STAT2602 Assignment 2

Cheng Ho Ming, Eric (3036216734) [Section 1A, 2024]

October 16, 2024

1. (i) The probability density function (p.d.f.) for the uniform distribution $U[-\alpha, 0]$ is:

$$f(x;\alpha) = \frac{1}{\alpha}\mathbb{I}_{[-\alpha,0]}(x),$$

where $\mathbb{I}_{[-\alpha,0]}(x)$ is the indicator function, which is defined as:

$$\mathbb{I}_{[-\alpha,0]}(x) = \begin{cases} 1 & \text{if } x \in [-\alpha, 0], \\ 0 & \text{otherwise.} \end{cases}$$

The likelihood function for the sample $X_1, X_2, \ldots, X_n$ is:

$$L(\alpha) = \prod_{i=1}^{n} \frac{1}{\alpha}\mathbb{I}_{[-\alpha,0]}(X_i).$$

This product implies that the likelihood is zero if any X_i lies outside the interval $[-\alpha, 0]$. Therefore, for the likelihood to be non-zero, all X_i must lie in $[-\alpha, 0]$, i.e., $-\alpha \leq X_i \leq 0$ for all i. Hence, the likelihood function becomes:

$$L(\alpha) = \frac{1}{\alpha^n}\mathbb{I}_{[-\alpha,0]}(\min(X_1, \ldots, X_n))\mathbb{I}_{[-\alpha,0]}(\max(X_1, \ldots, X_n)).$$

The log-likelihood function is:

$$\ell(\alpha) = -n\log(\alpha) + \log\left(\mathbb{I}_{[-\alpha,0]}(\min(X_1, \ldots, X_n))\right) + \log\left(\mathbb{I}_{[-\alpha,0]}(\max(X_1, \ldots, X_n))\right).$$

From the log-likelihood function, we can observe that it becomes larger when α is smaller.
Also, the indicator function must not return zero, as $\log(0)$ is undefined.
Therefore, $\min(X_1, \ldots, X_n) \geq -\alpha$ and $\max(X_1, \ldots, X_n) \leq 0$ must be satisfied when we maximizing the likelihood by finding the smallest value of α.
As a result:

$$\min(X_1, \ldots, X_n) \geq -\alpha \quad \Rightarrow \quad \alpha \geq -\min(X_1, \ldots, X_n).$$

Therefore, the MLE of α is:

$$\hat{\alpha} = -\min(X_1, \ldots, X_n).$$

(ii) The likelihood function is:

$$L(\alpha) = \frac{1}{\alpha^n}\mathbb{I}_{[-\alpha,0]}(\min(X_1, \ldots, X_n))\mathbb{I}_{(-a,0]}(\max(X_1, \ldots, X_n)).$$

By the **factorization theorem**, a sufficient statistic for α can be found by factorizing the likelihood function into two parts: one that depends on α and another that does not depend on α. Applying this theorem, we can write the likelihood function as:

$$g(T(X);\alpha) = \frac{1}{\alpha^n}\mathbb{I}_{[-\alpha,0]}(\min(X_1, \ldots, X_n)) \qquad h(X_1, \ldots, X_n) = \mathbb{I}_{(-a,0]}(\max(X_1, \ldots, X_n))$$

Therefore, the likelihood depends on α only through $\min(X_1, \ldots, X_n)$,
meaning that $\min(X_1, \ldots, X_n)$ is a sufficient statistic for α.

2. The MLE of θ is the value of θ that maximizes $f(x;\theta)$ for the observed x. For a given observation $x = x_{\text{obs}}$, we check the table to find which θ gives the largest probability.

- If $x_{\text{obs}} = 0$ or $x_{\text{obs}} = 1$, then the MLE is $\hat{\theta} = 1$ because $\frac{1}{3} > \frac{1}{4}$ and $f(x;3) = 0$.
- If $x_{\text{obs}} = 2$, then $\hat{\theta} = 2$ or $\hat{\theta} = 3$, because both give $f(x;\theta) = \frac{1}{4}$.
- If $x_{\text{obs}} = 3$, then $\hat{\theta} = 3$ because $\frac{1}{2}$ is the largest probability.
- If $x_{\text{obs}} = 4$, then $\hat{\theta} = 3$ because $\frac{1}{4}$ is the largest probability.

3. (i) The likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} \frac{\theta}{X_i^2} \mathbb{I}(0 < \theta \leq X_i < \infty)$$

The log-likelihood function is:

$$\begin{aligned}
\ell(\theta) &= \log\left(\frac{\theta^n \mathbb{I}(0 < \theta \leq X_1, \ldots, X_n < \infty)}{\prod_{i=1}^{n} X_i^2}\right) \\
&= n\log(\theta) + \log(\mathbb{I}(0 < \theta \leq \min(X_1, \ldots, X_n) < \infty)) - \log\left(\prod_{i=1}^{n} X_i^2\right) \\
&= n\log(\theta) + \log(\mathbb{I}(0 < \theta \leq \min(X_1, \ldots, X_n) < \infty)) - 2\sum_{i=1}^{n} \log(X_i)
\end{aligned}$$

Taking the gradient w.r.t. θ:

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} > 0 \quad (\text{for } 0 < \theta \leq \min(X_1, \ldots, X_n) < \infty)$$

Thus, the likelihood is increasing function w.r.t. θ.
The MLE is the maximum value of θ that satisfies the constraint $0 < \theta \leq \min(X_1, \ldots, X_n) < \infty$, which is $\hat{\theta} = \min(X_1, \ldots, X_n)$.

(ii) The expectation of $X_1^{1/3}$ is:

$$E(X_1^{1/3}) = \int_\theta^\infty x^{1/3} \frac{\theta}{x^2}\, dx = \int_\theta^\infty \theta x^{-5/3}\, dx = \theta \left[-\frac{3}{2} x^{-2/3}\right]_\theta^\infty = \theta(0 - (-\frac{3}{2}\theta^{-2/3})) = \frac{3}{2}\theta^{1/3}$$

(iii) Since the expectation of X:

$$E(X) = \int_\theta^\infty x\frac{\theta}{x^2}\, dx = \int_\theta^\infty \theta x^{-1}\, dx = \theta\left[-\log(x)\right]_\theta^\infty = \theta(-\log(\infty) - (-\log(\theta)))$$

diverges, we need to use $E(X_1^{1/3})$ for methods of moment estimator (MME).
From (ii), we have $E(X_1^{1/3}) = \frac{3}{2}\theta^{1/3}$. By equating $E(X_1^{1/3})$ to the 1/3-th sample moment of X $m_{1/3} = \frac{1}{n}\sum_{i=1}^{n} X_i^{1/3}$, we get:

$$\frac{3}{2}\theta^{1/3} = m_{1/3} \quad \Rightarrow \quad (\frac{3}{2})^3\theta = m_{1/3}^3 \quad \Rightarrow \quad \theta = (\frac{2}{3})^3 m_{1/3}^3 = \hat{\theta}_{MME}$$

As $n \to \infty$, the 1/3-th sample moment of X converges to $E(X_1^{1/3}) = \frac{3}{2}\theta^{1/3}$. Therefore,

$$\hat{\theta}_{MME} = (\frac{2}{3})^3 m_{1/3}^3 \to (\frac{2}{3})^3 E(X_1^{1/3})^3 = \frac{2}{3}^3 (\frac{3}{2}\theta^{1/3})^3 = \theta$$

Hence, $\hat{\theta}_{MME} \to_p \theta$.
$\therefore$ The MME is consistent.

4. (i) The likelihood function is:

$$L(p) = \prod_{i=1}^{n} p(1-p)^{X_i} = p^n(1-p)^{\sum_{i=1}^n X_i}.$$

By the factorization theorem, $T = \sum_{i=1}^n X_i$ is a sufficient statistic for p.
The p.d.f. can be rewritten as:

$$f(x;p) = \exp\left(x\ln(1-p) + \ln(p)\right).$$

Since the geometric distribution belongs to the exponential family and the parameter space $0 < p < 1$ is large enough, $T = \sum_{i=1}^n X_i$ is also complete for p.
Therefore, the statistic $T = \sum_{i=1}^n X_i$ is both sufficient and complete for p.

(ii) Given from (i), $T = \sum_{i=1}^n X_i$ is a complete and sufficient statistic for p.

$$\begin{aligned} E(X_1) &= \sum_{x=0}^{\infty} x\cdot p(1-p)^x = p\sum_{x=0}^{\infty} x(1-p)^x \\ &= p\sum_{x=0}^{\infty}(x+1-1)(1-p)^x = p\left(\sum_{x=0}^{\infty}(x+1)(1-p)^x - \sum_{x=0}^{\infty}(1)(1-p)^x\right) \\ &= p\left(\sum_{x=1}^{\infty} x(1-p)^{x-1} - \sum_{x=0}^{\infty}(1-p)^x\right) \\ &\text{Since } \sum_{x=1}^{\infty} xp(1-p)^{x-1} \text{ is the expectation of a geometric distribution which is } \frac{1}{p}, \\ &= p\left(\frac{1}{p^2} - \frac{1}{1-(1-p)}\right) = \frac{1}{p} - 1 = \frac{1-p}{p} \end{aligned}$$

As $X_1, X_2, \ldots, X_n$ are i.i.d., the expectation of T is:

$$\because E(T) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = n\cdot E(X_1) = n\cdot\frac{1-p}{p}$$

$$\therefore E(T) = n\cdot\left(\frac{1}{p} - 1\right) \Rightarrow \frac{E(T)}{n} + 1 = \frac{1}{p} \Rightarrow \frac{E(T)+n}{n} = \frac{1}{p} \Rightarrow \frac{n}{E(T)+n} = p$$

$\because E(\frac{n}{T+n}) = \frac{E(n)}{E(T)+E(n)} = \frac{n}{E(T)+n} = p$,
By Theorem 3.2 in the lecture notes, $\frac{n}{T+n}$ is the UMVUE of p.

5. (i) Since $X_i \sim N(\frac{p}{q}, \sigma_1^2)$ and $Y_i \sim N(q, \sigma_2^2)$, the expectation of T_1 is:

$$E(T_1) = \frac{1}{n}\sum_{i=1}^n E(X_iY_i) = \frac{1}{n}\sum_{i=1}^n \frac{p}{q}q = p$$

Therefore, T_1 is an unbiased estimator of p.

(ii) Since X and Y are independent,

$$\begin{aligned} Var(T_1) &= \frac{1}{n^2}\sum_{i=1}^n Var(X_iY_i) = \frac{1}{n^2}\sum_{i=1}^n \left(E(X_i^2)E(Y_i^2) - p^2\right) = \frac{1}{n}\left(\sigma_1^2 + \frac{p^2}{q^2}\right)(\sigma_2^2 + q^2) - \frac{p^2}{n} \\ &= \frac{1}{n}\left(\sigma_1^2\sigma_2^2 + \sigma_1^2 q^2 + \frac{p^2\sigma_2^2}{q^2}\right) \end{aligned}$$

(iii) Since X and Y are independent,

$$E(T_2) = E\left(\left(\frac{1}{n}\sum_{i=1}^n X_i\right)\left(\frac{1}{n}\sum_{i=1}^n Y_i\right)\right) = \left(\frac{1}{n}\sum_{i=1}^n E(X_i)\right)\left(\frac{1}{n}\sum_{i=1}^n E(Y_i)\right) = \frac{p}{q}\cdot q = p$$

Hence, T_2 is also an unbiased estimator of p.

(iv) By the weak law of large numbers,

$$\begin{aligned} E(T_2) &= E\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)\right) = E(\bar{X})E(\bar{Y}) \\ &\to_p \frac{p}{q}\cdot q = p \text{ as } n \to \infty \end{aligned}$$

$\therefore T_2$ is a consistent estimator of p.

(v) When $p = 0$ and $q^2 = \frac{\sigma_2^2}{n}$,

$$\begin{aligned} Var(T_1) &= \frac{1}{n}\left(\sigma_1^2\sigma_2^2 + \sigma_1^2 q^2 + \frac{p^2\sigma_2^2}{q^2}\right) = \frac{1}{n}\left(\sigma_1^2\sigma_2^2 + \sigma_1^2\frac{\sigma_2^2}{n}\right) = \frac{\sigma_1^2\sigma_2^2}{n} + \frac{\sigma_1^2\sigma_2^2}{n^2} \\ Var(T_2) &= Var(\bar{X}\bar{Y}) = E(\bar{X}^2)E(\bar{Y}^2) - 0 = (Var(\bar{X}) + E(\bar{X})^2)(Var(\bar{Y}) + E(\bar{Y})^2) \\ &\to_p (\sigma_1^2)(\sigma_2^2 + \frac{\sigma_2^2}{n}) = \sigma_1^2\sigma_2^2 + \frac{\sigma_1^2\sigma_2^2}{n} \end{aligned}$$

When $n = 1$, T_1 and T_2 has the same efficiency.
When $n > 1$, T_1 is more efficient than T_2.

6. (i) The expectation of $\bar{X}$ and $\frac{n}{n-1}S^2$ are:

$$\begin{aligned} E(\bar{X}) &= E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\cdot n\lambda = \lambda \\ E\left(\frac{n}{n-1}S^2\right) &= \frac{n}{n-1}E(S^2) = \frac{n}{n-1}E(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2) = \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n} E((X_i - \bar{X})^2) \\ &= \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i - \bar{X}) = \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i - \frac{X_1 + \cdots + X_n}{n}) \\ &= \frac{1}{n-1}\sum_{i=1}^{n} Var(\frac{n-1}{n}X_i - \sum_{j\neq i}\frac{1}{n}X_j) \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\left(\left(\frac{n-1}{n}\right)^2 Var(X_i) + \left(\frac{1}{n}\right)^2\sum_{j\neq i} Var(X_j)\right) \\ &= \frac{1}{n-1}\left(\left(\frac{n-1}{n}\right)^2 n\lambda + \left(\frac{1}{n}\right)^2 (n)(n-1)\lambda\right) \\ &= \frac{n-1}{n}\lambda + \frac{1}{n}\lambda = \lambda \end{aligned}$$

Both are unbiased estimators of λ.

(ii) Since $X \sim$ Poisson(λ) is from a exponential family, and the parameter $\lambda > 0$ contains an open set in $\mathbb{R}$,
the sufficient and complete statistic for λ would be $T = \sum_{i=1}^{n} X_i$.

(iii) The second derivative of the log-p.m.f. is:

$$\begin{aligned} f(x;\lambda) &= \frac{e^{-\lambda}\lambda^x}{x!} \\ \log(f(x;\lambda)) &= -\lambda + x\log(\lambda) - \log(x!) \\ \frac{\partial \log(f(x;\lambda))}{\partial\lambda} &= -1 + \frac{x}{\lambda} \\ \frac{\partial^2 \log(f(x;\lambda))}{\partial\lambda^2} &= -\frac{x}{\lambda^2} \end{aligned}$$

Therefore, the Fisher information about λ contained in data $X_1, \ldots, X_n$ is:

$$I_n(\lambda) = nI(\lambda) = -nE(\frac{\partial^2 log(f(X;\lambda))}{\partial\lambda^2}) = n\frac{E(X)}{\lambda^2} = \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda}$$

(iv) Using the Cramer-Rao Lower Bound (CRLB),

$$Var(\hat{\lambda}) \geq \frac{1}{I_n(\lambda)} = \frac{\lambda}{n}$$

(v) The variance of $\bar{X}$ and $\frac{n}{n-1}S^2$ are:

$$Var(\bar{X}) = Var(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} \cdot n\lambda = \frac{\lambda}{n}$$

$$\begin{aligned}
Var\left(\frac{n}{n-1}S^2\right) &= \frac{n^2}{(n-1)^2}Var(S^2) = \frac{n^2}{(n-1)^2}Var(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2) \\
&= \frac{n^2}{(n-1)^2}\frac{1}{n^2}\sum_{i=1}^{n} Var(X_i - \bar{X}) \\
&= \frac{1}{(n-1)^2}\sum_{i=1}^{n} Var(X_i - \bar{X}) = \frac{1}{(n-1)^2}\sum_{i=1}^{n} Var(X_i - \frac{X_1 + \cdots + X_n}{n}) \\
&= \frac{1}{(n-1)^2}\sum_{i=1}^{n} Var(\frac{n-1}{n}X_i - \sum_{j\neq i}\frac{1}{n}X_j) \\
&= \frac{1}{(n-1)^2}\sum_{i=1}^{n}\left(\left(\frac{n-1}{n}\right)^2 Var(X_i) + \left(\frac{1}{n}\right)^2\sum_{j\neq i} Var(X_j)\right) \\
&= \frac{1}{(n-1)^2}\left(\left(\frac{n-1}{n}\right)^2 n\lambda + \left(\frac{1}{n}\right)^2 (n)(n-1)\lambda\right) \\
&= \frac{1}{n}\lambda + \frac{n-1}{n}\lambda = \lambda
\end{aligned}$$

Since $Var(\bar{X}) < Var\left(\frac{n}{n-1}S^2\right)$, $\bar{X}$ is more efficient than $\frac{n}{n-1}S^2$.
$\because$ $\bar{X}$ also achieves the Cramer-Rao Lower Bound.
$\therefore$ $\bar{X}$ should be preferred as an estimator of λ.

7. (i) Let $\phi = (\theta, \theta^2)$. Denote the sample mean and sample variance as $\bar{X}$ and S^2 respectively. The joint p.d.f. of $X_1, X_2, \ldots, X_n \sim N(\theta, \theta^2)$ is:

$$\begin{aligned}
& f(x_1, x_2, \ldots, x_n; \phi) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta^2}} \exp\left(-\frac{(x_i - \theta)^2}{2\theta^2}\right) \\
&= \frac{1}{(2\pi\theta^2)^{n/2}} \exp\left(-\frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \theta)^2\right) \\
&= \frac{1}{(2\pi\theta^2)^{n/2}} \exp\left(-\frac{1}{2\theta^2}\left(\sum_{i=1}^{n} x_i^2 - 2\theta\sum_{i=1}^{n} x_i + n\theta^2\right)\right) \\
&= \frac{1}{(2\pi\theta^2)^{n/2}} \exp\left(-\frac{1}{2\theta^2}\left(\sum_{i=1}^{n} x_i^2 - 2\theta n\bar{X} + n\theta^2\right)\right) \\
&= \frac{1}{(2\pi\theta^2)^{n/2}} \exp\left(-\frac{1}{2\theta^2}(\sum_{i=1}^{n} x_i^2 - n\bar{X}^2 + n\bar{X}^2 - 2\theta n\bar{X} + n\theta^2)\right) \\
&= \frac{1}{(2\pi\theta^2)^{n/2}} \exp\left(-\frac{1}{2\theta^2}(\sum_{i=1}^{n} x_i^2 - n\bar{X}^2)\right) \exp\left(-\frac{1}{2\theta^2}(n\bar{X}^2 + n\theta^2 - 2\theta n\bar{X})\right) \\
&= \frac{1}{(2\pi\theta^2)^{n/2}} \exp\left(-\frac{n}{2\theta^2}(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{X}^2)\right) \exp\left(-\frac{n}{2\theta^2}(\bar{X} - \theta)^2\right) \\
&= \frac{1}{(2\pi\theta^2)^{n/2}} \exp\left(-\frac{n}{2\theta^2}S^2\right) \exp\left(-\frac{n}{2\theta^2}(\bar{X} - \theta)^2\right)
\end{aligned}$$

By the factorization theorem, $T = (\bar{X}, S^2)$ is a sufficient statistic for ϕ.

(ii) Let $\phi = (\theta, \theta^2) = (\theta, \xi)$ where $\xi = \theta^2$. The likelihood function based on the sample $X_1, X_2, \ldots, X_n$ is:

$$L(\phi) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\xi}} \exp\left(-\frac{(X_i - \theta)^2}{2\xi}\right)$$

Taking the logarithm of the likelihood:

$$\ell(\phi) = \log L(\phi) = -\frac{n}{2}\log(2\pi\xi) - \frac{1}{2\xi}\sum_{i=1}^{n}(X_i - \theta)^2$$

Taking the derivative of the log-likelihood with respect to θ:

$$\begin{aligned}\frac{\partial \ell}{\partial \theta} &= \frac{1}{2\xi}\sum_{i=1}^{n} 2(X_i - \theta) = \frac{1}{\xi}\sum_{i=1}^{n}(X_i - \theta) \\ &= \frac{1}{\xi}\left(\sum_{i=1}^{n} X_i - n\theta\right)\end{aligned}$$

Letting it to zero and solve for θ:

$$\begin{aligned}\frac{1}{\xi}\left(\sum_{i=1}^{n} X_i - n\theta\right) &= 0 \\ \sum_{i=1}^{n} X_i - n\theta &= 0 \\ \theta &= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}\end{aligned}$$

Given the distribution is normal, the MLE is indeed maximum.
Thus, the MLE of θ is:

$$\hat{\theta} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

(iii)

$$\begin{aligned}E(\bar{X}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\cdot n\theta = \theta \\ Var(\bar{X}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\cdot n\theta^2 = \frac{\theta^2}{n}\end{aligned}$$

By the Central Limit Theorem, as $n \to \infty$, $\bar{X} \sim_d N(\theta, \frac{\theta^2}{n})$.

8. (i) As $m \geq 2$, $\mu_2 = s_1$, $X_2 \sim N(s_1, \sigma_2^2)$ and thus $E(X_2) = s_1$.

$$\begin{aligned}\tilde{s_1} &= \frac{X_1 + 2X_2}{3} \\ E(\tilde{s_1}) &= \frac{E(X_1 + 2X_2)}{3} = \frac{E(X_1) + 2E(X_2)}{3} \\ &= \frac{s_1 + 2s_1}{3} = s_1\end{aligned}$$

As $E(\tilde{s_1}) = s_1$, $\tilde{s_1}$ is an unbiased estimator of s_1.

(ii) The likelihood function of $X_1, X_2, \ldots, X_{2m}$ is:

$$\begin{aligned}
L(s_1, s_2) &= \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(X_i - s_1)^2}{2\sigma_i^2}\right) \prod_{i=m+1}^{2m} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(X_i - s_2)^2}{2\sigma_i^2}\right) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \prod_{i=1}^{m} \exp\left(-\frac{(X_i - s_1)^2}{2\sigma_i^2}\right) \prod_{i=m+1}^{2m} \exp\left(-\frac{(X_i - s_2)^2}{2\sigma_i^2}\right)
\end{aligned}$$

Taking the logarithm of the likelihood:

$$\begin{aligned}
\ell(s_1, s_2) &= \log L(s_1, s_2) \\
&= \sum_{i=1}^{n} \log\left(\frac{1}{\sqrt{2\pi\sigma_i^2}}\right) - \sum_{i=1}^{m} \frac{(X_i - s_1)^2}{2\sigma_i^2} - \sum_{i=m+1}^{2m} \frac{(X_i - s_2)^2}{2\sigma_i^2}
\end{aligned}$$

Taking the derivative of the log-likelihood with respect to s_1 and s_2:

$$\frac{\partial \ell}{\partial s_1} = \sum_{i=1}^{m} \frac{X_i - s_1}{\sigma_i^2} \quad \text{and} \quad \frac{\partial \ell}{\partial s_2} = \sum_{i=m+1}^{2m} \frac{X_i - s_2}{\sigma_i^2}$$

Letting them to zero and solve for s_1 and s_2:

$$\begin{aligned}
\sum_{i=1}^{m} \frac{X_i - s_1}{\sigma_i^2} = 0 &\Rightarrow \sum_{i=1}^{m} \frac{X_i}{\sigma_i^2} = \sum_{i=1}^{m} \frac{s_1}{\sigma_i^2} \Rightarrow s_1 = \frac{\sum_{i=1}^{m} \frac{X_i}{\sigma_i^2}}{\sum_{i=1}^{m} \frac{1}{\sigma_i^2}} \\
\sum_{i=m+1}^{2m} \frac{X_i - s_2}{\sigma_i^2} = 0 &\Rightarrow \sum_{i=m+1}^{2m} \frac{X_i}{\sigma_i^2} = \sum_{i=m+1}^{2m} \frac{s_2}{\sigma_i^2} \Rightarrow s_2 = \frac{\sum_{i=m+1}^{2m} \frac{X_i}{\sigma_i^2}}{\sum_{i=m+1}^{2m} \frac{1}{\sigma_i^2}}
\end{aligned}$$

$$\therefore \hat{s_1} = \frac{\sum_{i=1}^{m} \frac{X_i}{\sigma_i^2}}{\sum_{i=1}^{m} \frac{1}{\sigma_i^2}} \quad \text{and} \quad \hat{s_2} = \frac{\sum_{i=m+1}^{2m} \frac{X_i}{\sigma_i^2}}{\sum_{i=m+1}^{2m} \frac{1}{\sigma_i^2}}$$

(iii) After setting $\sigma_i^2 = \frac{m}{i}$, the MLEs of s_1 is:

$$\hat{s_1} = \frac{\sum_{i=1}^{m} \frac{X_i}{\frac{m}{i}}}{\sum_{i=1}^{m} \frac{1}{\frac{m}{i}}} = \frac{\sum_{i=1}^{m} \frac{iX_i}{m}}{\sum_{i=1}^{m} \frac{i}{m}} = \frac{\sum_{i=1}^{m} iX_i}{\sum_{i=1}^{m} i} = \frac{\sum_{i=1}^{m} iX_i}{\frac{m(m+1)}{2}} = \frac{2}{m(m+1)} \sum_{i=1}^{m} iX_i$$

Variance of $\hat{s_1}$:

$$\begin{aligned}
Var(\hat{s_1}) &= Var\left(\frac{2}{m(m+1)} \sum_{i=1}^{m} iX_i\right) = \frac{4}{m^2(m+1)^2} \sum_{i=1}^{m} i^2 Var(X_i) = \frac{4}{m^2(m+1)^2} \sum_{i=1}^{m} i^2 \frac{m}{i} \\
&= \frac{4}{m(m+1)^2} \sum_{i=1}^{m} i = \frac{4}{m(m+1)^2} \cdot \frac{m(m+1)}{2} = \frac{2}{m+1}
\end{aligned}$$

Variance of $\tilde{s_1}$:

$$\begin{aligned}
Var(\tilde{s_1}) &= Var\left(\frac{X_1 + 2X_2}{3}\right) = \frac{1}{9} Var(X_1 + 2X_2) = \frac{1}{9}(Var(X_1) + 4Var(X_2)) \\
&= \frac{1}{9}(\sigma_1^2 + 4\sigma_2^2) = \frac{1}{9}\left(\frac{m}{1} + 4 \cdot \frac{m}{2}\right) = \frac{1}{9}(m + 2m) = \frac{3m}{9} = \frac{m}{3}
\end{aligned}$$

Comparing their variances:
At $m = 2$, $Var(\hat{s_1}) = \frac{2}{3}$ and $Var(\tilde{s_1}) = \frac{2}{3}$.
At $m > 2$, $Var(\hat{s_1}) = \frac{2}{m+1} < \frac{m}{3} = Var(\tilde{s_1})$.
$\therefore$ $\hat{s_1}$ is more efficient than $\tilde{s_1}$ for $m > 2$. Otherwise, they are equally efficient.

(iv) For $\hat{s_1}$, the asymptotic distribution of $\hat{s_1}$ is:

$$
\begin{aligned}
I(\hat{s_1}) &= \frac{1}{m} I_m(s_1) \\
&= \frac{1}{m} E\left[\left(\frac{d\ell(s_1)}{ds_1}\right)^2\right] = \frac{1}{m} E\left[\left(\sum_{i=1}^{m} \frac{i}{m(m+1)}(x_i - s_1)\right)^2\right] \\
&= \frac{1}{m} E\left[\left(\frac{2}{m(m+1)} \sum_{i=1}^{m} i(x_i - s_1)\right)^2\right] \\
&= \frac{1}{m^3}\left[\text{Var}\left(\sum_{i=1}^{m} i(x_i - s_1)\right) + E\left(\sum_{i=1}^{m} i(x_i - s_1)\right)^2\right] \\
&= \frac{1}{m^3}\left[\sum_{i=1}^{m} i^2 \text{Var}(x_i) + \sum_{i=1}^{m} i^2 E(x_i) E(x_i - s_1)^2\right] \\
&= \frac{1}{m^3}\left[\sum_{i=1}^{m} i^2 \frac{m}{i} + \sum_{i=1}^{m} i^2 E(x_i)\right] \\
&= \frac{\sum_{i=1}^{m} i}{m(m+1)} + o = \frac{m+1}{2m}
\end{aligned}
$$

Hence,

$$
\hat{s_1} \xrightarrow{d} N\left(s_1, \frac{2}{m+1}\right)
$$

Similarly,

$$
\begin{aligned}
I_n(\hat{s_2}) &= \frac{1}{m} I_m(s_2) \\
&= \frac{1}{m} E\left[\left(\sum_{i=m+1}^{2m} \frac{i}{m(m+1)}(x_i - s_2)\right)^2\right] \\
&= \frac{1}{m^3} \sum_{i=m+1}^{2m} i^2 \left(\text{Var}(x_i) + E(x_i)E(x_i - s_2)^2\right) \\
&= \frac{1}{m^3}\left[\sum_{i=m+1}^{2m} i^2 \frac{m}{i}\right] = \frac{3m+1}{2m}
\end{aligned}
$$

Hence,

$$
\hat{s_2} \xrightarrow{d} N\left(s_2, \frac{3m+1}{2m}\right)
$$

(v) As $m \to \infty$,

$$
\begin{aligned}
\lim_{m\to\infty}[E(\hat{s_1}) - s_1] &= \lim_{m\to\infty}\left[\frac{2}{m(m+1)} \sum_{i=1}^{m} iE(X_i) - s_1\right] = 0 \\
\lim_{m\to\infty}[E(\hat{s_2}) - s_2] &= \lim_{m\to\infty}\left[\frac{\sum_{i=m+1}^{2m} \frac{iE(X_i)}{m}}{\sum_{i=m+1}^{2m} \frac{i}{m}} - s_2\right] = 0
\end{aligned}
$$

$\therefore$ The MLE of s_1 and s_2 are (asymptotic) unbiased estimators of s_1 and s_2 respectively.
As $m \to \infty$, $Var(\hat{s_1}) = \frac{2}{m+1} \to 0$ and $Var(\hat{s_2}) = \frac{m}{3} \to 0$.
$\therefore$ The MLE of s_1 and s_2 are consistent estimators of s_1 and s_2 respectively.

9. (i) The likelihood function of θ based on $X_1, X_2, \ldots, X_n$ is:

$$L(\theta; X_1, \ldots, X_n) = \prod_{i=1}^{n} \frac{\theta}{X_i^{\theta+1}} = \theta^n \prod_{i=1}^{n} \frac{1}{X_i^{\theta+1}}.$$

(ii) For $x \geq 1$, the joint p.d.f. can be rewritten as:

$$\begin{aligned} f(x_1, x_2, \ldots, x_n; \theta) &= \prod_{i=1}^{n} \frac{\theta}{x_i^{\theta+1}} = \theta^n \prod_{i=1}^{n} x_i^{-(\theta+1)} \\ &= \theta^n e^{-(\theta+1)\sum_{i=1}^{n} \ln(x_i)} \\ &= g(T(x_1, x_2, \ldots, x_n), \theta)\ h(x_1, x_2, \ldots, x_n) \text{ where } h(\ldots) = 1 \end{aligned}$$

Using the **factorization theorem**, we got a scalar sufficient statistic T:

$$T = \sum_{i=1}^{n} \ln(X_i).$$

(iii) The Fisher information $I_n(\theta)$ is given by:

$$\begin{aligned} I_n(\theta) = nI(\theta) = -nE\left[\frac{\partial^2}{\partial\theta^2} \log f(X;\theta)\right] &= -nE\left[\frac{\partial^2}{\partial\theta^2} \log\left(\frac{\theta}{X^{\theta+1}}\right)\right] \\ = -nE\left[\frac{\partial^2}{\partial\theta^2}(\log(\theta) - (\theta+1)\log(X))\right] &= -nE\left[\frac{\partial}{\partial\theta}\left(\frac{1}{\theta} - \log(X)\right)\right] \\ = -nE\left[-\frac{1}{\theta^2}\right] = \frac{n}{\theta^2}. \end{aligned}$$

(iv) The Cramer-Rao Lower Bound (CRLB) for estimating θ is:

$$\text{CRLB} = \frac{1}{I_n(\theta)} = \frac{\theta^2}{n}.$$

(v) From the log-likelihood:

$$\log L(\theta) = n \log\theta - (\theta+1)\sum_{i=1}^{n} \log X_i$$

Taking the derivative with respect to θ and letting it to zero:

$$\frac{n}{\theta} - \sum_{i=1}^{n} \log X_i = 0 \quad \Rightarrow \quad \theta = \frac{n}{\sum_{i=1}^{n} \log X_i}$$

To confirm it is a MLE, we take the second derivative with respect to θ:

$$\frac{\partial^2 \log L(\theta)}{\partial\theta^2} = -\frac{n}{\theta^2} < 0 \text{ given } \theta > 0.$$

$\because$ The likelihood function is concave (open downward), the MLE is the maximum.
$\therefore$ The MLE of θ is $\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log X_i}$.

(vi) By Central Limit Theorem and Cramer-Rao Lower Bound, under regular conditions, the asymptotic distribution of the MLE is:

$$\frac{\hat{\theta} - \theta}{\sqrt{1/I_n(\theta)}} \sim N(0,1) \quad \Rightarrow \quad \hat{\theta} \sim N\left(\theta, \frac{\theta^2}{n}\right) \quad \text{as } n \to \infty.$$