BSDS (2024 II): Statistics II

Quiz1

NAME :	ROLL :

Time: 65 minutes **Total attainable marks:** 40

1. Let $X_1, \cdots, X_n \overset{iid}{\sim} U(\theta, 1)$ $(\theta < 1)$,

   (a) Find MLE of θ. [5]

$$f_{X_i}(x_i) = \frac{1}{(1-\theta)} \quad ; \quad \theta < x_i < 1 \quad , i=1,2,\ldots,n$$

Likelihood:

$$L(\theta \mid x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

$$= \frac{1}{(1-\theta)^n} \quad ; \quad \theta < x_i < 1 \quad \forall i$$

$$= \frac{1}{(1-\theta)^n} \quad ; \quad \theta < x_{(1)} < \cdots < x_{(n)} < 1$$

Observe that, $L(\theta \mid x_1, \ldots, x_n)$ is an increasing function of θ.

$\therefore$ Maximum value of θ, will maximize the likelihood.

$\therefore \hat{\theta}_{MLE} = X_{(1)}$.

(b) Find c and d such that $c + d\,\hat{\theta}_{\text{MLE}}$ is an unbiased estimator of θ. [5]

Let us first calculate, $E(X_{(1)})$ or $E(\hat{\theta}_{MLE})$

We know,

$$f_{X_{(1)}}(x) = n\,(1 - F(x))^{n-1} f(x)$$

for $X \sim U(\theta, 1)$,

$$F_X(x) = \int_{\theta}^{x} \frac{1}{1-\theta}\,dt = \frac{x-\theta}{1-\theta} \quad ; \theta < x < 1$$

$$\therefore f_{X_{(1)}}(x) = n\left(\frac{1-x}{1-\theta}\right)^{n-1} \frac{1}{1-\theta} \quad ; 1 \geq x > \theta$$

$$= \frac{n}{(1-\theta)^n}(1-x)^{n-1} \quad ; x > \theta$$

$$E(X_{(1)}) = \frac{n}{(1-\theta)^n}\int_{\theta}^{1} x\,(1-x)^{n-1}\,dx$$

Take, $z = 1 - x$

$$= \frac{n}{(1-\theta)^n}\int_{0}^{1-\theta} z^{n-1}(1-z)\,dz$$

$$= \frac{n}{(1-\theta)^n}\left[\frac{z^n}{n}\Big|_0^{1-\theta} - \frac{z^{n+1}}{n+1}\Big|_0^{1-\theta}\right]$$

$$= n\left[\frac{1}{n} - \frac{1-\theta}{n+1}\right]$$

$$E(X_{(1)}) = \frac{n}{n+1}\theta + \frac{1}{n+1}$$

$$\Rightarrow E\left[\frac{n+1}{n}\left(X_{(1)} - \frac{1}{n+1}\right)\right] = \theta$$

$$\Rightarrow E\left[\frac{n+1}{n}X_{(1)} - \frac{1}{n}\right] = \theta$$

$$\therefore d = \frac{n+1}{n} \quad \text{and} \quad c = -\frac{1}{n}$$

2. Let $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, 1)$, $\mu \in \mathbb{R}$, and $T(\mathbf{X})$ be an unbiased estimator of $\psi(\mu) = \mu^2$.

(a) Show that Cramer Rao Lower Bound for $\text{Var}(T(\mathbf{X}))$ is $\frac{4\mu^2}{n}$. [6]

From the definition of CRLB,

$$\text{Var}(T(\underset{\sim}{X})) \geq \frac{\{\psi'(\mu)\}^2}{I(\mu)} = \frac{(2\mu)^2}{I(\mu)} = \frac{4\mu^2}{I(\mu)}$$

Now, $I(\mu) = -E\left[\frac{\partial^2}{\partial\mu^2} \ln f(\underset{\sim}{x};\mu)\right]$

$$\ln f(x_1, \ldots, x_n; \mu) = \ln\left[\frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}}\right]$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}$$

$$\frac{\partial}{\partial\mu} \ln f(\underset{\sim}{x}, \mu) = \sum_{i=1}^{n}(x_i - \mu) = \sum_{i=1}^{n} x_i - n\mu$$

$$\frac{\partial}{\partial\mu^2} \ln f(\underset{\sim}{x};\mu) = -n$$

$\therefore I(\mu) = n$

Hence, CRLB for $\text{Var}(T(\underset{\sim}{X})) = \frac{4\mu^2}{n}$

(b) Show that, $T(\mathbf{X}) = \bar{X}_n^2 - \frac{1}{n}$ is the UMVUE of $\psi(\mu)$. [5]

- we want to show that, it belongs to an exponential family.

$$\ln f(\underset{\sim}{x}; \mu) = -\frac{n}{2} \ln(2\pi) - \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{2}$$

$$\Rightarrow f(\underset{\sim}{x}; \mu) = \exp\left\{-\frac{n}{2} \ln(2\pi) - \frac{\sum x_i^2}{2} + \mu \sum_{i=1}^{n} x_i - \frac{n\mu^2}{2}\right\}$$

$\therefore h(\underset{\sim}{x}) = \sum_{i=1}^{n} x_i$ is a complete and sufficient statistic.

Now, as $T(\underset{\sim}{X}) = g(h(\underset{\sim}{x}))$, it is sufficient to show that,

$E(T(\underset{\sim}{x})) = \psi(\mu)$. (Check tutorial for week 3)

$$E\left(\bar{X}_n^2 - \frac{1}{n}\right).$$

$$= E(\bar{X}_n^2) - \frac{1}{n}$$

$$= V(\bar{X}_n) + (E(\bar{X}_n))^2 - \frac{1}{n}$$

$$= \frac{1}{n} + (\mu)^2 - \frac{1}{n} = \mu^2$$

Hence, $T(\underset{\sim}{x})$ is the UMVUE ~~for~~ of $\psi(\mu) = \mu^2$.

(c) Find variance of $T(\mathbf{X})$. $\left(Hint: \mathbb{E}(\bar{X}_n^4) = \mu^4 + \frac{6\mu^2}{n} + \frac{3}{n^2}\right)$ [4]

$$V(T(\underset{\sim}{X}))$$

$$= V\left(\bar{X}_n^2 - \frac{1}{n}\right).$$

$$= V(\bar{X}_n^2)$$

$$= E(\bar{X}_n^4) - \left(E(\bar{X}_n^2)\right)^2$$

$$= \mu^4 + \frac{6\mu^2}{n} + \frac{3}{n^2} - \left(\mu^2 + \frac{1}{n}\right)^2 \quad \left(\text{as } E\left(\bar{X}^2 - \frac{1}{n}\right) = \mu^2\right).$$

$$= \mu^4 + \frac{6\mu^2}{n} + \frac{3}{n^2} - \mu^4 - \frac{2\mu^2}{n} - \frac{1}{n^2}$$

$$= \frac{4\mu^2}{n} + \frac{2}{n^2}$$

$\therefore$ Variance of $T(\underset{\sim}{X}) = \frac{4\mu^2}{n} + \frac{2}{n^2}$

[Recall, CRLB for $Var(T(\underset{\sim}{X})) = \frac{4\mu^2}{n}$, but variance of the UMVUE $= \frac{4\mu^2}{n} + \frac{2}{n^2}$,

Hence, ~~here~~ CRLB is not attained here.]

3. Let $X_1, \cdots, X_n$ be a random sample from

$$f_\theta(x) = \begin{cases} \dfrac{\theta}{x^{\theta+1}} & \text{if } x > 1;\ \theta > 0, \\ 0 & \text{o.w.} \end{cases}.$$

(a) Check weather it belongs to an exponential family, i.e. show that,

$$f_\theta(x) = \exp\{h(x) + c(\theta) + \theta T(x)\},$$

for some functions h, c and T. Find $T(\mathbf{X})$. [4]

$$f_{X_i}(x_i;\theta) = \frac{\theta}{x_i^{\theta+1}} \quad ; \quad x_i > 1 \ . \ \forall i = 1(1)n .$$

$$\Rightarrow f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = \prod_{i=1}^{n} f_{X_i}(x_i;\theta)$$

$$= \frac{\theta^n}{\left(\prod_{i=1}^{n} x_i\right)^{\theta+1}}$$

$$\Rightarrow \ln f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = n\ln\theta - (\theta+1)\sum_{i=1}^{n} \ln x_i$$

$$\Rightarrow f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = \exp\left\{n\ln\theta - \theta\sum_{i=1}^{n}\ln x_i - \sum \ln x_i\right\}$$

$$\therefore T(\underset{\sim}{X}) = \sum_{i=1}^{n} \ln x_i .$$

(b) Based on this T, find the UMVUE of $\frac{1}{\theta}$. [6]

As, $T(\underline{X})$ is complete and sufficient statistic, if $\exists$ a function g, such that $E[g(T(\underline{X}))] = \frac{1}{\theta}$, then $g(T(\underline{X}))$ is said to be the UMVUE of $\frac{1}{\theta}$.

Claim: $\frac{1}{n}\sum_{i=1}^{n} \ln X_i$ is the UMVUE of $\frac{1}{\theta}$.

$$E(\ln X_1) = \int_1^\infty \ln x_1 \frac{\theta}{x_1^{\theta+1}} dx_1$$

Let $\theta \ln x_1 = z$

$\Rightarrow \frac{\theta}{x_1} dx_1 = dz$

and $x_1 = e^{z/\theta}$.

$$= \int_0^\infty \frac{z}{\theta} \cdot \frac{1}{(e^{z/\theta})^\theta} dz$$

$$= \frac{1}{\theta} \int_0^\infty z e^{-z} dz.$$

$$= \frac{1}{\theta} \Gamma 2 = \frac{1}{\theta}.$$

$$\therefore E\left(\frac{1}{n}\sum_{i=1}^{n} \ln X_i\right) = \frac{1}{\theta}$$

Hence, $\frac{1}{n}\sum_{i=1}^{n} \ln X_i$ is the UMVUE of $\frac{1}{\theta}$.

4. (a) Let $Y \sim \text{Exponential}(1)$. Derive the pdf (say f_X) for $X = \left(\frac{Y}{\lambda}\right)^k$, where $\lambda > 0$. [5]

From the change of variable technique,

$$f_X(x) = f_Y(h^{-1}(x)) \left|\frac{dy}{dx}\right|.$$

where $X = h(Y) = \left(\frac{Y}{\lambda}\right)^k$.

$$\Rightarrow Y = \lambda X^{1/k} = h^{-1}(X).$$

and $\left|\frac{dy}{dx}\right| = \left|\frac{d(\lambda X^{1/k})}{dx}\right| = \lambda \cdot \frac{1}{k} \cdot X^{1/k - 1}$

$$\therefore f_X(x) = f_Y(\lambda x^{1/k}) \frac{\lambda}{k} x^{1/k-1}$$

$$= \frac{\lambda}{k} x^{1/k-1} e^{-\lambda x^{1/k}} \quad (\text{as } f_Y(y) = e^{-y}).$$

$\therefore$ The pdf of X is —

$$f_X(x) = \frac{\lambda}{k} x^{1/k-1} e^{-\lambda x^{1/k}} \quad ; x > 0, \lambda > 0.$$

(b) Provide a method to simulate from the distribution of X, starting from uniform$(0,1)$ random variables.

[5]

Here X is an absolutely continuous random variable,

CDF of X, $F_X(x) = U \sim$ Uniform $(0,1)$.

$$F_X(x) = P(X \leq x)$$
$$= P\left(\left(\frac{Y}{\lambda}\right)^K \leq x\right)$$
$$= P(Y \leq \lambda x^{1/K})$$
$$= F_Y(\lambda x^{1/K})$$
$$= 1 - e^{-\lambda x^{1/K}} \quad \left(\text{As } F_Y(y) = 1 - e^{-y}\right).$$

Now, $F_X(x) = u$

$$\Rightarrow 1 - e^{-\lambda x^{1/K}} = u$$
$$\Rightarrow e^{-\lambda x^{1/K}} = 1 - u$$
$$\Rightarrow \lambda x^{1/K} = -\ln(1-u)$$
$$\Rightarrow x = \left(-\frac{\ln(1-u)}{\lambda}\right)^K$$

$\therefore$ If we have a random sample $u_1, \ldots, u_n$ from uniform $(0,1)$, then $x_i = \left(-\frac{\ln(1-u_i)}{\lambda}\right)^K$; $i = 1 (1) n$ will be a random sample from the distribution of X.