University of Pennsylvania **Economics 705, Spring 2019**

Prelim Examination

Friday June 7, 2019. Time limit: 150 minutes

Instructions:

(i) The total number of points is 60. 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.**

Question 1: (6 Points) Suppose that $X \sim N(0,1)$. Define $Y = \exp(X)$.

(i) (2 Points) Derive the pdf and the cdf for Y. Hint: you can use $\phi(x)$ and $\Phi(x)$ to denote the pdf and cdf, respectively, of a $N(0,1)$ random variable.

(ii) (2 Points) Compute the mean of Y.

(iii) (2 Points) Compute the variance of Y.

Question 2: (5 Points) Consider the following model:

$$Y_i = \theta_i + U_i, \quad U_i \sim N(0,1), \quad i = 1, \dots, N. \tag{1}$$

Define the $N \times 1$ vector of means $\theta^N = [\theta_1, \dots, \theta_N]'$ and the $N \times 1$ vector of observations $Y^N = [Y_1, \dots, Y_N]'$

(i) (1 Point) Specify the likelihood function.

(ii) (1 Point) Derive the maximum likelihood estimator (MLE) of the vector θ^N.

(iii) (2 Points) Is the MLE of θ^N consistent as $N \longrightarrow \infty$?

(iv) (1 Point) Define the compound loss function

$$L(\hat{\theta}^N, \theta^N) = \frac{1}{N} \sum_{i=1}^{N} (\hat{\theta}_i - \theta_i)^2.$$

Compute the expected (frequentist) compound loss of the MLE.

Question 3: (20 Points) Now suppose that we add the following prior distribution to the model in (1):

$$\theta_i \sim iidN(0, 1/\lambda), \quad i = 1, \dots, N. \tag{2}$$

(i) (5 Points) Derive the posterior distribution for the vector θ^N. Are the elements of the θ^N vector *a posteriori* independent?

(ii) (3 Points) Derive the marginal likelihood

$$p(Y^N|\lambda)$$

by integrating out θ^N.

(iii) (6 Points) Let the

$$\hat{\lambda} = \text{argmax } p(Y^N|\lambda). \tag{3}$$

Is $\hat{\lambda}$ a consistent estimator of λ as $N \longrightarrow \infty$? Provide a formal proof and provide a discussion of the effect of the lower bound $\lambda \geq 0$.

(iv) (6 Points) Now consider the following posterior mean estimator (conditional on $\hat{\lambda}$) of θ^N:

$$\tilde{\theta}^N = \mathbb{E}[\theta^N|Y^N, \hat{\lambda}]. \tag{4}$$

Provide a formula for the compound loss (don't take expectations) of $\tilde{\theta}^N$, then let $N \longrightarrow \infty$ and evaluate the limit. Compare your result to Question 2(iv).

Question 4: (10 Points) Consider the following model

$$Y_i \sim N(0, \sigma_i^2), \quad \sigma_i^2 = \frac{2i}{N}, \quad i = 1, \ldots, N. \tag{5}$$

There are no unknown parameters here. The goal is to construct coverage intervals for Y_i (you can think of them as interval forecasts). We restrict ourselves to connected intervals of the form

$$C_i = [c_i, c_i + l_i],$$

where c_i is the lower bound of the interval and l_i is its length. Our goal is to minimize the average length of the coverage intervals:

$$\bar{l} = \frac{1}{N} \sum_{i=1}^{N} l_i.$$

(i) (5 Points) Suppose that we impose the coverage constraint

$$\mathbb{E}[Y_i \in C_i] = 0.95 \quad \text{for all} \quad i = 1, \ldots, N. \tag{6}$$

Using first-order conditions to a constrained optimization problem, show that the collection of intervals

$$C_i^* = [-1.96\sigma_i, \, +1.96\sigma_i], \quad i = 1, \ldots, N.$$

minimize the average length criterion.

(ii) (5 Points) Now consider the alternative coverage constraint which only restricts the average coverage probability for the vector of interval forecasts:

$$\frac{1}{N} \sum_{i=1}^{N} \mathbb{E}[Y_i \in C_i] = 0.95. \tag{7}$$

Show that the collection C_i^* is no longer optimal. What are the key features of the optimal intervals?

Question 5: (19 Points) Consider the following regression model:

$$\begin{aligned} y_i &= x_i\beta + u_{1i}, \quad |\beta| < M \\ x_i &= z_i'\gamma + u_{2i}, \end{aligned} \tag{8}$$

where x_i is an endogenous regressor, z_i is a $k \times 1$ vector of instruments that is independent of u_{1i} and u_{2i} and

$$\left[\begin{array}{c} u_{1i} \\ u_{2i} \end{array} \right] \sim iidN\left(\left[\begin{array}{c} 0 \\ 0 \end{array} \right], \left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array} \right] \right).$$

We assume that the covariance matrix Σ of $u_i = [u_{1i}, u_{2i}]'$ is known. You can use matrix notation and define Y, X, and Z as the vectors/matrices that stack y_i, x_i, and z_i', respectively.

(i) (4 Points) Is the OLS estimator of β consistent? Explain.

(ii) (6 Points) Propose a GMM estimator of β and denote it by $\hat{\beta}$. Show that $\hat{\beta}$ is consistent.

(iii) (7 Points) Derive the limit distribution of $\hat{\beta}$. Express the asymptotic covariance matrix as functions of the primitives of the model: β, γ, Σ, and $\mathbb{E}[z_i z_i']$.

(iv) (2 Points) What is the optimal weight matrix for your estimator?