University of Pennsylvania **Economics 705, Fall 2012**

Prelim Examination

Friday August 17, 2012, Time limit: 150 minutes

Instructions:

(i) The exam consists of two parts. The total number of points for each part is 50. The number of points for each question is given below.

(ii) The exam is closed book and closed notes.

(iii) To receive full credit for your answers you have to explain your calculations. **You may state additional assumptions.**

Part I

Question 1: Change of Variables (11 Points)

Suppose that $X \sim N(\mu, \sigma^2)$. Let $Y = \exp(X)$.

(i) (4 Points) Suppose that $X \sim N(\mu, \sigma^2)$. Let $Y = \exp(X)$. Compute the expected value $\mathbb{E}[Y]$.

(ii) (3 Points) Suppose that $X \sim N(\mu, \sigma^2)$. Let $Y = \exp(X)$. Derive the probability density function of Y.

(iii) (4 Points) Suppose that X_1 and X_2 are independent and have $N(0, 1)$ distributions. Define $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$. Use a change-of-variables argument to obtain the joint probability density for (Y_1, Y_2). Then compute the pdf for the marginal distribution of Y_1. What is the distribution of X_1+X_2?

Question 2: Linear Regression Model (14 Points)

Consider the linear regression model

$$y_i = x_i\theta + u_i, \quad u_i|x_i \sim iid(0, 1), \quad x_i \geq \epsilon > 0, \quad \mathbb{E}[x_i^2] = Q, \quad i = 1, \ldots, n$$

Moreover, the x_i's are also independent across i. Notice that we assumed that the conditional variance of u given x is known to be one. Moreover, $k = 1$ and both x_i and θ_i are scalar.

(i) (3 Points) Derive the likelihood function and the maximum likelihood estimator $\hat{\theta}$ under the assumption that the u_i's are in fact normally distributed.

(ii) (3 Points) Show that the MLE $\hat{\theta}$ is consistent even if the observations are not normally distributed.

(iii) (6 Points) Based on the assumption of normality, define the likelihood-ratio test statistic for the null hypothesis that $\theta = \theta_0$ and derive the large sample distribution of the likelihood ratio statistic (under the null hypothesis, without assuming that the data are normally distributed).

(iv) (2 Points) What are the acceptance and rejection regions for the LR test given a type-I error of $\alpha = 0.05$?

Question 3: Frequentist Inference on a Bounded Parameter Space (13 Points)

Consider the following experiment:

$$X \sim N(\theta, 1), \quad \theta \in \mathbb{R}^+.$$

Be aware of the constraint $\theta \geq 0$. You can use $\Phi(x)$ to denote the cumulative density function of a $N(0,1)$.

(i) (2 Points) Derive the likelihood function and the maximum likelihood estimator for this experiment.

(ii) (3 Points) What is the sampling distribution of the maximum likelihood estimator? Hint: it might be helpful to draw a graph.

(iii) (8 Points) Propose a test for the null hypotheses (i) $H_0: \ \theta = 0$; (ii) $H_0: \ \theta = 10$ versus the alternative that $\theta \neq 0$. For each test, report

- the test statistic;
- the sampling distribution of the test statistic under H_0;
- the critical value that guarantees a type-I error α;
- the power of the test against alternatives $\theta = \delta$ (for hypothesis (i)) and $\theta = 10 + \delta$ (for hypothesis (ii)).

Question 4: Bayesian Inference on a Bounded Parameter Space (12 Points)

Consider the following experiment:

$$X \sim N(\theta, 1), \quad 0 \leq \theta \leq C.$$

Be aware of the constraints on θ. The prior distribution of θ is uniform on the interval $[0, C]$. You can use $\Phi(x)$ to denote the cumulative density function of a $N(0,1)$.

(i) (4 Points) Derive the posterior density $p(\theta|X)$, including the normalization constants that ensure that the density integrates to one.

(ii) (4 Points) Consider the loss function

$$L_\epsilon(\theta, \delta) = \begin{cases} 0 & \text{if } |\theta - \delta| \leq \epsilon \\ 1 & \text{otherwise} \end{cases}$$

Find the Bayes estimator that minimizes the posterior expected loss. Hint: drawing the posterior for various choices of X might help.

(iii) (4 Points) Compare the posterior expected risk of the Bayes estimator and the maximum likelihood estimator for $X = 0$ and $X = 10$ (assuming that $C >> 10$).

Part II

Question 5: Linear Model with Endogeneity (20 points)

The model is

$$\begin{aligned} y_i &= x_i'\beta + u_i, \\ E(z_i u_i) &= 0, \end{aligned} \tag{1}$$

where $x_i, z_i, \beta \in R^k$, $y_i, u_i \in R$, $\{(x_i, z_i, y_i, u_i) : i = 1, ..., n\}$ are *iid*, and $Ex_i x_i'$ and $Ex_i z_i'$ both have full rank k. Let $\widehat{\beta}_n$ denote the least-squares estimator obtained by regressing y_i on x_i, and let $\widetilde{\beta}_n$ denote the two-stage-least-squares (TSLS) estimator using the instruments z_i.

(i) (4 points) Define $\delta_n = \widehat{\beta}_n - \widetilde{\beta}_n$. Derive the probability limit $\delta = \lim_{n\to\infty} \delta_n$.

(ii) (2 points) Propose a condition under which $\delta = 0$ in (i).

(iii) (5 points) Derive the asymptotic distribution of δ_n when $\delta = 0$.

(iv) (2 points) Propose an estimator of the asymptotic variance derived in (iii).

(v) (5 points) Suppose the condition in (1) is misspecified such that $E(z_i u_i) \neq 0$. Instead,

$$u_i = cn^{-1/2} + e_i \text{ and } E(z_i e_i) = 0 \text{ for some constant } c \neq 0.$$

Under this misspecification condition, derive the asymptotic distribution of the TSLS estimator.

Question 6: GMM Estimation (14 points)

The equation of interest is

$$\begin{aligned} y_i &= x_i'\beta + e_i, \text{ where} \\ E(x_i e_i) &= 0, \\ E(q_i e_i) &= 0, \end{aligned}$$

$y_i \in R$, $x_i \in R^l$, $q_i \in R^k$, and $l, k > 3$.

(i) (5 points) Show how to construct an *efficient* GMM estimator for β. Be clear on how to construct a feasible optimal weight matrix.

(ii) (5 points) Let $\beta_1, \beta_2, \beta_3$ denote the first three elements of β, respectively. We are interested in testing the null hypothesis:

$$\begin{aligned} H_0 &: \beta_1 = -\beta_2^2 \text{ and } \beta_1 = -\beta_3 \text{ vs} \\ H_1 &: \text{at least one of the equalities does not hold.} \end{aligned}$$

Provide a Wald statistic for this test and derive the asymptotic distribution of this Wald statistic.

(iii) (4 points) How to conduct a specification test on the exogeneity of x_i and q_i. Be clear on the test statistic, the critical value, and the acceptable/rejection rule. Suppose the size of the test is α.

Question 7: Truncated Observations (16 Points)

Consider a model on the purchase of durable goods. The latent variable y_i^* satisfies

$$y_i^* = x_i'\beta + u_i, \;\; u_i|x_i \sim N(0,1),$$

where $x_i \in R^k$ and $\{(x_i, u_i) : i = 1, ..., n\}$ are *iid*. The observations y_i are truncated because only consumers who purchased durable goods are sampled, i.e.,

$$y_i = y_i^* \text{ if } y_i^* > 0.$$

(i) (5 Point) Write down the log-likelihood function for the maximum likelihood estimator $\widehat{\beta}$.

(ii) (5 Points) Give the limit distribution of the maximum likelihood estimator $\widehat{\beta}$.

(iii) (3 Points) Estimate the standard error of the maximum likelihood estimator.

(iv) (3 Points) Write down the LR test statistic and its limit distribution for

$$H_0 : \beta = 0 \;\; vs \;\; H_1 : \beta \neq 0.$$

END OF EXAM