MATH5905 - Statistical Inference
Assignment 2

James Yang
z3415436

Question 1.

Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables from a population with a density

$$f(x;\theta) = \begin{cases} \frac{2\theta^2}{x^3}, & \text{if } x \geqslant \theta \\ 0, & \text{otherwise.} \end{cases}$$

where $\theta > 0$ is an unknown parameter.

a)

The likelihood function is

$$\begin{aligned} L(\mathbf{x}, \theta) &= \prod_{i=1}^{n} \frac{2\theta^2}{x_i^3} I_{[\theta,\infty)}(x_i) \\ &= 2^n \theta^{2n} \prod_{i=1}^{n} \frac{1}{x_i^3} I_{(0,x_{(1)}]}(\theta). \end{aligned}$$

Therefore

$$\begin{aligned} \frac{L(\mathbf{x}, \theta)}{L(\mathbf{y}, \theta)} &= \frac{2^n \theta^{2n} \prod_{i=1}^{n} \frac{1}{x_i^3} I_{(0,x_{(1)}]}(\theta)}{2^n \theta^{2n} \prod_{i=1}^{n} \frac{1}{y_i^3} I_{(0,x_{(1)}]}(\theta)} \\ &= \prod_{i=1}^{n} \frac{y_i^3}{x_i^3} \frac{I_{(0,x_{(1)}]}(\theta)}{I_{(0,x_{(1)}]}(\theta)}. \end{aligned}$$

This expression is independent of θ iff $x_{(1)} = y_{(1)}$. Hence $T = X_{(1)}$ is a minimal sufficient statistic for θ.

b)

For $x > \theta$,

$$\begin{aligned}
F_{X_{(1)}}(x) &= P(X_{(1)} \leqslant x) \\
&= 1 - P(X_1 \geqslant x, X_2 \geqslant x, \ldots, X_n \geqslant x) \\
&= 1 - [P(X_i \geqslant x)]^n \text{ since } X_i \text{ are i.i.d. for i = 1,2,\ldots,n} \\
&= 1 - \left[1 - \left(1 - \frac{\theta^2}{x^2}\right)\right]^n \\
&= 1 - \frac{\theta^{2n}}{x^{2n}}.
\end{aligned}$$

Differentiating with respect to x

$$f_{X_{(1)}}(x) = \frac{2n\theta^{2n}}{x^{2n+1}}.$$

Hence

$$f_{X_{(1)}}(x) = \begin{cases} \frac{2n\theta^{2n}}{x^{2n+1}}, & \text{if } x \geqslant \theta \\ 0, & \text{otherwise.} \end{cases}$$

c)

Recall that

$$L(\mathbf{x}, \theta) = 2^n\theta^{2n} \prod_{i=1}^{n} \frac{1}{x_i^3} I_{(0,x_{(1)}]}(\theta).$$

Since this is a monotonically increasing function in θ, the likelihood function is maximised when θ is largest. That is, when $\theta = x_{(1)}$. Thus the MLE is $\widehat{\theta} = x_{(1)}$.

d)

Suppose $E_\theta(g(X_{(1)})) = 0$. That is,

$$\int_\theta^\infty \frac{2ng(x)\theta^{2n}}{x^{2n+1}} dx = 0$$

$$\int_\theta^\infty \frac{g(x)}{x^{2n+1}} dx = 0 \text{ as } 2n\theta^{2n} \text{ is constant in } x,$$

$$\frac{d}{d\theta}\int_\theta^\infty \frac{g(x)}{x^{2n+1}} dx = \frac{d}{d\theta}(0)$$

$$\frac{g(\theta)}{\theta^{2n+1}} = 0.$$

Since $\theta > 0$, we have that $g(\theta) = 0$ for all $\theta > 0$. As $x > 0$, $g(x) = 0$ for all x. Hence $P(g(x) = 0) = 1$ and $X_{(1)}$ is complete for θ.
Note we have also proven that $X_{(1)}$ is sufficient in θ. Now,

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

$$= \frac{2n}{2n-1}\theta.$$

Hence an unbiased estimator for θ is $\frac{2n-1}{2n}\theta$ which is also a function of a complete and sufficient statistic for θ. Hence $\frac{2n-1}{2n}\theta$ is the UMVUE of θ.

e)

Let θ', θ'' be fixed such that $0 < \theta' < \theta''$.

$$\frac{L(\mathbf{X}, \theta'')}{L(\mathbf{X}, \theta')} = \frac{2^n \theta''^{2n} \prod_{i=1}^n \frac{1}{x_i^3} I_{(0,x_{(1)}]}(\theta'')}{2^n \theta'^{2n} \prod_{i=1}^n \frac{1}{x_i^3} I_{(0,x_{(1)}]}(\theta')}$$

$$= \left(\frac{\theta''}{\theta'}\right)^{2n} \frac{I_{(0,x_{(1)}]}(\theta'')}{I_{(0,x_{(1)}]}(\theta')}$$

$$= \left(\frac{\theta''}{\theta'}\right)^{2n} \frac{I_{(0,T]}(\theta'')}{I_{(0,T]}(\theta')}$$

$$= \left(\frac{\theta''}{\theta'}\right)^{2n}, \quad 0 < T \leqslant \theta'.$$

For $0 < \theta' < \theta''$, this is a non-decreasing function of T. Hence the family $L(\mathbf{X}, \theta)$ has a MLR in $T = X_{(1)}$.

f)

By the Theorem of Blackwell and Girshick the UMP α-size test ϕ^* of $H_0 : \theta \leqslant 1$ and $H_1 : \theta > 1$ has a structure:

$$\phi^* = \begin{cases} 1, & \text{if } T > k \\ 0, & \text{if } T < k. \end{cases}$$

where k is the upper $\alpha.100\%$ point of the P_θ distribution of T. That is,

$$\begin{aligned} \alpha &= P(T = X_{(1)} \geqslant k | \theta = 1) \\ &= \frac{1}{k^{2n}}. \end{aligned}$$

Hence $k = \alpha^{-\frac{1}{2n}}$.

g)

The power function of ϕ^* is

$$\begin{aligned} E_\theta[\phi^*] &= P(T > \alpha^{-\frac{1}{2n}}) \\ &= \frac{\theta^{2n}}{(\alpha^{-\frac{1}{2n}})^{2n}} \\ &= \alpha\theta^{2n}, \quad \theta > 0. \end{aligned}$$

Question 2.

Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be i.i.d. random variables each with the log-normal density

$$f(x,\theta) = \begin{cases} \frac{1}{\sqrt{2\pi}x\theta} e^{\left[-\frac{1}{2}[\frac{ln(x)}{\theta}]^2\right]}, x > 0 \\ 0, \quad \text{otherwise} \end{cases}$$

where $\theta > 0$ is a parameter.

a)

Let $T(\mathbf{X}) = \Sigma_{i=1}^n (\ln X_i)^2$ and let θ', θ'' be fixed such that $0 < \theta' < \theta''$. The likelihood function is

$$L(\mathbf{X}, \theta) = \frac{1}{(2\pi)^{n/2}\theta^n \prod_{i=1}^n x_i} exp\left[-\frac{\Sigma_{i=1}^n (\ln X_i)^2}{2\theta^2}\right].$$

The likelihood ratio is

$$\begin{aligned} \frac{L(\mathbf{X}, \theta'')}{L(\mathbf{X}, \theta')} &= \left(\frac{\theta'}{\theta''}\right)^n \exp\left[-\frac{1}{2}\left(\frac{1}{\theta''} - \frac{1}{\theta'}\right)\Sigma_{i=1}^n (\ln X_i)^2\right] \\ &= \left(\frac{\theta'}{\theta''}\right)^n \exp\left[\frac{1}{2}\left(\frac{\theta'' - \theta'}{\theta'\theta''}\right) T(\mathbf{X})\right]. \end{aligned}$$

Since $\frac{\theta''-\theta'}{\theta'\theta''} > 0$, the likelihood ratio is a strictly increasing function of $T(\mathbf{X})$. Hence the family $L(\mathbf{X}, \theta)$ has a monotone likelihood ratio in $T(\mathbf{X})$.

b)

The family $L(\mathbf{X}, \theta)$ is MLR in $T(\mathbf{X})$. By the Theorem of Blackwell and Girshick, there is a UMP α - size test ϕ^* of the hypothesis $H_0 : \theta \leqslant \theta_0$ against $H_1 : \theta > \theta_0$ and the structure of this test is

$$\phi^*(\mathbf{x}) = \begin{cases} 1 \text{ if } T(\mathbf{X}) \geqslant k \\ 0 \text{ if } T(\mathbf{X}) < k. \end{cases}$$

where k is the upper α 100% of the P_{θ_0} - distribution of $T(\mathbf{X})$.

c)

Let $Y = \ln X$. Then $x = \exp(y)$ and $\frac{dx}{dy} = \exp(y)$. Hence the density of Y is

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi}\exp(y)\theta}\exp\left(-\frac{y^2}{2\theta^2}\right) \times \exp(y) \\ &= \frac{1}{\sqrt{2\pi}\theta}\exp\left(-\frac{y^2}{2\theta^2}\right), \quad -\infty < y < \infty. \end{aligned}$$

Hence $Y \sim N(0, \theta^2)$.

d)

If $Y_i \sim N(0, \theta^2)$ then $\frac{Y_i}{\theta} \sim N(0, 1)$. Hence $\sum_{i=1}^n \left(\frac{Y_i}{\theta}\right)^2 = \frac{1}{\theta^2}\sum_{i=1}^n Y_i^2 \sim \chi_n^2$. We have that,

$$\begin{aligned} P(\Sigma_{i=1}^n(\ln X_i)^2 \geqslant k|\theta = \theta_0) &= P\left(\frac{1}{\theta_0^2}\Sigma_{i=1}^n(\ln X_i)^2 \geqslant \frac{k}{\theta_0^2}|\theta = \theta_0\right) \\ &= \int_{\frac{k}{\theta_0^2}}^{\infty} \frac{x^{\frac{n}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}. \end{aligned}$$

Hence the threshold is $k = \theta_0^2\chi_{n,\alpha}^2$ where $\chi_{n,\alpha}^2$ is the upper $\alpha \times 100\%$ point of the χ_n^2 distribution.
The complete UMP α - size test ϕ^* of $H_0 : \theta \leqslant \theta_0$ versus $H_1 : \theta > \theta_0$ is

$$\phi^* = \begin{cases} 1, \text{ if } T(\mathbf{X}) \geqslant \theta_0^2\chi_{n,\alpha}^2 \\ 0 \text{ if } T(\mathbf{X}) < \theta_0^2\chi_{n,\alpha}^2. \end{cases}$$

The power function of ϕ^* is

$$\begin{aligned} E_\theta[\phi^*] &= P(\Sigma_{i=1}^n(\ln X_i)^2 \geqslant \theta_0^2\chi_{n,\alpha}^2) \\ &= P\left(\frac{1}{\theta^2}\Sigma_{i=1}^n(\ln X_i)^2 \geqslant \frac{\theta_0^2\chi_{n,\alpha}^2}{\theta^2}\right) \\ &= P\left(\chi_n^2 \geqslant \frac{\theta_0^2\chi_{n,\alpha}^2}{\theta^2}\right) \end{aligned}$$

Hence $E_\theta[\phi^*]$ is an increasing function of $\theta \in (0, \infty)$, and satisfies $E_{\theta=0}[\phi^*] = 0, E_{\theta=\theta_0}[\phi^*] = \alpha$ and $\lim_{\theta\to\infty} E_\theta[\phi^*] = 1$.

Question 4.

Suppose $X_{(1)} < X_{(2)} < X_{(3)}$ are the order statistics based on a random sample of size 3 from the standard exponential density $f(x) = e^{-x}, x > 0$.

i)

The general formula for the density of an order statistic is

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!}[F(x)]^{i-1}[1-F(x)]^{n-i}f(x)$$

The density of $X_{(3)}$ is

$$\begin{aligned} f_{X_{(3)}}(x) &= \frac{3!}{(3-1)!(3-3)!}[1-e^{-x}]^{3-1}[e^{-x}]^{3-3}e^{-x} \\ &= 3[1-e^{-x}]^2 e^{-x}, 0 < x < \infty \end{aligned}$$

So then

$$\begin{aligned} EX_{(3)} &= 3\int_0^\infty x[1-e^{-x}]^2 e^{-x} dx \\ &= 3\left[\Gamma(2) - \frac{1}{2}\Gamma(2) + \frac{1}{9}\Gamma(2)\right] \\ &= \frac{11}{6}. \end{aligned}$$

ii)

It holds in general that

$$f_{X_{(1)},X_{(n)}}(x_{(1)}, x_{(n)}) = n(n-1)[F_X(x_{(n)}) - F_X(x_{(1)})]^{n-2} f_X(x_{(1)}) f_X(x_{(n)}), \quad x_{(1)} < x_{(n)}.$$

So the joint density of $X_{(1)}$ and $X_{(3)}$ is

$$f_{X_{(1)},X_{(3)}}(x_{(1)}, x_{(3)}) = 6[e^{-x_{(1)}} - e^{-x_{(3)}}]e^{-x_{(1)}}e^{-x_{(3)}}, \quad 0 < x_{(1)} < x_{(3)} < \infty.$$

Let $A = X_{(1)}$ and $B = \frac{1}{2}(X_{(1)} + X_{(3)})$. Rearranging, we have that $x_{(1)} = a$ and $x_{(3)} = 2b - a$ with a Jacobian of transformation equal to 2. Hence the joint density of A and B is

$$\begin{aligned} f_{A,B}(a,b) &= 6[e^{-a} - e^{a-2b}]e^{-2b} \times 2 \\ &= 12[e^{-a-2b} - e^{a-4b}]. \end{aligned}$$

The relationship $0 < x_{(1)} < x_{(3)} < \infty$ transforms to $0 < a < b < \infty$. Hence the density of the median B is

$$\begin{aligned} f_B(b) &= 12\int_0^b [e^{-a-2b} - e^{a-4b}]da \\ &= 12(e^{-2b} - 2e^{-3b} + e^{-4b}), \quad 0 < b < \infty. \end{aligned}$$

Now,

$$\begin{aligned} P(B > 1) &= 1 - 12\int_0^1 (e^{-2b} - 2e^{-3b} + e^{-4b})db \\ &= 0.468662. \end{aligned}$$

as required.