國立臺灣大學 112 學年度碩士班招生考試試題

※ 注意：請於試卷內之「非選擇題作答區」作答，並應註明作答之題號。

1. (10 points) Let $(Z_1, Z_2) \sim^{i.i.d.} N(0,1)$.

   (a) Define
   $$Y = |Z_1|$$
   Use the distribution function technique to find the pdf of Y.

   (b) Define
   $$X_1 = \frac{Z_1 + Z_2}{2}, \quad X_2 = \frac{Z_1 - Z_2}{2}, \quad X_3 = \frac{Z_2 - Z_1}{2}$$
   Find the joint distribution of (X_1, X_2, X_3).

2. (10 points) Let X be a continuous random variable with pdf $f_X(c)$ and distribution function $F_X(c)$. Suppose X is truncated to satisfy $X > a$.

   (a) Find the conditional density $f_{X|X>a}(c)$.

   (b) Now suppose that $X \sim U[0,1]$ and $0 < a < 1$. Find $Var[X|X > a]$.

3. (30 points) Suppose we want to estimate θ, the population proportion of Taiwanese citizens who have ever driven after drinking alcohol, where $\theta \in [0,1]$. However, people may not be truthful in their responses; i.e., they may respond NO when, in fact, they are guilty of drunk driving. This is called "evasive answer bias".

   To solve this problem, we use the following randomized response methodology of interviewing people about sensitive issues. A random sample of n people is drawn from the population. We ask respondents to first flip a unbiased coin, and have them hide the result from us (i.e., do not reveal to us the outcome of the coin tossing).

   - Respondent i who flip head ($H_i = 1$) will answer YES/NO to the question: "Have you ever driven after drinking alcohol?"
   - Respondent i who flip tail ($H_i = 0$) will answer YES/NO to the question: "Is the last digit of your National ID number a 0, 1, 2, or 3?"

   Let $X_i = 1$ represent the response is YES from this questionnaire, and $X_1, X_2, \ldots, X_n$ denote the random sample. The novelty of this strategy is that when a respondent answers YES, we do not know which question he or she answers. Hence, respondents have the incentive to simply tell the truth, and we assume that these YES and NO reports are made truthfully.

   (a) Find the probability $P(X_i = 1)$, denoted by p.

   (b) Use $X_1, X_2, \ldots, X_n$ to find the maximum likelihood estimator of θ.

   (c) Find the limiting distribution of $\sqrt{n}(\hat{\theta} - \theta)$. You should mention what properties and theorems you are using.

見背面

(d) In order to construct a confidence interval, propose a pivot, and derive its limiting distribution. You should mention what properties and theorems you are using.

(e) Now suppose that $n = 100$, and $\sum_{i=1}^{100} X_i = 53$. Use the proposed pivot to construct a 95% approximate confidence interval (interval estimate) of θ.

(f) Use the above confidence interval to test $H_0 : \theta = 0.6$ vs. $H_1 : \theta \neq 0.6$.

Note:

$$P(N(0,1) \leq 1.28) = 0.90, \quad P(N(0,1) \leq 1.64) = 0.95$$
$$P(N(0,1) \leq 1.96) = 0.975, \quad P(N(0,1) \leq 2.33) = 0.99$$

4. (25 points(5,10,10))
Consider the model

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + u$$
$$E(u|x) = 0,$$

where $E(x) = 0$, $E(x^2) = 1$, and $E(x^3) = 0$.
If we write

$$y = \alpha_0 + \beta_1 x + v.$$

(1) Find α_0 and v.
(2) If $\hat{\beta}_1$ is an OLS estimator for β_1 from the regression y on x, is $\hat{\beta}_1$ a consistent estimator for β_1? Explain.
(3) Is $\hat{\beta}_1$ an unbiased estimator for β_1? Explain.

5. (10 points)
Consider the model

$$y_1 = z_1\delta_1 + \alpha_1 y_2 + \alpha_2 y_2^2 + u_1$$
$$E(u_1|z) = 0$$

where z is $1 \times L$ vector of exogenous explanatory variables which includes unity as its first element, z_1 is $1 \times L_1$ strict sub-vector of z that also includes unity as its first element. z contains at least one element not in z_1. y_2 is a scalar endogenous explanatory variable. In the first stage, we run the regression of y_2 on z and get the fitted value $\hat{y}_2$, and in the second stage we run the regression of y_1 on z_1, $\hat{y}_2$ and $\hat{y}_2^2$. Is this a correct procedure for consistently estimating δ_1, α_1 and α_2? Why or why not.

6. (15 points (5,5,5))

In the panel data models with a single explanatory variable, the between estimator is the OLS estimator of the following equation

$$\overline{y}_i = \beta_0 + \beta_1 \overline{x}_i + a_i + \overline{u}_i$$

where the overbar represents the average over tome. Assume that $E(a_i) = 0$ and $\overline{u}_i$ is uncorrelated with $\overline{x}_i$, but $Cov(x_{it}, a_i) = \sigma_{xa}$ for all t and i. Letting $\widetilde{\beta}_1$ be the OLS estimator (between estimator) using the time average.

(1) Show that $plim \widetilde{\beta}_1 = \beta_1 + \sigma_{xa}/Var(\overline{x}_i)$, where $plim$ is the probability limit defined as $N \to \infty$.

(2) Assume also that the x_{it} are uncorrelated with a constant variance σ_x^2 for all $t = 1, 2, ..., T$. Show that $plim \widetilde{\beta}_1 = \beta_1 + T(\sigma_{xa}/\sigma_x^2)$.

(3) Is the inconsistency for the between estimator smaller when there are more time periods and the explanatory variables are not very highly correlated over time? What about the sign of the inconsistency for the between estimator?

試題隨卷繳回