Recall that the “asymptotic distribution” of a sequence $X_1, X_2, \ldots$ is a sequence $c_n$ of constants, a sequence $k_n$ of constants, and a nondegenerate random variable $Y$ such that $c_n(X_n - k_n) \xrightarrow{L} Y$. (Often, $k_n \equiv k$ and $c_n$ will be $n^b$ for some $b$ such as $b = 1/2$.)

**Problem 1. [6 points]** Let $X_1, \ldots, X_n$ be a simple random sample from $f_\theta(x)$ and suppose that $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ satisfies

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{L} N(0, \tau^2)$$

for a positive, finite constant $\tau^2$.

(a) Prove that if $\theta = 0$, then $P(|\hat{\theta}_n| < n^{-1/3}) \to 1$.

**Solution:** Since $\sqrt{n}\hat{\theta}_n \xrightarrow{L} N(0, \tau^2)$, Slutsky’s theorem implies that

$$(n^{-1/6})\sqrt{n}\hat{\theta}_n = n^{1/3}\hat{\theta}_n \xrightarrow{P} 0.$$  

By definition, this means that

$$P(n^{1/3}|\hat{\theta}_n| < 1) = P(|\hat{\theta}_n| < n^{-1/3}) \to 1.$$  

(b) Assuming that $\theta > 0$, find (with proof) the set of all $\alpha > 0$ such that $P(|\hat{\theta}_n| < n^{-\alpha}) \to 0$.

**Solution:** We know that $\hat{\theta}_n \xrightarrow{P} \theta > 0$, which implies $P(|\hat{\theta}_n| < \theta/2) \to 0$. Take any $\alpha > 0$. Since $n^{-\alpha} \to 0$ for any $\alpha > 0$, we can find $N$ such that $n^{-\alpha} < \theta/2$ for $n > N$. This implies that for $n < N$,

$$P(|\hat{\theta}_n| < n^{-\alpha}) \leq P(|\hat{\theta}_n| < \theta/2).$$

Since the right hand side tends to zero and $\alpha > 0$ was arbitrary, we conclude that $P(|\hat{\theta}_n| < n^{-\alpha}) \to 0$ for any $\alpha > 0$. ■
Problem 2. [3 points] Suppose that for $\theta > 2$, $X_1, X_2, \ldots$ are independent and identically distributed from a Pareto distribution with density function $f_\theta(x) = \frac{\theta}{x^{\theta+1}}$, $x > 1$. This implies that

$$E X_i = \frac{\theta}{\theta - 1} \text{ and } \text{Var } X_i = \frac{\theta}{(\theta - 2)(\theta - 1)^2}.$$ 

Let $Y_i = X_i + X_{i+1}$. Find the asymptotic distribution of $Y_n$.

Solution: Because $Y_1, Y_2, \ldots$, is a stationary 1-dependent sequence, we may apply the central limit theorem for $m$-dependent sequences.

Alternatively, just notice that $Y_n = 2X_n + (X_{n+1} - X_1)/n$, and since $\sqrt{n}(X_{n+1} - X_1)/n \xrightarrow{P} 0$, we can apply Slutsky’s theorem to conclude that $Y_n$ has the same asymptotic distribution as $2X_n$, which is obtained using the central limit theorem:

$$\sqrt{n}(Y_n - 2E X_1) \xrightarrow{L} N(0, 4 \text{Var } X_1).$$

The values of $E X_1$ and $\text{Var } X_1$ are given in the problem.

Problem 3. [3 points] Suppose that for $\theta > 0$, $X_1, \ldots, X_n$ is a simple random sample from an exponential distribution with distribution function $F_\theta(x) = 1 - \exp\{-x/\theta\}$, $x > 0$. This implies that $E X_i = \theta$ and $\text{Var } X_i = \theta^2$. Let $U_n$ denote the U-statistic corresponding to the kernel function $\phi(x, y) = I\{x + y > 1\}$. In other words,

$$U_n = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \phi(X_i, X_j).$$

Find the asymptotic distribution of $U_n$.

Solution: We know that

$$\sqrt{n}(U_n - E U_n) \xrightarrow{L} N(0, 4\sigma_1^2),$$

where $E U_n = E \phi_1(X)$ and $\sigma_1^2 = \text{Var } \phi_1(X)$. We find that

$$\phi_1(x) = P(x + Y > 1) = 1 - F_\theta(1 - x) = 1 - [1 - \exp\{(x - 1)/\theta\}]I\{x < 1\}.$$

Thus,

$$E \phi_1(X) = 1 - \frac{1}{\theta} \int_0^1 (e^{-x/\theta} - e^{-1/\theta}) \, dx = e^{-1/\theta} \left(1 + \frac{1}{\theta}\right)$$

and

$$\text{Var } \phi_1(X) = E \exp\{-2X - 2/\theta\} - [E \phi(X)]^2 = \frac{1}{\theta} \int_0^1 (1 - e^{(x-1)/\theta})^2 e^{-x/\theta} \, dx - \frac{1}{\theta^2} \left[\int_0^1 (e^{-x/\theta} - e^{-1/\theta}) \, dx \right]^2 = 2\gamma - \gamma^2 \left(2 + \frac{2}{\theta} + \frac{1}{\theta^2}\right).$$
where \( \gamma = \exp\{-1/\theta\} \) (sorry, that last calculation is annoyingly complicated for a final exam). We merely plug the computed values of \( E U_n \) and \( \sigma_1^2 \) into (1) to obtain

\[
\sqrt{n} \left[ U_n - \gamma \left( 1 + \frac{1}{\theta} \right) \right] \xrightarrow{L} N \left[ 0, 8 \gamma - 4 \gamma^2 \left( 2 + \frac{2}{\theta} + \frac{1}{\theta^2} \right) \right].
\]

**Problem 4. [8 points]** Suppose that for \( \theta > 2 \), \( X_1, \ldots, X_n \) is a simple random sample from a Pareto distribution with density function \( f_\theta(x) = \theta/x^{\theta+1}, \ x > 1 \). This implies that

\[
E X_i = \frac{\theta}{\theta - 1} \quad \text{and} \quad \text{Var} \ X_i = \frac{\theta}{(\theta - 2)(\theta - 1)^2}.
\]

Let \( a \) be some constant and define

\[
\delta_n = \hat{\theta}_n + \frac{a}{\sum_{i=1}^n \log X_i} = \frac{n + a}{\sum_{i=1}^n \log X_i}.
\]

(a) Find (with proof) the asymptotic distribution of \( \delta_n \). You may assume (without proof) the facts that the MLE is an efficient estimator and that \( \log X_i \) has an exponential distribution with mean \( 1/\theta \) and variance \( 1/\theta^2 \). (Also see part (c).)

**Solution:** First, we find \( I(\theta) \). Since \( (\partial^2 / \partial \theta^2) \log f_\theta(x) = -1/\theta^2 \), we know that \( I(\theta) = 1/\theta^2 \).

Since \( n/\sum_{i=1}^n \log X_i \xrightarrow{P} \theta \) by the weak law of large numbers, we know that \( a\sqrt{n}/\sum_{i=1}^n \log X_i \xrightarrow{P} 0 \).

Therefore, since \( \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{L} N(0, \theta^2) \), we conclude by Slutsky’s theorem that

\[
\sqrt{n}(\delta_n - \theta) = \sqrt{n}(\hat{\theta}_n - \theta) + \frac{a\sqrt{n}}{\sum_{i=1}^n \log X_i} \xrightarrow{L} N(0, \theta^2).
\]

(b) Find \( a \) if \( \delta_n \) is the Bayesian estimator equal to the posterior mean of \( \theta \) under the Jeffreys prior.

**Hint:** The Gamma(\( \alpha, \beta \)) distribution has expectation \( \alpha/\beta \), variance \( \alpha/\beta^2 \), and density function

\[
f(y) = \beta^\alpha y^{\alpha-1}e^{-\beta y}I\{y > 0\}/\Gamma(\alpha).
\]

**Solution:** The Jeffreys prior density on \( \theta \), which is improper in this case, is proportional to \( 1/\theta \).

Ignoring proportionality constants not involving \( \theta \), the likelihood times the prior is therefore

\[
\theta^{n-1} \left( \prod_{i=1}^n X_i \right)^{-\theta} = \theta^{n-1} \exp \left\{ -\theta \sum_{i=1}^n \log X_i \right\}.
\]

Thus, the posterior distribution of \( \theta \) is gamma \( (n, \sum_{i=1}^n \log X_i) \), which means that the posterior mean in this case is the same as the maximum likelihood estimator \( \hat{\theta}_n \). In other words, \( a = 0 \).
(c) The method of moments estimator and maximum likelihood estimator are

\[ M_n = \frac{X_n}{X_n - 1} \quad \text{and} \quad \hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} \log X_i, \]

respectively. Find the joint asymptotic distribution of \((M_n, \hat{\theta}_n)\). You may use the fact that

\[ \int_{1}^{\infty} x^{-\theta} \log x \, dx = \frac{1}{(\theta - 1)^2} \]

in addition to the facts stated in part (a).

**Solution:** Let \( Y_i = \log X_i \). Then

\[ \mathbb{E} X_i Y_i = \theta \int_{1}^{\infty} \frac{x \log x}{x^{\theta+1}} \, dx = \frac{\theta}{(\theta - 1)^2} \]

implies that \( \text{Cov}(X_i, Y_i) = \theta/(\theta - 1)^2 - 1/(\theta - 1) = 1/(\theta - 1)^2 \). Therefore, the multivariate central limit theorem implies

\[ \sqrt{n} \left[ \left( \frac{X_n}{Y_n} \right) - \left( \frac{\theta}{\theta - 1} \right) \right] \xrightarrow{d} N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta \theta^2 \\ \frac{1}{(\theta - 1)^2} \frac{\theta}{(\theta - 1)^2} \end{pmatrix} \right] \]

Now define the vector-valued function \( g(a, b) = [a/(a - 1), 1/b] \). We calculate that

\[ \nabla g \left( \frac{\theta}{\theta - 1}, \frac{1}{\theta} \right) = \begin{pmatrix} -\theta + 1 \\ 0 \end{pmatrix} / -\theta^2 \]

Therefore,

\[ \nabla g \left( \frac{\theta}{\theta - 1}, \frac{1}{\theta} \right) \left( \begin{pmatrix} \theta \theta^2 \\ \frac{1}{(\theta - 1)^2} \frac{\theta}{(\theta - 1)^2} \end{pmatrix} \right) = \begin{pmatrix} \theta \theta^2 \\ \frac{\theta}{\theta - 2} \frac{\theta}{\theta^2} \end{pmatrix} \]

We conclude that

\[ \sqrt{n} \left[ \left( \frac{M_n}{\hat{\theta}_n} \right) - \left( \frac{\theta}{\theta} \right) \right] \xrightarrow{d} N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta \theta^2 \\ \frac{\theta}{\theta - 2} \frac{\theta}{\theta^2} \end{pmatrix} \right] \]