 1 minute.

Your own data shows that of 1000 vehicles sampled over the last week, the sum of all wait times was 800 minutes.

1. Setting $b = 10$, find the prior shape parameter a to set the prior expectation such that it corresponds to the *inverse* of average waiting time from the other toll booths (given the relationship between $E(y|\lambda)$ and λ in (1).

2. Solve for the MLE estimate $\hat{\lambda}$, and the posterior expectation $E(\lambda|\mathbf{y})$, with precision to the third decimal. Why are they different? (Hint: You can use your results from part (b) for guidance). Solve for the corresponding expected waiting times for MLE and Bayes (use again the relationship in (1)).

3. Now assume your sample was much smaller, say 10 vehicles with a total waiting time of 8 minutes. Solve again for $\hat{\lambda}$ and $E(\lambda|\mathbf{y})$, with precision to 3 decimals. How do they compare to the larger-sample results? Has the difference between them increased? If so, why? Solve for the corresponding expected waiting times for MLE and Bayes (use again the relationship in (1)). How do they compare to the waiting times based on the larger sample?