You may use the following facts on this Theory I question set.

- If $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with marginal pdf f and cdf F, then the joint pdf of order statistics $(X_{(i)}, X_{(j)})$ (for given $1 \le i < j \le n$) is
$$f_{X_{(i)},X_{(j)}}(x,y) =$$
$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} f(x)f(y)[F(x)]^{i-1}[F(y)-F(x)]^{j-i-1}[1-F(y)]^{n-j}, \quad x < y,$$
with $f_{X_{(i)},X_{(j)}}(x,y) = 0$ for any other $x, y \in \mathbb{R}$.

- Based on the random sample above, the marginal pdf of the order statistic $X_{(i)}$ (for some $1 \le i \le n$) is
$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} f(x)[F(x)]^{i-1}[1-F(x)]^{n-i}, \quad x \in \mathbb{R}.$$

- A Gamma(α, β) random variable X has pdf $f(x) = x^{\alpha-1}e^{-x/\beta}/[\beta^\alpha \Gamma(\alpha)]$, $x > 0$ and moment generating function $M(t) = [1-\beta t]^{-\alpha}$ for $t < 1/\beta$, where $\alpha, \beta > 0$.

- A Gamma(α, β) random variable X has mean $\mathrm{E}(X) = \alpha\beta$ and variance $\mathrm{Var}(X) = \alpha\beta^2$.

- A chi-square random variable with ν degrees of freedom is Gamma$(\nu/2, 2)$-distributed.

Part I

Let $X_1, X_2, \ldots, X_n$ $(n > 2)$ be i.i.d. random variables with marginal pdf f, marginal cdf F, and support $\mathcal{X} = \{x \in \mathbb{R} : f(x) > 0\}$.
Let $Y_i = X_{(i)}$ denote the ith order statistic, $i = 1, \ldots, n$, whereby $Y_1 < Y_2 < \cdots < Y_n$.

1. Show that the conditional pdf of Y_n given $Y_{n-1} = y_{n-1} \in \mathcal{X}$ is given by
$$f_{Y_n|y_{n-1}}(y) = \frac{f(y)}{1 - F(y_{n-1})}, \quad y > y_{n-1}$$
with $f_{Y_n|y_{n-1}}(y) = 0$ for $y \le y_{n-1}$.

2. For fixed $y_1 < y_2 < \ldots < y_{n-1} \in \mathcal{X}$ and any $y \in \mathbb{R}$, show that
$$P(Y_n \le y | Y_{n-1} = y_{n-1}) = P(Y_n \le y | Y_1 = y_1, Y_2 = y_2, \cdots, Y_{n-1} = y_{n-1})$$
(i.e., the conditional cdf of Y_n given $Y_1, \ldots, Y_{n-1}$ matches the conditional cdf of Y_n given Y_{n-1}).

3. If $X_1 \sim$Uniform$(0,1)$, find $\mathrm{E}[\mathrm{E}(Y_n Y_{n-1} | Y_{n-1})]$.

4. Let $F_i(y; y_{i-1})$, $y \in \mathbb{R}$, denote the conditional cdf of Y_i given $Y_{i-1} = y_{i-1} \in \mathcal{X}$ for $i = 2, \ldots, n$, and let $F_1(y)$, $y \in \mathbb{R}$, denote the cdf of Y_1. Define random variables $R_1 = F_1(Y_1)$ and $R_i = F_i(Y_i; Y_{i-1})$ for $i = 2, \ldots n$. Determine $P(R_1 \le r_1, R_2 \le r_2, \ldots, R_n \le r_n)$ as a function of $r_1, r_2, \ldots, r_n \in \mathbb{R}$.

Part II

Let $X_1, \ldots, X_n$ be i.i.d. standard exponential random variables with pdf $f(x) = e^{-x}$, $x > 0$. The joint pdf of the order statistics $(X_{(1)}, \ldots, X_{(n)})$ is

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1, \ldots, x_n) = n! \exp\left(-\sum_{i=1}^{n} x_i\right), \quad 0 < x_1 < \ldots < x_n.$$

Define the spacing random variables as $D_1 = X_{(1)}$ and $D_i = X_{(i)} - X_{(i-1)}$ for $i = 2, \ldots, n$.

5. Show that $D_1, \ldots, D_n$ are *independent* random variables, where D_i is Exponential with pdf $f_{D_i}(d_i) = (n-i+1)e^{-(n-i+1)d_i}$, $d_i > 0$, for $i = 1, \ldots, n$. Justify your answer.

Note: $X_{(i)} = \sum_{k=1}^{i} D_k$ for each $i = 1, \ldots, n$.

6. If $M(t)$ denotes the moment generating function of $\sum_{i=1}^{n} D_i$, determine the second derivative $[d^2 \log M(t)/dt^2]|_{t=0}$ as a function of $\sum_{i=1}^{n} i^{-2}$.

7. Show that $D_1 + D_n \xrightarrow{d} W$ as $n \to \infty$, for some random variable W. Identify the distribution of W.

8. For a fixed $r \in \{1, \ldots, n\}$, let

$$T_r \equiv \sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)} = \sum_{i=1}^{r}(n-i+1)D_i.$$

Find the moment generating function of $2T_r$.

9. Show that the distribution of $2T_r$ is the same as that of $\sum_{i=1}^{2r} Z_i^2$, where $Z_1, \ldots, Z_{2r}$ are i.i.d. standard normal random variables. State any standard results that you use.

10. Show that $\left(\sum_{i=1}^{n} X_{(i)}/n\right)^2 \xrightarrow{d} 1$ as $n \to \infty$. Justify your answer.

11. Determine the distribution of $\sqrt{r}(\log T_r - \log r)$ when $r \to \infty$ as $n \to \infty$.

Hint: Consider $\log(T_r/r) = \log T_r - \log r$ for the quantity $2T_r/(2r)$ along with the implication of Question 9.

12. For $Z_i = X_{(i)} - X_{(1)}$, $i > 1$, show that $(Z_2, \ldots, Z_n)$ has the same joint distribution as the order statistics defined by a size $n-1$ random sample of standard exponential random variables.

13. Using Question 12 and letting $\bar{X}_n = \sum_{i=1}^{n} X_i/n$, show that the distribution of $2(\bar{X}_n - X_{(1)})$ is χ^2_{2n-2}. State any standard results that you use.

Part I

1. From the facts given, the joint pdf of $(Y_{n-1}, Y_n) = (X_{(n-1)}, X_{(n)})$ is given by

$$f_{Y_{n-1},Y_n}(x, y) = n(n-1)f(x)f(y)[F(x)]^{n-2}, x < y$$

and the marginal pdf of Y_{n-1} is

$$f_{Y_{n-1}}(x) = n(n-1)f(x)[F(x)]^{n-2}[1 - F(x)].$$

Given $Y_{n-1} = x$ (where $f(x) > 0$), the conditional pdf of Y_n is then

$$f_{Y_n|x}(y) = \frac{f_{Y_n,Y_{n-1}}(x,y)}{f_{Y_{n-1}}(x)} = \frac{f(y)}{1-F(x)}, \quad y > x.$$

Set $x = y_{n-1}$ above for conditioning on $Y_{n-1} = y_{n-1}$.

2. To show that the conditional distributions are the same given $Y_{n-1} = y_{n-1}$, it suffices to show the conditional pdfs match. We already have the conditional pdf of Y_n given $Y_{n-1} = y_{n-1}$. To find the conditional pdf of Y_n given $Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1}$ (where each $f(y_i) > 0$), note that the joint pdf of $Y_1, \ldots, Y_n$ is

$$g(y_1, \cdots, y_n) = n!f(y_1)\cdots f(y_n), \quad y_1 < \ldots < y_n$$

and the joint pdf of $Y_1, \ldots, Y_{n-1}$ is (by integrating Y_n out)

$$h(y_1, \ldots, y_{n-1}) = n!f(y_1)\cdots f(y_{n-1}) \int_{y_{n-1}}^{\infty} f(y_n)dy_n = n!f(y_1)\cdots f(y_{n-1})[1 - F(y_{n-1})]$$

for $y_1 < \ldots < y_{n-1}$. Hence, the conditional pdf of Y_n given $Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1}$ is

$$f_{Y_n|y_1,\ldots,y_{n-1}}(y_n) = \frac{g(y_1,\cdots,y_n)}{h(y_1,\ldots,y_{n-1})} = \frac{f(y_n)}{1-F(y_{n-1})}, \quad y_n > y_{n-1}$$

which is the same conditional pdf $f_{Y_n|y_{n-1}}(y_n)$.

3. $\mathrm{E}[\mathrm{E}(Y_nY_{n-1}|Y_{n-1})] = \mathrm{E}[Y_nY_{n-1}] = \mathrm{E}[X_{(n)}X_{(n-1)}]$, where, if $f(x) = 1$, $0 < x < 1$,

$$\mathrm{E}[X_{(n)}X_{(n-1)}] = \int_0^1 \int_0^y n(n-1)xy[x]^{n-2}dxdy = \int_0^1 (n-1)y^{n+1}dy = \frac{n-1}{n+2}.$$

Or, $\mathrm{E}[\mathrm{E}(Y_nY_{n-1}|Y_{n-1})] = \mathrm{E}[Y_{n-1}\mathrm{E}(Y_n|Y_{n-1})]$ where $Y_n|Y_{n-1} = y_{n-1} \sim$Uniform$(y_{n-1}, 1)$ by $f_{Y_n|Y_{n-1}=y_{n-1}}(y_n) = 1/(1-y_{n-1})$, $y_{n-1} < y_n < 1$ so that

$$\mathrm{E}(Y_n|Y_{n-1} = y_{n-1}) = \int_{y_{n-1}}^1 y_n \frac{1}{1-y_{n-1}} dy_n = \frac{1}{2}(1 + y_{n-1})$$

and, by $f_{Y_{n-1}}(y) = n(n-1)y^{n-2}(1-y)$, $0 < y < 1$,

$$\frac{1}{2}\mathrm{E}[Y_{n-1}(1+Y_{n-1})] = \frac{1}{2}\int_0^1 n(n-1)y(1+y)y^{n-2}(1-y)dy = \frac{1}{2}n(n-1)\left(\frac{1}{n} - \frac{1}{n+2}\right) = \frac{n-1}{n+2}.$$

4. In analogy to question 2, the conditional cdf of Y_i given Y_{i-1} is the same as the conditional cdf of Y_i given $Y_{i-1}, \ldots, Y_1$ for each $i = 2, \ldots, n$. By the probability integral transform, $F_i(Y_i; y_{i-1})$ is uniform$(0, 1)$ distributed for each $i = 2, \ldots, n$ (regardless of the value of y_{i-1}) and so is $F_1(Y_1)$; that is,

$$P(F_i(Y_i; y_{i-1}) \leq r_i | Y_1 = y_1, \ldots, Y_{i-1} = y_{i-1}) = r_i \mathbb{I}(r_i \in [0, 1])$$

for each $i = 2, \ldots, n$ and $r_i \in \mathbb{R}$, where $\mathbb{I}(\cdot)$ denotes the indicator function. Hence,

$$\begin{aligned}
& P(R_1 \leq r_1, \ldots, R_n \leq r_n) \\
= \; & \mathrm{E}[\mathbb{I}(R_1 \leq r_1, \ldots, R_{n-1} \leq r_{n-1}) P(F_{n-1}(Y_n; y_{i-1}) \leq r_n | Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1})] \\
= \; & \mathrm{E}[\mathbb{I}(R_1 \leq r_1, \ldots, R_{n-1} \leq r_{n-1}) r_n \mathbb{I}(r_n \in [0, 1])] \\
= \; & r_n \mathbb{I}(r_n \in [0, 1]) P(R_1 \leq r_1, \ldots, R_{n-1} \leq r_{n-1}),
\end{aligned}$$

$$\begin{aligned}
& P(R_1 \leq r_1, \ldots, R_{n-1} \leq r_{n-1}) \\
= \; & \mathrm{E}[\mathbb{I}(R_1 \leq r_1, \ldots, R_{n-2} \leq r_{n-2}) P(F_{n-1}(Y_{n-1}; y_{i-2}) \leq r_{n-1} | Y_1 = y_1, \ldots, Y_{n-2} = y_{n-2})] \\
= \; & \mathrm{E}[\mathbb{I}(R_1 \leq r_1, \ldots, R_{n-2} \leq r_{n-2}) r_{n-1} \mathbb{I}(r_{n-1} \in [0, 1])] \\
= \; & r_{n-1} \mathbb{I}(r_{n-1} \in [0, 1]) P(R_1 \leq r_1, \ldots, R_{n-2} \leq r_{n-2}),
\end{aligned}$$

$$\vdots$$

$$\begin{aligned}
& P(R_1 \leq r_1) \\
= \; & P(F_1(Y_1) \leq r_1) = r_1 \mathbb{I}(r_1 \in [0, 1])
\end{aligned}$$

That is, $R_1, \ldots, R_n$ are iid uniform$(0, 1)$ variables and

$$P(R_1 \leq r_1, \ldots, R_n \leq r_n) = \prod_{i=1}^{n} r_i \mathbb{I}(r_i \in [0, 1]), \quad r_1, \ldots, r_n \in \mathbb{R}.$$

Part II

5. As $X_{(i)} = \sum_{k=1}^{i} D_k$ for $i = 1, \ldots, n$, the transformation is one-to-one and the Jacobian is given by

$$J = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

(i.e., 1's on the diagonal and lower-diagonal with 0's elsewhere); note $\det(J) = 1$. The range of each D_i is $(0, \infty)$ and so the joint pdf of $(D_1, \ldots, D_n)$ is

$$\begin{aligned} f_{D_1,\ldots,D_n}(d_1, \ldots, d_n) &= n! \exp\left(-\sum_{i=1}^{n}\sum_{k=1}^{i} d_k\right) |\det(J)|, \quad d_1, \ldots, d_n > 0 \\ &= n! \exp\left(-\sum_{k=1}^{n}\sum_{i=k}^{n} d_k\right), \quad d_1, \ldots, d_n > 0 \\ &= n! \exp\left(-\sum_{k=1}^{n}(n-k+1) d_k\right), \quad d_1, \ldots, d_n > 0 \\ &= \prod_{k=1}^{n}(n-k+1) \exp\left(-\sum_{k=1}^{n}(n-k+1) d_k\right), \quad d_1, \ldots, d_n > 0 \end{aligned}$$

using $\sum_{i=k}^{n} 1 = n-k+1$ for each $k = 1, \ldots, n$ and $n! = n(n-1)\cdots 1 = \prod_{k=1}^{n}(n-k+1)$. The joint pdf factors into a product of marginal pdfs $f_{D_i}(d_i) = (n-i+1)\exp(-(n-i+1)d_i)$, $d_i > 0$, $i = 1, \ldots, n$ so that $D_1, \ldots, D_n$ are independent and D_i is exponential with mean $(n-i+1)^{-1}$.

6. Let $D = \sum_{i=1}^{n} D_i$ so that $M_D(t) = M(t)$ by definition. Using $M_D(0) = 1$, $[dM_D(t)/dt]|_{t=0} = \mathrm{E}(D)$, $[d^2 M_D(t)/dt^2]|_{t=0} = \mathrm{E}(D^2)$ by definition of the moment generating function,

$$\begin{aligned} \frac{d^2 \log M_D(t)}{dt^2}\Big|_{t=0} &= \frac{d\{d[M_D(t)]/dt \cdot [M_D(t)]^{-1}\}}{dt}\Big|_{t=0} \\ &= d^2[M_D(t)]/dt^2 \cdot [M_D(t)]^{-1}|_{t=0} - \{d[M_D(t)]/dt\}^2 \cdot [M_D(t)]^{-2}|_{t=0} \\ &= \mathrm{E}(D^2) - [\mathrm{E}(D)]^2 = \mathrm{Var}(D) \\ &= \sum_{i=1}^{n} \mathrm{Var}(D_i) = \sum_{i=1}^{n} \frac{1}{(n-i+1)^2} = \sum_{i=1}^{n} \frac{1}{i^2} \end{aligned}$$

7. Note for any n, D_n $\sim$Exponential(1) so that $D_n \xrightarrow{d} W$ trivially as $n \to \infty$, where W $\sim$Exponential(1). Also, D_1 is Exponential with mean $1/n$ so that $D_1 \xrightarrow{p} 0$ as $n \to \infty$;

i.e., pick $\epsilon > 0$, then

$$P(|D_1| > \epsilon) \leq \frac{\mathrm{E}|D_1|}{\epsilon} = \frac{\mathrm{E}(D_1)}{\epsilon} = \frac{1}{n\epsilon} \to 0.$$

Hence, by Slutsky's theorem $D_1 + D_n \xrightarrow{d} 0 + W = W$.

8. As $T_r = \sum_{i=1}^{r}(n-i+1)D_i$ with $D_1, \ldots, D_r$ as independent and the moment generating function of Exponential D_i is $M_{D_i}(t) = \mathrm{E}e^{tD_i} = [1-(n-i+1)^{-1}t]$, $t < n-i+1$, then

$$M_{2T_r}(t) = \mathrm{E}e^{t2T_r} = \mathrm{E}e^{2t\sum_{i=1}^{r}(n-i+1)D_i} = \mathrm{E}\prod_{i=1}^{r} e^{t2(n-i+1)D_i} = \prod_{i=1}^{r} M_{D_i}[2t(n-i+1)] = [1-2t]^{-r}$$

which is valid for $2(n-i+1)t < (n-i+1)$, $i = 1, \ldots, n$ or $t < 1/2$.

9. By its moment generating function, the distribution of T_r is Gamma$(r, 2)$ or chi-square with $2r$ degrees of freedom. If $Z_1, \ldots, Z_{2r}$ are iid standard normal, then $Z_1^2, \ldots, Z_n^2$ are iid chi-square with 1 degree of freedom, and $\sum_{i=1}^{2r} Z_i^2$ is also chi-square with $2r$ degrees of freedom.

10. Note $\sum_{i=1}^{n} X_{(i)}/n = \sum_{i=1}^{n} X_i/n = \bar{X}_n$. By the WLLN, $\bar{X}_n \xrightarrow{p} \mathrm{E}X_1 = 1$. By the continuous mapping theorem (i.e., $g(x) = x^2$ is continuous), $(\bar{X}_n)^2 \xrightarrow{p} 1^2 = 1$. Convergence in distribution and in probability are equivalent when the limit is a constant, so $(\bar{X}_n)^2 \xrightarrow{d} 1$.

11. We have $T_r/r = 2T_r/(2r) \stackrel{d}{=} \bar{Y}_{2r} \equiv \sum_{i=1}^{2r} Y_i/(2r)$, where $Y_1, \ldots, Y_r$ are iid chi-square random variables (1 degree of freedom). By the CLT,

$$\sqrt{2r}\left(\bar{Y}_{2r} - \mathrm{E}Y_1\right) = \sqrt{2r}\left(\bar{Y}_{2r} - 1\right) \xrightarrow{d} N(0, \mathrm{Var}(Y_1) = 2)$$

as $r \to \infty$ when $n \to \infty$. By the Delta method and $g(x) = \log x$ (note $g'(x) = 1/x$)

$$\sqrt{2r}\left(g(\bar{Y}_{2r}) - g(1)\right) \xrightarrow{d} N(0, [g'(1)]^2 2 = 2).$$

Finally, $g(\bar{Y}_{2r}) - g(1) = \log(T_r/r) - \log 1 = \log T_r - \log r$, so that

$$\sqrt{2r}\,(\log T_r - \log r) \xrightarrow{d} N(0, 2)$$

and

$$\sqrt{r}\,(\log T_r - \log r) \xrightarrow{d} \frac{1}{\sqrt{2}} \cdot N(0, 2) \stackrel{d}{=} N(0, 1).$$

12. The joint pdf of $(Z_1, \ldots, Z_n)$ with $Z_1 = X_{(1)}$ is

$$\begin{aligned} f_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) &= n! \exp\left(-z_1 - \sum_{i=2}^{n}(z_i + z_1)\right), \quad z_1 > 0, z_n > \cdots > z_2 > 0 \\ &= (n-1)! \exp\left(-\sum_{i=2}^{n} z_i\right) \cdot n \exp\left(-nz_1\right), \quad z_1 > 0, z_n > \cdots > z_2 > 0. \end{aligned}$$

Integrating out $Z_1 = X_{(1)}$ (having an Exponential distribution with mean n^{-1}) gives the joint pdf of $(Z_2, \ldots, Z_n)$ as

$$f_{Z_2,\ldots,Z_n}(z_2, \ldots, z_n) = (n-1)! \exp\left(-\sum_{i=2}^{n} z_i\right), \quad z_n > \cdots > z_2 > 0.$$

13. We have $\bar{X}_n - X_{(1)} = \sum_{i=2}^n Z_i \stackrel{d}{=} \sum_{i=2}^n Z_i^*$ where $Z_2^*, \ldots, Z_n^*$ are iid standard exponential. Hence, $2\sum_{i=2}^n Z_i^*$ is Gamma$(n-1, 2)$ distributed or $\chi^2_{2(n-1)}$ distributed.

Part I

Suppose that $X_1, X_2, \ldots, X_n$ are iid with common pdf

$$f(x|\theta) = \begin{cases} \frac{1}{\sqrt{2\pi\theta}} \frac{1}{x} \exp\Big\{ -\frac{1}{2\theta}(\log x - \theta)^2 \Big\} & \text{if } x > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$.

1. Find a one-dimensional sufficient statistic for θ based on $X_1, \ldots, X_n$.

2. Show that the two-dimensional statistic $(\sum_{i=1}^n \log X_i, \sum_{i=1}^n (\log X_i)^2)$ is *not* complete.

3. Argue that there is a unique maximizer of the likelihood function, call it $\hat{\theta}_n$. Find $\hat{\theta}_n$.

4. Find the asymptotic normal distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$.

5. Show that from the result of Question 4, it follows that $\hat{\theta}_n$ is a consistent estimator of θ.

Part II

Assume that $X_1, X_2, \ldots, X_n$ are iid with common pdf

$$f(x|\mu, \sigma) = \begin{cases} \frac{1}{\sigma} \exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right] & \text{if } x \geq \mu, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$. Assume that both μ and σ are unknown.

6. Find the MLE of (μ, σ).

7. Let $\sigma_0 > 0$ be a fixed number. Show that the likelihood ratio test (LRT) statistic for testing $H_0 : \sigma = \sigma_0$ against $H_1 : \sigma \neq \sigma_0$ can be expressed in terms of $Y \equiv n(\bar{X} - X_{(1)})/\sigma_0$.

8. For a given $\alpha \in (0, 1)$, and fixed $\sigma_0 > 0$, find the size α LRT for testing $H_0 : \sigma = \sigma_0$ against $H_1 : \sigma \neq \sigma_0$.

Hint: You can use the following result without proof to determine the cutoff points. Let $Z_{(1)} < Z_{(2)} < \cdots < Z_{(n)}$ be the order statistics based on a random sample $Z_1, Z_2, \ldots, Z_n$ from Exponential (1). Then $2n(\bar{Z} - Z_{(1)}) \sim \chi^2_{2n-2}$, where $\bar{Z} \equiv \sum_{i=1}^n Z_i/n$.

9. Show that the LRT statistic for testing $H_0 : \mu = \sigma$ against $H_1 : \mu \neq \sigma$ can be expressed in terms of $W \equiv (\bar{X} - X_{(1)})/X_{(1)}$.

10. Show that under $H_0 : \mu = \sigma$, the distribution of the LRT statistic in Question 9 does not depend on the parameters.

For Questions 11-12, assume that $\mu = \sigma = \theta$, where θ is unknown.

11. Show that $X_{(1)}/\theta$ is a pivotal quantity.

12. For a given $\alpha \in (0,1)$, use the pivotal quantity in Question 11 to construct a $(1-\alpha)$ confidence interval for θ.

For questions in **Part III**, you may assume the following without proof.
The beta function is defined as

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx,$$

for $\alpha > 0, \beta > 0$. Also, $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$.

Part III

Suppose that X has the pmf

$$f(x|\theta) = \begin{cases} (1-\theta)^{x-1}\theta & \text{if } x = 1, 2, \ldots, \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta \in (0,1)$. Consider the estimation of θ with the loss function $L(\theta, a) = (\theta - a)^2/\theta$.

13. Let $\pi(\theta)$ be the prior density of θ on $(0,1)$. Show that the Bayes estimator of θ based on X is given by

$$\delta(x) = 1 - \frac{\int_0^1 (1-\theta)^x \pi(\theta)d\theta}{\int_0^1 (1-\theta)^{x-1}\pi(\theta)d\theta}.$$

For Questions 14-15, assume that

$$\pi(\theta) = \frac{\Gamma(2\alpha)}{(\Gamma(\alpha))^2}\theta^{\alpha-1}(1-\theta)^{\alpha-1}, \;\; \theta \in (0,1),$$

is the prior density of θ where $\alpha > 0$.

14. Show that the Bayes estimator of θ based on X is

$$\delta(x) = \frac{\alpha}{x + 2\alpha - 1}, \; x = 1, 2, \cdots .$$

15. Let δ_0 be an estimator, where

$$\delta_0(1) = \frac{1}{2}, \text{ and } \delta_0(x) = 0 \text{ for all } x > 1.$$

Show that δ_0 is a limit of the Bayes estimators defined in Question 14.

Part IV

Let X be an unbiased estimator of θ, and let T be a sufficient statistic for θ. Define $\phi(T) \equiv E(X|T)$.

16. Prove that $\phi(T)$ is an unbiased estimator of θ.

17. Prove that $\text{Var}_\theta(\phi(T)) \leq \text{Var}_\theta(X)$ for all θ.

18. Let $L(\theta, d)$ be a convex function of d for each θ. Show that $E_\theta(L(\theta, \phi(T))) \leq E_\theta(L(\theta, X))$ for all θ.

19. Show that if W is a UMVUE of θ under squared error loss, then W is unique.

1. Since

$$\prod_{i=1}^{n} f(x_i|\theta) = \tfrac{1}{(2\pi)^{n/2}[\prod_{i=1}^{n} x_i]}\tfrac{1}{\theta^{n/2}} \exp\Big\{ -\tfrac{1}{2\theta}[\textstyle\sum_{i=1}^{n}(\log x_i)^2 + n\theta^2 - 2\theta \sum_{i=1}^{n} \log x_i]\Big\}$$
$$= \tfrac{1}{(2\pi)^{n/2}}\tfrac{1}{\theta^{n/2}} \exp\Big\{ -\tfrac{1}{2\theta}[\textstyle\sum_{i=1}^{n}(\log x_i)^2]\Big\} \exp(-n\theta/2),$$

from the factorization theorem, it follow that $\sum_{i=1}^{n}(\log X_i)^2$ is sufficient for θ.

2. Let $Y = \log X$. Note that $Y \sim N(\theta, \theta)$, and $(\sum_{i=1}^{n} \log X_i, \sum_{i=1}^{n}(\log X_i)^2) \equiv (\sum_{i=1}^{n} Y_i, \sum_{i=1}^{n} Y_i^2)$. Let $\bar{Y} = \sum_{i=1}^{n} Y_i/n$. Since $E(g(\sum_{i=1}^{n} Y_i, \sum_{i=1}^{n} Y_i^2)) = 0$, where

$$g(\sum_{i=1}^{n} Y_i, \sum_{i=1}^{n} Y_i^2) = \bar{Y} - \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1},$$

it implies that$(\sum_{i=1}^{n} \log X_i, \sum_{i=1}^{n}(\log X_i)^2)$ is not complete.

3. The log likelihood function is given by

$$l(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\theta) - \frac{1}{2\theta}\Big(\sum_{i=1}^{n}(\log X_i)^2\Big) - \frac{n\theta}{2}.$$

Equating $l'(\theta) = 0$, we have

$$\sum_{i=1}^{n}(\log X_i)^2 - n\theta^2 - n\theta = 0.$$

Since $\theta > 0$, the unique MLE of θ is given by

$$\hat{\theta}_n = \frac{\sqrt{1 + 4\sum_{i=1}^{n}(\log X_i)^2/n} - 1}{2}.$$

4. Since

$$-E(l''(\theta)) = -E\Big(\frac{n}{2\theta^2} - \frac{\sum_{i=1}^{n}(\log X_i)^2}{\theta^3}\Big) = \frac{n(2\theta+1)}{2\theta^2},$$

$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, 2\theta^2/(2\theta+1))$.

5. Since

$$(\hat{\theta}_n - \theta) = \frac{1}{\sqrt{n}}\sqrt{n}(\hat{\theta}_n - \theta),$$

and $\frac{1}{\sqrt{n}} \to 0$, by Slutsky's theorem $(\hat{\theta}_n - \theta) \xrightarrow{d} 0$, which is equivalent to $(\hat{\theta}_n - \theta) \xrightarrow{P} 0$.

6. The likelihood function is

$$\ell(\mu,\sigma|x)=\sigma^{-n}\exp\Big[-\frac{n\bar{x}}{\sigma}+\frac{n\mu}{\sigma}\Big]I_{(-\infty,x_{(1)}]}(\mu),$$

where $\bar{x}$ is the sample mean, $x_{(1)}$ is the smallest order statistic, and $I_A(\cdot)$ is the indicator function of A. For any fixed σ, the likelihood is an increasing function of μ and hence $\hat{\mu}=X_{(1)}$ is the MLE of μ.

Solving the following equation

$$\frac{d}{d\sigma}\log\ell(x_{(1)},\sigma|x)=0,$$

we get $\sigma=\bar{x}-x_{(1)}$. The estimator $\hat{\sigma}=\bar{X}-X_{(1)}$ is the MLE of σ since

$$\frac{d^2}{d\sigma^2}\log\ell(x_{(1)},\sigma|x)\Big|_{\sigma=\hat{\sigma}}<0.$$

7. Under H_0, $\hat{\mu}_0=X_{(1)}$. The likelihood ratio is

$$\lambda(x)=\frac{\frac{1}{\sigma_0^n}\exp(-\sum_{i=1}^n(x_i-x_{(1)})/\sigma_0)}{\frac{1}{\hat{\sigma}^n}\exp(-n)}=e^n\Big(\frac{\hat{\sigma}}{\sigma_0}\Big)^n e^{-n\frac{\hat{\sigma}}{\sigma_0}}=e^n n^{-n}y^n e^{-y}.$$

8. Then $\lambda(x)<c$ is equivalent to $y<c_1$ or $y>c_2$, where $c_1<n<c_2$ with $\lambda(c_1)=\lambda(c_2)$ or $c_1^n e^{-n}=c_2^n e^{-n}$. Under H_0,

$$Y=\frac{n(\bar{X}-X_{(1)})}{\sigma_0}=n\Big(\frac{\bar{X}-\mu}{\sigma_0}-\frac{X_{(1)}-\mu}{\sigma_0}\Big)\overset{d}{=}n(\bar{Z}-Z_{(1)}),$$

where Z_i's are iid Exp(1). From the hint, it follows that $2Y\sim\chi^2_{2n-2}$. Let $\chi^2_{r,\alpha}$ be the αth quantile of χ^2 distribution with r degrees of freedom. Hence, an LRT of size α rejects H_0 when $2Y<\chi^2_{2n-2,\alpha_1}$ or $2Y>\chi^2_{2n-2,1-\alpha_2}$ where α_1,α_2 such that $1-\alpha_1-\alpha_2=1-\alpha$ and $a_1^n e^{-a_1}=a_2^n e^{-a_2}$ for $a_1=\chi^2_{2n-2,\alpha_1}/2, a_2=\chi^2_{2n-2,1-\alpha_2}/2$.

9. Under H_0, $\hat{\mu}_0=X_{(1)}=\hat{\sigma}_0$. The likelihood ratio is

$$\lambda(x)=\frac{\frac{1}{x_{(1)}^n}\exp(-\sum_{i=1}^n(x_i-x_{(1)})/x_{(1)})}{\frac{1}{\hat{\sigma}^n}\exp(-n)}=e^n\Big(\frac{\hat{\sigma}}{x_{(1)}}\Big)^n e^{-n\frac{\hat{\sigma}}{x_{(1)}}}=e^n w^n e^{-nw}.$$

10. Under $H_0:\mu=\sigma$,

$$W=\frac{\bar{X}-X_{(1)}}{X_{(1)}}=\frac{\frac{\bar{X}-\sigma}{\sigma}-\frac{X_{(1)}-\sigma}{\sigma}}{\frac{X_{(1)}-\sigma}{\sigma}+1}\overset{d}{=}\frac{\bar{Z}-Z_{(1)}}{Z_{(1)}+1},$$

where Z_i's are iid Exp(1).

11. Let $Z_i = X_i/\theta - 1$. Since Z_i's are iid Exp(1), $Z_{(1)} \equiv X_{(1)}/\theta - 1 \sim$ Exp $(1/n)$.

12. Since $P(X_{(1)}/\theta - 1 \leq -1/n \log(1-\alpha)) = \alpha$,

$$\left[\frac{X_{(1)}}{1 - n^{-1}\log(\alpha/2)}, \frac{X_{(1)}}{1 - n^{-1}\log(1-\alpha/2)}\right]$$

is a $(1-\alpha)$ confidence interval for θ.

13. The probability mass function of X is $(1-\theta)^{x-1}\theta$ for $x = 1, 2, \ldots$. Hence, for given $X = x$, the Bayes estimator of θ with respect to the loss function $(\theta - a)^2/\theta$ is

$$\delta(x) = \frac{\int_0^1 \theta^{-1}\theta(1-\theta)^{x-1}\theta\pi(\theta)d\theta}{\int_0^1 \theta^{-1}(1-\theta)^{x-1}\theta\pi(\theta)d\theta} = 1 - \frac{\int_0^1 (1-\theta)^{x}\pi(\theta)d\theta}{\int_0^1 (1-\theta)^{x-1}\pi(\theta)d\theta}.$$

14. The Bayes estimator is

$$\delta(x) = 1 - \frac{\int_0^1 (1-\theta)^{x+\alpha-1}\theta^{\alpha-1}d\theta}{\int_0^1 (1-\theta)^{x+\alpha-2}\theta^{\alpha-1}d\theta} = 1 - \frac{x+\alpha-1}{x+2\alpha-1} = \frac{\alpha}{x+2\alpha-1}.$$

15. As $\alpha \to 0$, $\delta(x) \to 1/2$ if $x = 1$ and $\delta(x) \to 0$ if $x > 1$.

16. Since T is sufficient, $\phi(T)$ is a function of the sample only. Then $E_\theta(\phi(T)) = E_\theta(E(X|T)) = E_\theta(X) = \theta$.

17. We have $\text{Var}_\theta(X) = \text{Var}_\theta(E(X|T)) + E(\text{Var}(X|T)) = \text{Var}_\theta(\phi(T)) + E(\text{Var}(X|T)) \geq \text{Var}_\theta(\phi(T))$.

18. Using Jensen's inequality, we have $E_\theta(L(\theta, X)) = E_\theta(E(L(\theta, X)|T)) \geq E_\theta(L(\theta, E(X|T))) = E_\theta(L(\theta, \phi(T)))$.

19. If possible, let W' be another UMVUE of θ. Let $\tilde{W} = (W + W')/2$. Note that $E_\theta(\tilde{W}) = \theta$, and

$$\begin{aligned}
\text{Var}_\theta(\tilde{W}) &= \frac{1}{4}\text{Var}_\theta(W) + \frac{1}{4}\text{Var}_\theta(W') + \frac{1}{2}\text{Cov}_\theta(W, W') \\
&\leq \frac{1}{4}\text{Var}_\theta(W) + \frac{1}{4}\text{Var}_\theta(W') + \frac{1}{2}\sqrt{\text{Var}_\theta(W), \text{Var}_\theta(W')} \\
&= \text{Var}_\theta(W),
\end{aligned}$$

where the inequality follows from the Cauchy-Schwartz inequality. The above inequality cannot be strict as W is UMVUE. Hence $W' = a(\theta)W + b(\theta)$ implying $\text{Cov}_\theta(W, W') = a(\theta)\text{Var}_\theta(W)$. So $a(\theta) = 1$, but then it follows that $\beta(\theta) = 0$ as W' is an unbiased estimator of θ.

1. Let $\Omega \neq \emptyset$ and $\mathcal{F} \neq \emptyset$ be a collection of subsets of Ω. Then, define the following:

(a) Semi-algebra $\mathcal{F}$,

(b) algebra $\mathcal{F}$,

(c) σ-algebra $\mathcal{F}$,

(d) measurable space,

(e) measure μ,

(f) finite measure μ and probability measure μ,

(g) σ-finite measure μ,

(h) measure space and probability space,

(i) measurable transformation T between two measurable spaces,

(j) random variable.

2. Let Ω be an uncountable set and let $\mathcal{F}$ be the countable-cocountable collection, i.e. $\mathcal{F} \equiv \{A \subset \Omega : A \text{ is countable or } A^c \text{ is countable}\}$. Define a real-valued function μ on $\mathcal{F}$ as: $\mu(A) = 1$ if A is uncountable and $\mu(A) = 0$ if A countable, $A \in \mathcal{F}$.

(a) Show $\mathcal{F}$ is a σ-algebra of subsets of Ω.

(b) Show μ is a probability measure on $(\Omega, \mathcal{F})$.

Let X be a random variable on the probability space $(\Omega, \mathcal{F}, \mu)$.

(c) Show that there exists a real constant a and a countable $S \subset \Omega$ such that $X(\omega) = a$ if $\omega \notin S$.

(d) Compute $E(X)$.

(e) Let Y be another random variable defined on $(\Omega, \mathcal{F}, \mu)$. Argue that X and Y are independent.

3. Let $\{B_i : i = 1, 2, \ldots\}$ be mutually independent Bernoulli random variables with $P(B_i = 1) = \frac{1}{i}$ for $i = 1, 2, \ldots$ and define $R_n = \sum_{i=1}^{n} B_i$, for all $n \geq 1$.

(a) Show $\gamma_n \equiv E(R_n) - \log n \in [0, 1]$ for all $n \geq 1$.

(b) Prove $Var(R_n)/\log n \to 1$ as $n \to \infty$.

(c) Show that

$$\frac{R_n}{\log n} \xrightarrow{p} 1, \quad \text{as } n \to \infty.$$

(d) Show that the Lindeberg CLT applies for the triangular array

$$X_{n,i} \equiv \frac{\left(B_i - \frac{1}{i}\right)}{\sqrt{\log n}}, \quad 1 \leq i \leq n, n \geq 1.$$

(e) Argue that

$$\frac{(R_n - \log n)}{\sqrt{\log n}} \xrightarrow{d} N(0,1), \quad \text{as } n \to \infty$$

(using the results in any of the parts (a)-(d) above, if you think they are relevant).

1. *Answer:*

(a) $\Omega \neq \emptyset$, $\mathcal{C} \subset \mathcal{P}(\Omega)$ is semi-algebra if

(i) $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$($\pi$-system);

(ii) $\forall A \in \mathcal{C} \Rightarrow A^C = \bigcup_{i=1}^k B_i, \{B_i\}_{i=1}^k \subset \mathcal{C}$ are disjoint;

(b) $\Omega \neq \emptyset$, $\mathcal{F} \subset \mathcal{P}(\Omega)$ is an algebra if

(i) $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$;

(ii) $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$;

(iii) $\Omega \in \mathcal{F}$.

(c) $\Omega \neq \emptyset$, $\mathcal{F} \subset \mathcal{P}(\Omega)$ is an σ-algebra if

(i) $\{A_i\} \subset \mathcal{F} \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathcal{F}$;

(ii) $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$;

(iii) $\Omega \in \mathcal{F}$.

(d) $(\Omega, \mathcal{F})$ is called a measurable space if $\Omega \neq \emptyset$ and $\mathcal{F}$ is a σ-algebra.

(e) A set function μ defined on an algebra/semi-algebra/σ-algebra $\mathcal{F}$ ($\Omega \neq \emptyset$) is called a measure if

(i) $\mu : \mathcal{F} \to [0, \infty]$;

(ii) $\mu(\emptyset) = 0$;

(iii) $\forall$ disjoint collection of sets $\{A_i\}_i \subset \mathcal{F}, \mu(\bigcup_i A_i) = \sum_{i=1}^\infty \mu(A_i)$.

(f) A measure μ is called finite if $\mu(\Omega) < \infty$. A finite measure μ with $\mu(\Omega) = 1$ is called a probability measure.

(g) A measure μ is σ-finite if $\exists$ $\{A_i, i = 1, 2, \ldots\} \subset \mathcal{F}$, s.t $\bigcup_i A_i = \Omega$ and $\mu(A_i) < \infty, \forall\ i$.

(h) If μ is a measure on a measurable space $(\Omega, \mathcal{F})$, then $(\Omega, \mathcal{F}, \mu)$ is called a measure space. In addition, if μ is a probability measure, then $(\Omega, \mathcal{F}, \mu)$ is a probability space.

(i) If $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are two measurable spaces and if $X : \Omega_1 \to \Omega_2$ satisfies $X^{-1}(B) \in \mathcal{F}_1, \forall B \in \mathcal{F}_2$, then X is called a measurable transformation.

(j) If $(\Omega_1, \mathcal{F}_1)$ has a probability measure defined on it, and if $(\Omega_2, \mathcal{F}_2) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then the measurable transformation X is called a random variable.

2. *Answer:*

(a) First of all, $\mathcal{F}$ satisfies the first two conditions of being a σ-algebra . Secondly, for any $\{A_n\}_{n\geq 1} \subset \mathcal{F}$, if $\cup_{n\geq 1}A_n$ is countable then $\cup_{n\geq 1}A_n \in \mathcal{F}$. If $\cup_{n\geq 1}A_n$ is not countable, then $\exists$ uncountable set $A_{n_0} \in \{A_n\}_{n\geq 1}$. Because $A_{n_0} \in \mathcal{F}$, so $A^c_{n_0}$ has to be countable, and consequently $A^c_{n_0} \supset \cap_{n\geq 1}A^c_n = (\cup_{n\geq 1}A_n)^c$ is countable, i.e. $\cup_{n\geq 1}A_n \in \mathcal{F}$.

(b) As we can see, we only need to verify the *countable additivity* property for μ on $\mathcal{F}$. For any disjoint collection $\{A_n\}_{n\geq 1} \subset \mathcal{F}$, by the argument in (a), we see that there can be at most one uncountable set in $\{A_n\}_{n\geq 1}$, because they are disjoint (therefore, $\cup_{n\neq j}A_n \subset A^c_j, \forall j$). Thus $\mu(\cup_{n\geq 1}A_n) = \sum_{n\geq 1}\mu(A_n)$, and $(\Omega, \mathcal{F}, \mu)$ is a probability space.

(c) Since X is $\langle \mathcal{F}, \mathcal{B}(I\!\!R)\rangle$-measurable, $\{X < c\} \equiv \{\omega \in \Omega : X(\omega) < c\} \in \mathcal{F}$ for all $c \in I\!\!R$. That is for any $c \in I\!\!R$, $\{X < c\}$ is either countable or cocountable. Take $a = \sup\{c \in I\!\!R : \{X < c\} \text{ is countable}\}$. Then for all $n \in I\!\!N$, $\{X < a + \frac{1}{n}\}$ is co-countable and $\{X < a - \frac{1}{n}\}$ is countable. Hence, being a countable union of countable sets, S is countable where

$$S \doteq \{\omega \in \Omega : X(\omega) \neq a\} = \bigcup_{n\in I\!\!N}\left[\left\{X < a - \frac{1}{n}\right\}\bigcup\left\{X \geq a + \frac{1}{n}\right\}\right]$$

(d) Note that $X = a$ almost surely(μ), since $\mu(X \neq a) = \mu(S) = 0$ because S is countable. So, $E(X) = E(a) = a$.

(e) Since X is a constant random variable, it is independent of any other random variable.

3. *Answer:*

(a) Note $E(R_n) = \sum_{i=1}^n E(B_i) = \sum_{i=1}^n \frac{1}{k}$ and observe that

$$\sum_{i=2}^{n}\frac{1}{k} \;\leq\; \log n \;\leq\; \sum_{i=1}^{n-1}\frac{1}{k} \;\leq \sum_{i=1}^{n}\frac{1}{k} = E(R_n) \tag{1}$$

So,

$$0 \;\leq\; \gamma_n \equiv E(R_n) - \log n \;\leq\; \sum_{i=1}^{n}\frac{1}{k} - \sum_{i=2}^{n}\frac{1}{k} \;=\; 1$$

(b) Note that $Var(R_n) = \sum_{i=1}^n \frac{1}{k}\left(1 - \frac{1}{k}\right) = \sum_{i=1}^n \frac{1}{k} - \sum_{i=1}^n \frac{1}{k^2}$. Hence, from (1) and the facts that $\sum_{i=1}^\infty \frac{1}{k^2} < \infty$ and that $\log n \to \infty$ as $n \to \infty$, we get the result.

(c) For all $\epsilon > 0$, we get from Markov's Inequality,

$$P\left(\left|\frac{R_n}{\log n} - 1\right| > \epsilon\right) \;\leq\; \frac{2\,(Var(R_n) + \gamma_n^2)}{(\log n)^2\epsilon^2} \;\to\; 0$$

as $n \to 0$ (from part (a) and (b)).

(d) Note that

$$|X_{n,i}| \leq (\log n)^{-\frac{1}{2}}, \tag{2}$$

for all i, n (and for al $\omega \in \Omega$). Also, $S_n = \sum_{i=1}^{N} X_{n,i} = \frac{R_n - E(R_n)}{\sqrt{\log n}}$ and $s_n^2 = \sum_{i=1}^{N} E\left(X_{n,i}\right)^2 = \frac{Var(R_n)}{\log n} \to 1$, as $n \to \infty$ (Note that $E\left(X_{n,i}\right) = 0$). So, for any $\epsilon > 0$,

$$\left|\frac{X_{n,i}}{s_n}\right| \leq \frac{1}{s_n\sqrt{\log n}} \; < \; \frac{\epsilon}{2}.2 = \epsilon,$$

for all $n \geq \max(n_1, n_2)$, where $n_1 = \exp\left(4\epsilon^{-2}\right)$ and $n_2 \geq 1$ is such that for all $n \geq n_2$, we have $s_n > 1/2$ (note: $s_n \to 1$ as $n \to \infty$ as shown above).

(e) Using part (e), we have from Lindeberg's CLT,

$$\frac{R_n - E(R_n)}{\sqrt{\log n}} \xrightarrow{d} N(0,1),$$

as $n \to \infty$. Since from part (a) we have

$$\frac{E(R_n) - \log n}{\sqrt{\log n}} = \frac{\gamma_n}{\sqrt{\log n}} \to 0$$

as $n \to \infty$, the result follows from the last two displays above.

Part I

Let (X, Y) be bivariate normally distributed as $N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$ for $|\rho| < 1$. Consider the transformed random variables $U = \Phi\left(\frac{X-\rho Y}{\sqrt{1-\rho^2}}\right)$ and $V = \Phi\left(\frac{Y-\rho X}{\sqrt{1-\rho^2}}\right)$, where Φ is the cumulative distribution function of the standard normal random variable.

1. Show that the conditional distribution of U given $Y = y$ is uniform$(0, 1)$.
2. Show that the marginal distribution of U is uniform$(0, 1)$.
3. Show that U and Y are independent.
4. Find the joint distribution of U and V. Are U and V independent?

   Hint: *You may find it helpful to consider using the intermediate transformation* $Z \equiv \left(\frac{X-\rho Y}{\sqrt{1-\rho^2}}\right)$ *and* $W \equiv \left(\frac{Y-\rho X}{\sqrt{1-\rho^2}}\right)$.

Part II

Suppose that $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ for $n > 3$ is a random sample drawn from the Pareto distribution with probability density function

$$f(x; \theta, \nu) = \frac{\theta \nu^\theta}{x^{\theta+1}} I_{[\nu,\infty)}(x), \qquad \theta > 0, \nu > 0, \tag{1}$$

where $I_{[\nu,\infty)}(x)$ is the indicator function for the interval $[\nu, \infty)$. Write the order statistics from $\boldsymbol{X}$ as $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$. *You may use without proof the fact that variables* $[\log X_{(i+1)} - \log X_{(1)}]/\theta$, $i = 1, 2, \ldots, n-1$, *are distributed as the order statistics of* $n-1$ *independent exponential*(1) *random variables.*

5. Show that, for a fixed $\nu = \nu_0 > 0$, the Pareto density $f(x; \theta, \nu_0)$ is part of the regular exponential family. What about the form of the family in equation (1) for unknown $\nu > 0$ prevents it from being a 2-parameter exponential family?
6. Provide a two-dimensional sufficient statistic for (θ, ν).
7. Find the maximum likelihood estimator of (θ, ν). Call it $(\hat{\theta}_n, \hat{\nu}_n)$.
8. **a)** Show that $\hat{\theta}_n$ in problem (7) is biased for θ, but is asymptotically unbiased.

   b) Show that $\hat{\theta}_n$ in problem (7) is consistent for θ.
9. **a)** Provide the sampling distribution of $X_{(1)}$.

   b) Show that $\hat{\nu}_n$ in problem (7) is consistent for ν.

10. For a given $\alpha \in (0, 1)$, show that the size-α likelihood ratio test of

$$H_0 : \theta = 1 \qquad vs. \qquad H_a : \theta \neq 1$$

for unspecified ν has critical region of the form $\{\boldsymbol{x} | S(\boldsymbol{x}) < c_1 \text{ or } S(\boldsymbol{x}) > c_2\}$, for appropriate $0 < c_1 < c_2$, depending on α and

$$S(\boldsymbol{X}) = \sum_{i=2}^{n} \left(\log X_{(i)} - \log X_{(1)}\right) .$$

11. Find the distribution under H_0 of $2S(\boldsymbol{X})$ (for $S(\boldsymbol{X})$ and H_0 as in problem (10)).

Solutions

Part I.

1. $U = \Phi\left(\frac{X-\rho Y}{\sqrt{(1-\rho^2)}}\right)$. $P\left[U \leq u \mid Y=y\right] = P\left[\Phi\left(\frac{X-\rho Y}{\sqrt{1-\rho^2}}\right) \leq u \mid Y=y\right]$

Note that $\mathcal{L}(X \mid Y=y) \equiv N(\rho Y, 1-\rho^2)$ so that

$$\mathcal{L}\left(\frac{X-\rho Y}{\sqrt{1-\rho^2}} \,\Big|\, Y=y\right) \equiv N(0,1).$$

From the standard result: If $X \sim F(x)$, $F(X) \sim U(0,1)$.

So, the above holds.

2. We have $U = \Phi\left(\frac{X-\rho Y}{\sqrt{1-\rho^2}}\right)$, $V = \Phi\left(\frac{Y-\rho X}{\sqrt{1-\rho^2}}\right)$.

Want the joint distn of (U,V).

Let $Z = \dfrac{X-\rho Y}{\sqrt{1-\rho^2}}$, $W = \dfrac{Y-\rho X}{\sqrt{1-\rho^2}}$.

Then $\begin{pmatrix} Z \\ W \end{pmatrix} \sim BN(0,0,1,1,-\rho)$. So, $U = \Phi(Z)$, $V = \Phi(W)$

[This follows from $\begin{pmatrix} Z \\ W \end{pmatrix} = \frac{1}{\sqrt{1-\rho^2}}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$. Then $\begin{pmatrix} Z \\ W \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ - \end{matrix}, \frac{1}{\sqrt{1-\rho^2}}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}\right)$]

Then the inverse transformation is given by $Z = \Phi^{-1}(U)$, $W = \Phi^{-1}(V)$.

The Jacobian is given by

$$|J| = \begin{vmatrix} \frac{1}{\varphi(\Phi^{-1}(u))} & 0 \\ 0 & \frac{1}{\varphi(\Phi^{-1}(v))} \end{vmatrix} = \frac{1}{\varphi(\Phi^{-1}(u))\,\varphi(\Phi^{-1}(v))}.$$

The joint distn of (U,V) is given by

$$f_{U,V}(u,v) = \frac{1}{2\pi(1-\rho^2)} \exp\left\{ \frac{-1}{2(1-\rho^2)}\left[\left(\Phi^{-1}(u)\right)^2 + 2\rho\,\Phi^{-1}(u)\Phi^{-1}(v) + \left(\Phi^{-1}(v)\right)^2\right]\right\} \frac{1}{\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}$$

The product term in the exponent means that we have that unless $\rho = 0$, $f_{U,V}(u,v) \neq f_U(u)\, f_V(v)$ so that U and V are not independent.

3. $U = \Phi(Z)$, $Z \sim N(0,1)$. So, again, it is clear that $U \sim U(0,1)$

4. From (1) and (3) since $\mathcal{L}(U)$ and $\mathcal{L}(U \mid Y)$ are both $U(0,1)$, $U \perp\!\!\!\perp Y$.

Part II

5. $X_1, X_2, \ldots, X_n$ iid $f(x;\theta,\nu) = \dfrac{\theta \nu^{\theta}}{x^{\theta+1}} I_{(\nu,\infty)}(x)$, $\theta>0, \nu>0$.

For given ν, we have $f(x;\theta,\nu) = \exp\{(\theta+1)\ln\nu + \ln\theta - (\theta+1)\ln x\} \cdot I_{(\nu,\infty)}(x)$.

The above is a REF.

However, for unknown ν, the range of the density depends on the parameter (ν). So, this cannot be a REF density.

6. The joint density of $X_1, X_2, \ldots, X_n$ is given by

$$f(x_1, x_2, \ldots, x_n; \theta, \nu) = \frac{\theta^n \nu^{n\theta}}{\prod_{i=1}^{n} x_i^{\theta+1}} I_{(\nu,\infty)}(x_{(1)}).$$

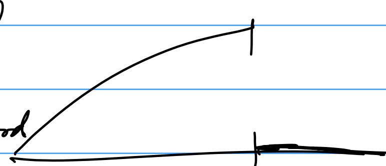
So the 2-D sufficient statistic for (θ,ν) is given by $\left(\prod_{i=1}^{n} X_i, X_{(1)}\right)$, or equivalently $\left(\sum_{i=1}^{n} \ln X_i, X_{(1)}\right)$.

7. Consider the loglikelihood function of the parameters given the observations

$$\ell(\theta,\nu; X_1, X_2, \ldots, X_n) = n\ln\theta + n\theta\ln\nu - (\theta+1)\sum_{i=1}^{n}\ln X_i + \ln I_{(\nu,\infty)}(X_{(1)}).$$

To find the MLE for (θ,ν), first consider $\ell(\theta,\nu; X_1, X_2, \ldots, X_n)$ as a function of ν for each θ.

Note that for each θ, $\ell(\theta,\nu; X_1, X_2, \ldots, X_n)$ is an increasing function in ν as long as ν is not above $X_{(1)}$. Above $X_{(1)}$, the loglikelihood function tends to $-\infty$. So the MLE of ν is $\hat{\nu} = X_{(1)}$.



$X_{(1)}$

Next, find the profile loglikelihood of θ for $\hat{\nu} = X_{(1)}$.

$$\frac{\partial \ell(\theta,\hat{\nu}; X_1, X_2, \ldots, X_n)}{\partial\theta} = \frac{n}{\theta} + n\ln X_{(1)} - \sum_{i=1}^{n}\ln X_i = 0$$

$$\Rightarrow \frac{1}{\theta} = \frac{1}{n}\sum_{i=1}^{n}\ln X_i - \ln X_{(1)}.$$

$$\Rightarrow \hat{\theta} = \left[\frac{1}{n}\sum_{i=2}^{n}\left[\ln X_{(i)} - \ln X_{(1)}\right]\right]^{-1} = \frac{n}{\sum_{i=1}^{n}(\ln X_{(i)} - \ln X_{(1)})}$$

8 (a) We have for $i=2,3,\dots,n$, $\ln X_{(i)} - \ln X_{(1)}$ is the order stat. of size $(n-1)$ from $\exp(\theta)$.
Therefore $T = \sum_{i=2}^{n} (\ln X_{(i)} - \ln X_{(1)}) \sim \Gamma(n-1, \theta)$, since this is the sum of $(n-1)$ indep. $\exp(1)$ r.v.s.

Then $f(t) = \dfrac{1}{\Gamma(n-1)\,\theta^{n}} t^{n-2} e^{-t/\theta}$

So $E[T^{-1}] = \dfrac{\Gamma(n-2)\,\theta}{\Gamma(n-1)} = \dfrac{\theta}{n-2}$.

Thus $E(\hat{\theta}) = \dfrac{n}{n-2}\theta$.

So that $\hat{\theta}$ is biased for θ, but asymptotically not so.

(b) Also, $\operatorname{Var}(\hat{\theta}) = E(\hat{\theta}^2) - [E(\hat{\theta})]^2$
$$= n^2 E[T^{-2}] - \left[\frac{\theta}{(n-2)}\right]^2$$

Now $E[T^{-2}] = \dfrac{\theta^2}{(n-2)(n-3)}$ so that $\operatorname{Var}(\hat{\theta}) = n^2\theta^2\left[\dfrac{1}{(n-2)(n-3)} - \dfrac{1}{(n-2)^2}\right]$
$$= n^2\theta^2\left[\frac{(n-2)-(n-3)}{(n-2)^2(n-3)}\right]$$
$$= \frac{n^2\theta^2}{(n-2)^2(n-3)} \longrightarrow 0 \text{ as } n\to\infty$$

So $\hat{\theta}$ is consistent for θ.

9 (a) Sampling distn of $\hat{\nu} \equiv X_{(1)}$.
$$G(x) = P(X_{(1)} \leq x) = 1 - P[X_{(1)} > x]$$
$$= 1 - \prod_{i=1}^{n} P(X_i > x)$$
$$= 1 - [1 - F(x)]^n$$
$$F(x) = \int_{\nu}^{x} \frac{\theta \nu^{\theta}}{y^{\theta+1}}\,dy = \theta \nu^{\theta} \int_{\nu}^{x} \frac{dy}{y^{\theta+1}} = \left[1 - \left(\frac{\nu}{x}\right)^{\theta}\right] I_{(\nu,\infty)}(x)$$

So $G(x) = 1 - \left(\dfrac{\nu}{x}\right)^{n\theta} I_{(\nu,\infty)}(x)$

Thus $X_{(1)} \sim$ Pareto with parameters ν and $n\theta$.

(b) $E(X_{(1)}) = \int_{\nu}^{\infty} x \cdot \frac{(n\theta)\nu^{n\theta}}{x^{n\theta+1}} dx = \nu \int_{\nu}^{\infty} \frac{(n\theta)\nu^{n\theta-1}}{x^{n\theta}} dx = \frac{\nu n\theta}{(n\theta-1)} \int \frac{(n\theta-1)\nu^{n\theta-1}}{x^{n\theta}} dx$

$= \frac{n\nu\theta}{n\theta-1}.$

$E(X_{(1)}^2) = \int_{\nu}^{\infty} x^2 \frac{(n\theta)\nu^{n\theta}}{x^{n\theta+1}} dx = \nu^2 \int_{\nu}^{\infty} \frac{n\theta \nu^{n\theta-2} dx}{x^{n\theta-1}}$

$= \frac{\nu^2 n\theta}{n\theta-2}$

Thus $Var(\hat{\nu}_n) = Var(X_{(1)}) = E(X_{(1)}^2) - [E(X_{(1)})]^2$

$= n\theta\nu^2 \left[\frac{1}{n\theta-2} - \frac{n\theta}{(n\theta-1)^2}\right] = \frac{n\theta\nu^2}{(n\theta-1)^2(n\theta-2)} \to 0$ as $n \to \infty$

$E(\hat{\nu}_n)$ is asymptotically unbiased and $Var(\hat{\nu}_n) \to 0$ as $n \to \infty$.
So $\hat{\nu}_n$ is consistent for ν.

10. The LRT for testing $H_0: \theta = 1$ vs. $H_a: \theta \neq 1$.
For $\theta = 1$, $L(\nu; X_1, X_2, \ldots, X_n)$ has maximized value
$\frac{X_{(1)}}{\sum_{i=1}^{n} X_i^2} I(0 < \nu < X_{(1)})$

For θ, $L(\theta, \nu; X_1, X_2, \ldots, X_n)$ has maximized value:
$\frac{\hat{\theta}^n X_{(1)}^{n\hat{\theta}}}{\prod_{i=1}^{n} X_i^{\hat{\theta}+1}} I(0 < \nu < X_{(1)})$

The likelihood ratio test is given by

$$\Lambda = \frac{\sup_{\theta \in \Theta_0} L(\theta, \nu; X_1, X_2, \ldots, X_n)}{\sup_{\theta \in \Theta_0 \cup \Theta_1} L(\theta, \nu; X_1, X_2, \ldots, X_n)} = \frac{X_{(1)}^n \prod_{i=1}^{n} X_i^{\hat{\theta}}}{\hat{\theta}^n X_{(1)}^{n\hat{\theta}} \prod_{i=1}^{n} X_i}$$

Reject H_0 in favor of H_a at level α if $\Lambda < \Lambda_\alpha$.

i.e, $n \ln X_{(1)} + \hat{\theta} \sum_{i=1}^{n} (\ln X_{(i)} - \ln X_{(1)}) - n \ln \hat{\theta} - \sum_{i=1}^{n} \ln X_i <$ Const.

i.e, $-n \ln \hat{\theta} - \sum_{i=2}^{n} [\ln X_{(i)} - \ln X_{(1)}] <$ const.

i.e, $-n \ln \hat{\theta} - \frac{n}{\hat{\theta}} <$ const.

Consider the function $h(r) = \ln r + \frac{1}{r}$.

$h'(r) = \frac{1}{r} - \frac{1}{r^2} \lesseqgtr 0 \iff r \lesseqgtr 1$

Note that $-h(r)$ is $\uparrow$ for $r < 0$ and $\downarrow$ for $r > 0$.

$-h(r)$ is small for r small and for r large.
Thus the critical region in terms of S is of the form $S < c_1$ or $S > c_2$.
where $S = \ln\left(\prod_{2}^{n} X_i \Big/ X_{(1)}^{n-1}\right) \equiv \frac{n}{\hat{\theta}}$

11. From the above, we have that the distn of

$$2S = 2\sum_{i=2}^{n} \exp(\theta) \text{ r.v.'s}$$

so that $2S \sim \chi^2_{(n-1)}$.