Advanced Econometrics 2:

June 28, 2020

Final Exam

This exam has 7 questions. It is necessary to answer all questions in order to get full marks (100 points). Please show your derivations for each question (answers without derivations will NOT be given marks). Please write down your name and page numbers on your answer sheet.

1. [10 points] Consider a classical linear regression model in the form of
$$y_i = \beta_1 + x_i'\beta_2 + u_i$$
where y_i is a scalar, x_i is a $k \times 1$ vector, and u_i is error term.
Please show that $\beta_2 = [Var(x_i)]^{-1}Cov(x_i, y_i)$.

2. Consider the logit regression model. The conditional probability is modeled as
$P(y_i = 1|x_i) = \frac{e^{x_i'\beta}}{1+e^{x_i'\beta}}$.

   a) [5 points] Please construct the log likelihood function for the logit regression model.

   b) [15 points] Please prove to identify whether the GMM and MLE will generate the same estimate of the unknown parameters of the logit regression model?

3. [10 points] You are given the following function:
$$L(y-c) = \begin{cases} \alpha|y-c|, & y-c \geq 0 \\ (1-\alpha)|y-c|, & y-c < 0 \end{cases}$$
where y is a continuous random variable and α is a constant. Please prove to identify whether minimizing the expectation of $L(y-c)$ with respect to c will give us: $c = Quant_\alpha(y)$?

4. Suppose a time series $\{X_t\}_{t=1}^{n}$ has the following representation:

$$X_t = \phi X_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_t \sim WN(0, \sigma^2), t = 1,2, \ldots, n$$

, where $|\phi| < 1$ and $|\theta| < 1$.

   a) [5 points] In class we declare $\hat{\phi}_{OLS}$ is an inconsistent estimator of ϕ. Why does the OLS estimation fail to be consistent?

   b) [10 points] Actually $\hat{\phi}_{OLS} \xrightarrow{P} \rho$ as $n \to \infty$. What is this ρ? (You are asked to derive the expression of ρ that only contains number, θ, and ϕ)

5. [10 points] For the classical random-effect model that we have learned in the class:

$$y_{it} = x_{it}'\beta + \alpha + \epsilon_{it} + u_i$$

where, x_{it} is $k \times 1$ and strictly exogenous; $\epsilon_{it} + u_i$ is the error term; $Var(\epsilon_{it}) = \sigma_\epsilon^2$ and σ_ϵ^2 is unknown. We can estimate the σ_ϵ^2 by:

$$\hat{\sigma}_\epsilon^2 = \frac{1}{n(T-1)-k} \sum_{i=1}^{n} \sum_{t=1}^{T} \ddot{\hat{\epsilon}}_{it}^2$$

where, $\ddot{\hat{\epsilon}}_{it} = \hat{\epsilon}_{it} - \bar{\hat{\epsilon}}_i$.

Please prove that $\hat{\sigma}_\epsilon^2$ is a consistent estimator of σ_ϵ^2.

6. [15 points] Suppose the observable y is generated as follows:

$$y = \begin{cases} a, & y^* \le a \\ y^*, & y^* > a \end{cases}$$

where y^* is the latent continuous variable, and $y^* \sim N(\mu, \sigma^2)$.

Please derive: $E(y) =?$

7. [20 points] Suppose there is a CLT that you can use for the following derivation. Please derive the asymptotic distribution for the multivariate nonparametric kernel density estimator $\hat{f}(x_1, x_2) = \frac{1}{nh_1h_2} \sum_{i=1}^{n} k\left(\frac{x_{i1}-x_1}{h_1}\right) k\left(\frac{x_{i2}-x_2}{h_2}\right)$, where h_1, h_2 are bandwidths, x_1, x_2, x_{i1}, x_{i2} are scalars and $k(\,)$ is the 2nd order Gaussian kernel function.