Part I - 203A

Question I-1

Let $U_1, U_2, \ldots$ be a sequence of i.i.d. random variables having the uniform distribution on $[0, 1]$ and define

$$Y_n = \left( \prod_{i=1}^{n} U_i \right)^{-\frac{1}{n}}$$

Show that

$$\sqrt{n}(Y_n - \exp(1)) \overset{d}{\to} N(0, \exp(1)^2)$$

**Hint:** The integration by parts formula may be useful

$$\int_{[a,b]} u(x)v'(x)dx = u(x)v(x)|^b_a - \int_{[a,b]} u'(x)v(x)dx$$

Question I-2

Consider a sample $\{X_i\}_{i=1}^{n}$ where $X_i$ are independent random variables and $X_i$ is distributed as bivariate normal with mean vector and covariance matrix given by

$$\mu_i = \begin{bmatrix} \mu_i \\ \mu_i \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_i^2 \end{bmatrix}$$

where $\mu_i \in \mathbb{R}$ and $\sigma_i^2 > 0$. Find the MLE of $\{\mu_i\}_{i=1}^{n}$ and $\sigma_i^2$. Show that the MLE for $\sigma_i^2$ is inconsistent as $n \to \infty$.

Question I-3

Missing data is pervasive in economics. This problem studies what can be identified in linear regression models in the presence of missing data.

Consider first the regression model

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \epsilon$$

$$\mathbb{E} [\epsilon | X, Z] = 0$$
If we observed an i.i.d. sample from the joint distribution \((Y, X, Z)\), the vector \(\beta \equiv (\beta_0, \beta_1, \beta_2)\) is identified and we can estimate it using OLS. Suppose now that we observe data on \((Y, X, Z)\) but that some of the data on \(X\) (only) is missing.

One way to formalize this is the following: Let \(\{Y_i, X_i, Z_i, M_i\}_{i=1}^{\infty}\) be an i.i.d. sequence from a distribution \(F\). For each observation \(i\) we observe \((Y, Z, M)\) and \(X(1 - M)\). The variable \(M\) is a binary variable equal to 0 if \(X\) is observed and 1 otherwise. We assume that the data is missing at random in the sense that

\[
E[\epsilon | M, X, Z] = 0
\] (1)

1. A common practice in empirical work (according to a recent estimate, about 20% of papers with missing data problems resort to this method) in such situations is the following: replace \(X\) by 0 whenever it is missing and add a dummy equal to 1 if the observation has a missing \(X\). Formally, researchers run the OLS regression

\[
Y = \alpha_0 + \alpha_1 M + \alpha_2 (1 - M) X + \alpha_3 Z + u
\] (2)

where \(E[uM] = E[u(1 - M)X] = E[uZ] = E[u] = 0\). In this part we will see whether the OLS estimator for \(\alpha_2\) converges to \(\beta_1\). To make some progress, we write

\[
X = \gamma_0 + \gamma_1 Z + v
\]

and assume that in addition to (1)

\[
E[v | M, Z] = 0
\]

Show that the OLS estimator for \(\alpha_2\) will be consistent for \(\beta_1\) if either \(\gamma_1 = 0\) or \(\beta_1 = 0\).

2. Show that if we add the regressor \(MZ\) to (2), then the OLS estimator for \(\alpha_2\) will be consistent for \(\beta_1\).
Part II - 203B

Question II-1

Suppose that
\[ y_i = x_i^3 + \epsilon_i \]
such that \((x_i, \epsilon_i)' i = 1, 2, \ldots\) is i.i.d. and
\[
\begin{pmatrix} x_i \\ \epsilon_i \end{pmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)
\]

1. Let
\[
\tilde{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}
\]
denote the OLS estimator of \(\beta\). What is the probability limit of \(\tilde{\beta}\) as \(n \to \infty\)?

From the model \(y = x_i^3 + \epsilon\), we can calculate the derivative \(d y / d x = 3x_i^2\), based on which we can define the average derivative
\[
E \left[ 3x_i^2 \right]
\]
Is the average derivative equal to the probability limit of \(\tilde{\beta}\)?

2. Suppose now that
\[
\begin{pmatrix} x_i \\ \epsilon_i \end{pmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)
\]

Does your conclusion in the previous question change or not?

Hint: The moment generating function of \(\mathcal{N}(\mu, \sigma^2)\) is \(\exp \left( \mu t + \frac{\sigma^2}{2} t^2 \right)\).

Question II-2

Consider the model
\[ y_i = x_i \beta + \epsilon_i \]
\[ x_i = z_{i1} \pi_1 + z_{i2} \pi_2 + v_i \]
where we assume that (i) \((z_{i1}, z_{i2})'\) is independent of \((\epsilon_i, v_i)';\) (ii) \((\epsilon_i, v_i)'\) has a mean equal to zero; and (iii) \(z_{i1}\) and \(z_{i2}\) are independent of each other with both means equal to zero. We assume that \((x_i, z_{i1}, z_{i2}, \epsilon_i, v_i)'\) is iid. Let
\[
\tilde{\beta}_1 = \frac{\sum_{i=1}^n z_{i1} y_i}{\sum_{i=1}^n z_{i1} x_i}, \quad \tilde{\beta}_2 = \frac{\sum_{i=1}^n z_{i2} y_i}{\sum_{i=1}^n z_{i2} x_i}
\]
You may assume for simplicity that the second moments of \(z_{i1}, z_{i2}, \epsilon_i, v_i\) are all 1.

1. Derive the asymptotic distribution of
\[
\left[ \frac{\sqrt{n}}{} \left( \tilde{\beta}_1 - \beta \right) \right. \left. \frac{\sqrt{n}}{} \left( \tilde{\beta}_2 - \beta \right) \right]
\]
2. Assume that the above vector converges to $\mathcal{N}(0, \Sigma)$. Consider a class of estimators $t\hat{\beta}_1 + (1 - t)\hat{\beta}_2$ indexed by $t$. Find $t$ that minimizes the asymptotic variance of $\sqrt{n}(t\hat{\beta}_1 + (1 - t)\hat{\beta}_2 - \beta)$. What is the minimized asymptotic variance?

3. Derive the asymptotic distribution of 2SLS for this model. (You do not need to show how to derive the asymptotic distribution of 2SLS for the general case. You are allowed to state the asymptotic variance formula of 2SLS, and then use specific values from this question.) Compare the asymptotic variance of 2SLS with the minimized variance derived in (2) above.

**Question II-3**

Suppose that $Z_1, \ldots, Z_n$ are i.i.d., and their common PDF is given by $f(z; \theta_1, \theta_2)$. Both $\theta_1$ and $\theta_2$ are scalars. Let

$$s_1 = \frac{\partial \log f(Z; \theta_1, \theta_2)}{\partial \theta_1}$$

$$s_2 = \frac{\partial \log f(Z; \theta_1, \theta_2)}{\partial \theta_2}$$

and let

$$\begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix} = E \begin{bmatrix} s_1^2 & s_1 s_2 \\ s_1 s_2 & s_2^2 \end{bmatrix}$$

Assume that $|\rho| < 1$. Here, the $Z$ denotes a random variable with distribution identical to that of $Z_i$.

We consider two estimators of $\theta_1$. The first estimator $\hat{\theta}_1$ is a feasible estimator that solves

$$\max_{c_1, c_2} \sum_{i=1}^{n} \log f(Z_i; c_1, c_2)$$

The second estimator $\tilde{\theta}_1$ is an infeasible estimator that solves

$$\max_{c_1} \sum_{i=1}^{n} \log f(Z_i; c_1, \theta_2)$$

i.e., it is based on the assumption that $\theta_2$ is known. Prove that the asymptotic variance of $\sqrt{n} \left( \tilde{\theta}_1 - \theta_1 \right)$ is smaller than or equal to the asymptotic variance of $\sqrt{n} \left( \hat{\theta}_1 - \theta_1 \right)$. When are they equal to each other?
Question III-1

Let $X_1, ..., X_n$ be a random sample from a location-exponential family with density function

$$f(x; \theta) = \begin{cases} \exp(\theta - x), & \text{if } x \geq \theta \\ 0, & \text{otherwise} \end{cases},$$

where $\theta$ can be any finite constant, i.e. $\theta \in (-\infty, \infty)$. We are interested in testing

$$H_0 : \theta \leq \theta_0 \text{ versus } H_1 : \theta > \theta_0.$$

1. Find the maximum likelihood estimator of $\theta$.
2. Find likelihood ratio test with size $\alpha \in (0, 1)$ (note: you are supposed to find the explicit critical value and hence the explicit critical region.)
3. Find the power function of the likelihood ratio test in question (2).

Question III-2

Suppose that $\{Y_t\}$ is an auto-regressive process, i.e.

$$Y_t = \rho_0 Y_{t-1} + u_t,$$

where $|\rho_0| < 1$ and $u_t \sim i.i.d.(0, \sigma_u^2)$. Suppose that $Y_t$ is only observable at the even period, i.e. when $t$ is an even number, and there exists a sample of observation for $Y_t$, i.e. $\{Y_{2t}\}_{t=1}^n$.

1. Given the sample $\{Y_{2t}\}_{t=1}^n$, we can define an estimator as

$$\hat{\theta}_n = \frac{\sum_{t=2}^n Y_{2t}Y_{2t-2}}{\sum_{t=1}^n Y_{2t}^2}. \quad (3)$$

Is $\hat{\theta}_n$ a consistent estimator for $\rho_0$? Show the consistency of $\hat{\theta}_n$ if your answer is "yes". Otherwise, construct a consistent estimator for $\rho_0$.

2. Find the asymptotic distribution of the consistent estimator you find in question (1).

Question III-3

Suppose that $\{e_t\}$ is generated from the following equation

$$e_t = \gamma_0 + e_{t-1} + u_t,$$
$$u_t = \rho_0 u_{t-1} + v_t,$$

where $e_0 = 0$, $\rho_0$ and $\gamma_0$ are some finite constants and $v_t \sim i.i.d.(0, \sigma_v^2)$. 

5
1. Suppose that $\gamma_o$ and $\rho_o$ are known and you have data $\{e_t\}_{t=0}^n$ where $n \geq 2$. What is the optimal prediction of $e_{n+1}$? What is the related mean square prediction error?

2. Given the data $\{e_t\}_{t=1}^n$, one could estimate $\gamma_o$ and $\rho_o$ by

$$
\hat{\gamma}_n = \frac{1}{n-1} \sum_{t=1}^{n-1} (e_{t+1} - e_t) \quad \text{and} \quad \hat{\rho}_n = \frac{\sum_{t=1}^{n-1} \hat{u}_{t+1} \hat{u}_t}{\sum_{t=1}^{n-1} \hat{u}_t^2}
$$

where $\hat{u}_t = e_t - e_{t-1} - \hat{\gamma}_n$. If $|\rho_o| < 1$, are the estimators $\hat{\gamma}_n$ and $\hat{\rho}_n$ defined in equation (4) consistent? Justify your answer.

3. If $\rho_o \geq 1$, are the estimators $\hat{\gamma}_n$ and $\hat{\rho}_n$ defined in equation (4) still consistent? Justify your answer.

Some Useful Theorems and Lemmas

**Theorem 1 (Martingale Convergence Theorem)** Let $\{(X_t, \mathcal{F}_t)\}_{t \in \mathbb{Z}_+}$ be a martingale in $L^2$. If $\sup_t E[X_t^2] < \infty$, then $X_n \to X_\infty$ almost surely, where $X_\infty$ is some element in $L^2$.

**Theorem 2 (Martingale CLT)** Let $\{X_{t,n}, \mathcal{F}_{t,n}\}$ be a martingale difference array such that $E[|X_{t,n}|^{2+\delta}] < \Delta < \infty$ for some $\delta > 0$ and for all $t$ and $n$. If $\sigma_n^2 > \delta_1 > 0$ for all $n$ sufficiently large and $\frac{1}{n} \sum_{t=1}^n X_{t,n}^2 - \sigma_n^2 \to_p 0$, then $n^{1/4} X_n / \sigma_n \to_d N(0,1)$.

**Theorem 3 (LLN of Sample Variance)** Suppose that $Z_t$ is i.i.d. with mean zero and $E[Z_0^2] = \sigma_Z^2 < \infty$. Let $X_t = \sum_{k=0}^\infty \varphi_k Z_{t-k}$, where $\varphi_k$ is a sequence of real numbers with $\sum_{k=0}^\infty k \varphi_k^2 < \infty$. Then

$$
\frac{1}{n} \sum_{t=1}^n X_t X_{t-h} \to_p \Gamma_X(h) = E[X_t X_{t-h}].
$$

(5)

**Theorem 4 (Donsker)** Let $\{u_t\}$ be a sequence of random variables generated by $u_t = \sum_{k=0}^\infty \varphi_k \varepsilon_{t-k} = \varphi(L) \varepsilon_t$, where $\{\varepsilon_t\} \sim \text{iid} (0, \sigma_\varepsilon^2)$ with finite fourth moment and $\{\varphi_k\}$ is a sequence of constants with $\sum_{k=0}^\infty k |\varphi_k| < \infty$. Then $B_{u,n}(\cdot) = n^{-\frac{1}{2}} \sum_{t=1}^{[n]} u_t \to_d \lambda B(\cdot)$, where $\lambda = \sigma_\varepsilon \varphi(1)$. \]