University of Pennsylvania **Economics 705, Fall 2013**

Prelim Examination

Friday August 9, 2013, 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: Testing and Coverage Sets (12 Points)

(i) (4 Points) Suppose you have a test $\varphi(Y;\theta_0)$ of a null hypothesis $H_0: \ \theta = \theta_0$ with type-I error α. Explain how this test can be "inverted" to obtain a $1-\alpha$ confidence set.

(ii) (4 Points) Consider the location model $Y \sim N(\theta, 1)$. Propose a test for the hypothesis $H_0: \ \theta = 0$ with type-I error α and derive its power against the alternative $\theta = \tilde{\theta} \neq 0$.

(iii) (4 Points) Provide an example in which Bayesians and frequentists would report credible and confidence sets that are numerically identical. Explain carefully.

Question 2: Point Estimation in a Non-standard Setting (20 Points)

Consider the model

$$Y \sim N(\theta, 1)$$

with the constraint on the parameter space $\theta \geq 0$. Moreover, for the Bayesian analysis below suppose the prior distribution is $\theta \sim U[0, M]$ where $M < \infty$ is a large number.

(i) (3 Points) Derive the posterior distribution $p(\theta|Y)$.

(ii) (2 Points) Derive the sampling distribution of $\hat{\theta}|\theta$, where $\hat{\theta}$ is the maximum likelihood estimator of θ. (Make sure to consider the constraint on the parameter space: $\theta \geq 0$).

(iii) (5 Points) Now consider the following form of "limited-information" Bayesian inference. Derive a limited-information posterior by inverting the sampling distribution of $\hat{\theta}|\theta$:

$$p(\theta|\hat{\theta}) \propto p(\hat{\theta}|\theta)p(\theta),$$

where $\propto$ denotes proportionality and $\hat{\theta}$ is the MLE from Part (ii). Hint: when deriving $p(\theta|\hat{\theta})$ distinguish the case $\hat{\theta} = 0$ from the case $\hat{\theta} > 0$.

(iv) (2 Points) Compare the actual posterior $p(\theta|Y)$ and the limited-information posterior $p(\theta|\hat{\theta})$ for $Y = 0$ and $Y = -1$. What is a key difference?

(v) (3 Points) Consider the loss function $L(\theta, \delta) = (\theta - \delta)^2$ and characterize the Bayes estimator based on the actual posterior, denoted by $\hat{\theta}_B$, and based on the limited-information posterior, denoted by $\hat{\theta}_{LI}$.

(vi) (5 Points) Provide a ranking of $\hat{\theta}$ (the MLE), $\hat{\theta}_B$, and $\hat{\theta}_{LI}$ in terms of the integrated risk $r(\mathbb{P}^\theta, \delta(\cdot))$.

Problem 3: Linear Regression Model (18 Points)

Consider the linear regression model

$$y_i = x_i'\theta + u_i, \quad u_i|x_i \sim iidN(0,1), \quad x_i \sim iidN(0,Q). \tag{1}$$

Notice that we assumed that the conditional variance of u given x is known to be one. Suppose that $k = 2$ and we can write $x_i = [x_{1,i}, x_{2,i}]'$ and $\theta = [\theta_1, \theta_2]'$. Our goal is to test the null hypothesis $H_0: \theta_2 = 1$.

(i) (2 Points) Derive the unconstrained maximum likelihood estimator (MLE)

$$\hat{\theta} = \text{argmax}_{\theta \in \mathbb{R}^2} \, l(\theta|Y).$$

(ii) (3 Points) Derive the constrained MLE of θ_1, imposing that $\theta_2 = 1$:

$$\bar{\theta}_1 = \text{argmax}_{\theta_1 \in \mathbb{R}} \, l\big([\theta_1, 1]'|Y\big).$$

(iii) (3 Points) Provide an explicit formula for the likelihood ratio statistic.

(iv) (10 Points) Under the null hypothesis that $\theta_2 = 1$, i.e., $Y = X_1\theta_1 + X_2 + U$) derive an expression for the likelihood ratio statistic of the form

$$LR = U'\big(??\big)U,$$

where ?? does not depend on U. Then derive the limit distribution for the likelihood ratio statistic under the null hypothesis:

$$LR \Longrightarrow ???.$$

Hint: use the Frisch-Waugh-Lovell Theorem.

Note: The Frisch-Waugh-Lovell theorem implies that the residuals from the regressions

$$Y = X_1\beta_1 + X_2\beta_2 + residuals$$

and

$$M_1 Y = M_1 X_2 \beta_2 + residuals$$

are identical. Here $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.

Part II

Question 4: Linear Model with Endogeneity (25 points)

Consider the single equation

$$\begin{aligned} y_i &= x_i\beta + e_i, \\ E(e_i|z_i) &= 0. \end{aligned}$$

The observations $\{(x_i, z_i', y_i) : i = 1, ..., n\}$ are i.i.d., $x_i \in R$, $y_i \in R$, $z_i \in R^k$ with $k > 1$. Assume $E(e_i^2|z_i) = \sigma^2$.

(i) (2 points) Show how to construct a two-stage least squares (TSLS) estimator for β.

(ii) (3 points) Show the TSLS estimator is consistent.

(iii) (5 points) Develop the asymptotic distribution of the TSLS estimator.

(iv) (5 points) Show how to construct a GMM estimator for β.

(v) (5 points) Show the efficient GMM estimator is equivalent to the TSLS estimator in this problem.

(vi) (3 points) How to test whether the instruments z_i are all exogeneous.

(vii) (2 points) Suppose there is an extra variable z_i^* that is independent of z_i and x_i. Explain whether z_i^* should be used as an instrumental variable together with z_i.

Question 5: Covariance Matrix Estimation (5 Points)

Consider the model $y_i = x_i'\beta + e_i$, where $E(z_i e_i) = 0$. Let $\widehat{e}_i = y_i - x_i'\widetilde{\beta}$, where $\widetilde{\beta}$ is a consistent estimator of β. Define a GMM weighting matrix

$$W_n = n^{-1} \sum_{i=1}^{n} z_i z_i' \widehat{e}_i^2.$$

The observations $\{(x_i', z_i', y_i) : i = 1, ..., n\}$ are i.i.d., $x_i \in R^\ell$, $y_i \in R$, $z_i \in R^k$ with $k > \ell$. Show

$$W_n \rightarrow_p \Omega, \text{ where } \Omega = E z_i z_i' e_i^2.$$

Question 6. Confidence Interval (10 points)

Consider the model

$$\begin{aligned} y_i &= x_{1i}\beta_1 + x_{2i}\beta_2 + e_i, \\ E(x_i e_i) &= 0, \end{aligned}$$

where $x_{1i}, x_{2i} \in R$, $x_i = (x_{1i}, x_{2i})' \in R^2$. Define the parameter

$$\theta = \beta_1 \beta_2.$$

(i) (2 points) What is the appropriate estimator of $\widehat{\theta}$ for θ.

(ii) (5 points) Find the asymptotic distribution for $\widehat{\theta}$ under standard regularity conditions.

(iii) (3 points) Show how to construct an asymptotic 95% confidence interval for θ.

Question 7: Censored Observations (10 Points)

Consider the following model

$$y_i^* = x_i'\beta + u_i, \; u_i | x_i \sim iid \; N(0, \sigma^2),$$

where $x_i \in R^k$ with $k > 1$. We do not observe y_i^*, instead we observe

$$y_i = \begin{cases} y_i^* \text{ if } y_i^* < \tau \\ \tau \text{ if } y_i^* \geq \tau, \end{cases}$$

where τ is a known constant. Assume the observations are i.i.d.

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

(ii) (3 Points) Give the asymptotic distribution of the maximum likelihood estimator $\widehat{\beta}$.

(iii) (2 Points) Estimate the standard error of the maximum likelihood estimator.

END OF EXAM