STAT2602/STAT3902 Assignment 2

Q1) $X_i \overset{iid}{\sim} U[0, \beta]$.

(i) $\because$ $X_i$'s are i.i.d. with $E(X_i) = \frac{\beta}{2} < \infty$.

$\therefore$ By W.L.L.N., $\overline{X}_n \xrightarrow{P} E(X_i) = \frac{\beta}{2}$.

(ii) $E(X_i^2) = \int_0^\beta x^2 f_X(x)\,dx$

$$= \int_0^\beta x^2 \frac{1}{\beta}\,dx$$

$$= \frac{1}{\beta}\left[\frac{x^3}{3}\right]_0^\beta$$

$$= \frac{\beta^2}{3}$$

$\because$ $X_i^2$'s are i.i.d. with $E(X_i^2) = \frac{\beta^2}{3}$

$\therefore$ By W.L.L.N., $W_n = \frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{P} E(X_i^2) = \frac{\beta^2}{3}$. $\square$

Q2) (a) Let's prove by mathematical induction.

Define $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Then, when $n=1$, $\overline{X}_1 = \frac{1}{1} X_1 = X_1$

And $\overline{X}_1$ follows the Cauchy distribution.

Suppose $\overline{X}_k$ follows the Cauchy distribution.

Let's prove $\overline{X}_{k+1}$ also follows the Cauchy distribution.

$$\overline{X}_{k+1} = \frac{1}{k+1}\sum_{i=1}^{k+1} X_i = \frac{1}{k+1}\left[\sum_{i=1}^{k} X_i + X_{k+1}\right]$$

$$= \frac{1}{k+1}\sum_{i=1}^{k} X_i + \frac{1}{k+1} X_{k+1}$$

$$= \frac{k}{k+1}\,\frac{1}{k}\sum_{i=1}^{k} X_i + \frac{1}{k+1} X_{k+1}$$

$$= \frac{k}{k+1}\overline{X}_k + \frac{1}{k+1} X_{k+1}$$

$$\because \quad \frac{k}{k+1} + \frac{1}{k+1} = 1$$

$\therefore$ According to the theorem given by the question,

$\overline{X}_{k+1}$ follows the Cauchy distribution.

$\therefore$ $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ follows the Cauchy distribution.

Q2) (a) cont.

$\because$ $E(X_i)$ is not defined

$\therefore$ $\bar{X}_n$ does not converge in probability to anything.

(b) To apply the Law of large numbers (both weak & strong versions), we require $E(X_i) = \mu < \infty$, that is, the expectation of $X_i$ must be finite.

However, for Cauchy distribution, $E(X_i)$ is not defined

Note 1: $E(X_i) = \int_{-\infty}^{\infty} x \frac{1}{\pi(1+x^2)} dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x}{1+x^2} dx$

$$= \frac{1}{\pi} \frac{1}{2} \Big[ \ln(1+x^2) \Big]_{-\infty}^{\infty}$$

$\because$ $\lim_{x \to \infty} \ln(1+x^2)$ and $\lim_{x \to -\infty} \ln(1+x^2)$ do not exist

$\therefore$ $E(X_i)$ is undefined.

Q3) Let $X$ be the score of a given student.

$$X \sim N(53.6,\ 18.5^2)$$

(i)
$$P(X > 60) = P\left(Z > \frac{60 - 53.6}{18.5}\right)$$
$$\approx P(Z > 0.35)$$
$$= 1 - P(Z < 0.35)$$
$$= 1 - 0.6368$$
$$= 0.3632$$

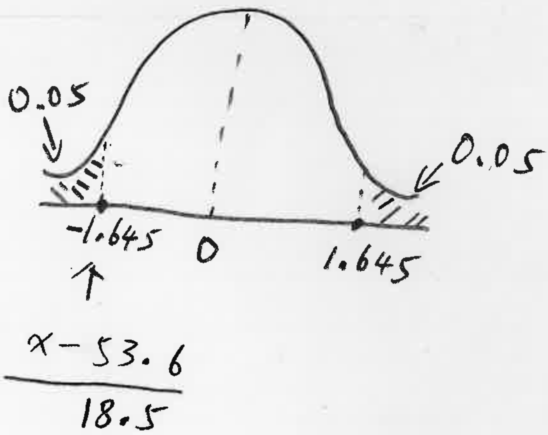
(ii)
$$X_i \overset{iid}{\sim} N(53.6,\ 18.5^2)$$
$$\bar{X} \sim N\left(53.6,\ \frac{18.5^2}{72}\right)$$
$$P(\bar{X} > 60) = P\left(Z > \frac{60 - 53.6}{\frac{18.5}{\sqrt{72}}}\right) \approx P(Z > 2.94)$$
$$= 1 - P(Z < 2.94)$$
$$= 1 - 0.99836$$
$$= 0.00164$$

Q3) (iii) $95\% = P(X > x) = P\left(Z > \frac{x - 53.6}{18.5}\right)$

$\therefore \frac{x - 53.6}{18.5} = -1.645$

$\therefore x = -1.645 \times 18.5 + 53.6$

$= 23.1675$




(iv) If the normal assumption is violated, then the results of (i) & (iii) are incorrect. However, the result of (ii) still holds because the sample size is large (i.e. $n = 60 > 30$) and the Central Limit Theorem can be applied.

Q4) The MLE of $\theta$ depends on the actual observation of $X$.

$$\hat{\theta}_{MLE} = \begin{cases} 1 & \text{if } x = 0 \\ 1 & \text{if } x = 1 \\ 2 \text{ or } 3 & \text{if } x = 2 \\ 3 & \text{if } x = 3 \\ 3 & \text{if } x = 4 \end{cases}$$

Q5) $X_i \overset{iid}{\sim} Poi(\lambda)$, $E(X_i) = Var(X_i) = \lambda$.

(i) $E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$

$\overset{\text{identical}}{=} \frac{1}{n}\sum_{i=1}^{n}\lambda = \frac{1}{n} n\lambda = \lambda$.

$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$

$$E(S^2) = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2\right]$$

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

$$= \frac{1}{n}\sum_{i=1}^{n}\left[Var(X_i) + E(X_i)^2\right] - \left[Var(\bar{X}) + E(\bar{X})^2\right]$$

$$\overset{iid}{=} \frac{1}{n}\sum_{i=1}^{n}\left[\lambda + \lambda^2\right] - \left[\frac{Var(X_i)}{n} + \lambda^2\right]$$

$$= \frac{1}{n} n\left[\lambda + \lambda^2\right] - \left[\frac{\lambda}{n} + \lambda^2\right]$$

$$= \lambda\left(\frac{n-1}{n}\right)$$

$\therefore$ $E\left(\frac{n}{n-1} S^2\right) = \lambda$. □

Q5) (ii) $f(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}$.

$$\ln f(x;\lambda) = -\lambda + x\ln\lambda - \ln(x!)$$

$$\frac{d}{d\lambda}\ln f(x;\lambda) = -1 + \frac{x}{\lambda}$$

$$\frac{d^2}{d\lambda^2}\ln f(x;\lambda) = -\frac{x}{\lambda^2}$$

$$\therefore I(\theta) = -E\left[\frac{d^2}{d\lambda^2}\ln f(x;\lambda)\right]$$

$$= -E\left[-\frac{X}{\lambda^2}\right]$$

$$= \frac{1}{\lambda^2}E(X)$$

$$= \frac{\lambda}{\lambda^2}$$

$$= \frac{1}{\lambda}$$

$$\therefore CRLB = \frac{1}{I_n(\theta)} = \frac{1}{nI(\theta)} = \frac{\lambda}{n}.$$

(iii) $Var(\overline{X}) = \frac{Var(X_i)}{n} = \frac{\lambda}{n} = CRLB$.

$\because E(\overline{X}) = \lambda$ and $Var(\overline{X}) = CRLB$

$\therefore \overline{X}$ is the UMVUE

$\therefore \overline{X}$ is preferred.

Q6) (i) $L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\left[\frac{\theta}{x_i^2}\right] = \frac{\theta^n}{\prod_{i=1}^{n} x_i^2}$ if $x_{min} \geq \theta$.

$$\frac{dL}{d\theta} = \frac{n\theta^{n-1}}{\prod_{i=1}^{n} x_i^2}, \quad \forall \theta > 0.$$

$\therefore$ As $\theta \uparrow$, $L(\theta)$ always $\uparrow$.

$\because \theta \leq x_{min}$

$\therefore \hat{\theta}_{MLE} = x_{min}$.

(ii) $E\left(X_1^{\frac{1}{3}}\right) \overset{iid}{=} E\left(X_i^{\frac{1}{3}}\right) = \int_{\theta}^{\infty} x^{\frac{1}{3}} \frac{\theta}{x^2} dx = \int_{\theta}^{\infty} \theta x^{-\frac{5}{3}} dx$

$$= \theta \left[x^{-\frac{2}{3}}\right]_{\theta}^{\infty}\left(-\frac{3}{2}\right)$$

$$= -\frac{3}{2}\theta\left[0 - \theta^{-\frac{2}{3}}\right]$$

$$= \frac{3}{2}\theta^{\frac{1}{3}}$$

(iii) $\frac{1}{n}\sum_{i=1}^{n} x_i^{\frac{1}{3}} = E\left(X^{\frac{1}{3}}\right) = \frac{3}{2}\theta^{\frac{1}{3}}$

$$\Longrightarrow \theta^{\frac{1}{3}} = \frac{2}{3}\frac{1}{n}\sum_{i=1}^{n} x_i^{\frac{1}{3}}$$

$$\Longrightarrow \theta = \frac{8}{27}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^{\frac{1}{3}}\right)^3$$

$$\therefore \hat{\theta}_{MME} = \frac{8}{27}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^{\frac{1}{3}}\right)^3$$

Q6) (iii) cont.

$\because$ $X_i^{\frac{1}{3}}$'s are i.i.d. with $E(X_i^{\frac{1}{3}}) = \frac{3}{2}\theta^{\frac{1}{3}} < \infty$

$\therefore$ By W.L.L.N., $\frac{1}{n}\sum_{i=1}^{n} X_i^{\frac{1}{3}} \xrightarrow{P} \frac{3}{2}\theta^{\frac{1}{3}}$.

$\because$ $g(y) = \frac{8}{27}y^3$ is a continuous function

and $\frac{1}{n}\sum_{i=1}^{n} X_i^{\frac{1}{3}} \xrightarrow{P} \frac{3}{2}\theta^{\frac{1}{3}}$

$\therefore$ By continuous mapping theorem,

$$\hat{\theta}_{MME} = g\left(\frac{1}{n}\sum_{i=1}^{n} X_i^{\frac{1}{3}}\right) \xrightarrow{P} g\left(\frac{3}{2}\theta^{\frac{1}{3}}\right) = \frac{8}{27}\left(\frac{3}{2}\theta^{\frac{1}{3}}\right)^3 = \theta \quad \square$$

Q7) (i) $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$$

$$\therefore E\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right) = \sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)$$

$$= n\left[Var(X_i) + E(X_i)^2\right] - n\left[Var(\bar{X}) + E(\bar{X})^2\right]$$

$$= n\left[\sigma^2 + \mu^2\right] - n\left[\frac{\sigma^2}{n} + \mu^2\right]$$

$$= (n-1)\sigma^2$$

$$\because S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$$

$$\therefore E\left(\frac{n}{n-1}S^2\right) = E\left(\frac{\Sigma(X_i-\bar{X})^2}{n-1}\right) = \frac{1}{n-1}E\left(\Sigma(X_i-\bar{X})^2\right)$$

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

(ii) $$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

$$\ln f(x) = -\frac{1}{2}\ln(\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2} - \frac{1}{2}\ln 2\pi$$

$$\frac{d}{d(\sigma^2)}\ln f(x) = -\frac{1}{2}\frac{1}{\sigma^2} + \frac{(x-\mu)^2}{2(\sigma^2)^2}$$

$$\frac{d^2}{d(\sigma^2)^2}\ln f(x) = \frac{1}{2(\sigma^2)^2} - \frac{(x-\mu)^2}{(\sigma^2)^3}$$

$$\therefore I(\theta) = -E\left(\frac{d^2}{d(\sigma^2)^2}\ln f(x)\right) = \frac{E[(X-\mu)^2]}{(\sigma^2)^3} - \frac{1}{2(\sigma^2)^2}$$

$$= \frac{\sigma^2}{(\sigma^2)^3} - \frac{1}{2(\sigma^2)^2} = \frac{1}{2(\sigma^2)^2}$$

Q7) (ii) cont.

$\therefore$ $CRLB = \frac{1}{nI(\theta)} = \frac{2(\sigma^2)^2}{n}$

(iii) $\because$ $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$

$\therefore$ $\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$

$\therefore$ $Var\left(\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2}\right) = 2(n-1)$

$\therefore$ $Var\left(\frac{n}{n-1} S^2\right) = Var\left(\frac{n}{n-1} \frac{\sum(X_i - \bar{X})^2}{n}\right)$

$= Var\left(\frac{\sum(X_i - \bar{X})^2}{n-1}\right) = \frac{1}{(n-1)^2} Var\left(\sum(X_i - \bar{X})^2\right)$

$= \frac{(\sigma^2)^2}{(n-1)^2} Var\left(\frac{\sum(X_i - \bar{X})^2}{\sigma^2}\right) = \frac{(\sigma^2)^2}{(n-1)^2} 2(n-1)$

$= \frac{2(\sigma^2)^2}{n-1} > CRLB$

Q8) (i) $X_i \overset{iid}{\sim} N(\theta, \theta^2)$, $\theta \neq 0$.

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

$$= (2\pi\theta^2)^{-\frac{n}{2}} \exp\left\{-\frac{\Sigma(x_i-\theta)^2}{2\theta^2}\right\}$$

$$= (2\pi)^{-\frac{n}{2}} |\theta|^{-n} \exp\left\{-\frac{\Sigma x_i^2 - 2\Sigma x_i\theta + n\theta^2}{2\theta^2}\right\}$$

$$= |\theta|^{-n} \exp\left\{-\frac{\Sigma x_i^2 - 2\Sigma x_i\theta}{2\theta^2}\right\} \cdot (2\pi)^{-\frac{n}{2}} e^{-\frac{n}{2}}$$

$$= g(T_1, T_2; \theta)\, h(\vec{x}^T)$$

where $g(T_1, T_2; \theta) = |\theta|^{-n} \exp\left\{-\frac{\Sigma x_i^2 - 2\Sigma x_i\theta}{2\theta^2}\right\}$, $T_1 = \Sigma x_i^2$, $T_2 = \Sigma x_i$

and $h(\vec{x}^T) = (2\pi)^{-\frac{n}{2}} e^{-\frac{n}{2}}$

$\because$ $h(\vec{x}^T)$ does not depend on $\theta$, $g(T_1, T_2; \theta)$ is a function of $T_1, T_2$ and $\theta$, and it depends on $x_1, x_2, \ldots, x_n$ only through $T_1$ and $T_2$.

$\therefore$ By Factorization Theorem, $(T_1, T_2)$ are jointly sufficient for $\theta$.

However, we can prove that $\{T_1 = \sum_{i=1}^{n} X_i^2 \;, \; T_2 = \sum_{i=1}^{n} X_i\}$ are not complete for the estimation of $\theta$.

Define $g(T_1, T_2) = 2S^2 - \frac{1}{n}\sum_{i=1}^{n} X_i^2$ where $S^2 = \dfrac{\sum(X_i - \bar{X})^2}{n-1}$

$$= \frac{\sum X_i^2 - n\bar{X}^2}{n-1}$$

Then, $E(g(T_1, T_2)) = 2E(S^2) - \frac{1}{n} n E(X_i^2)$

$$= 2\theta^2 - [Var(X_i) + E(X_i)^2]$$

$$= 2\theta^2 - [\theta^2 + \theta^2]$$

$$= 0$$

for all $\theta \in \Theta$

However, it is possible that $2S^2 - \frac{1}{n}\sum_{i=1}^{n} X_i^2 \neq 0$,

that is, $Pr(g(T_1, T_2) = 0) \neq 1$

Hence, $\{\sum_{i=1}^{n} X_i^2 \;, \; \sum_{i=1}^{n} X_i\}$ are <u>not</u> complete for the estimation of $\theta$.

Theorem: If the <u>minimum</u> sufficient statistics is not complete, then the complete statistics does not exist.

Note: You will also study this theorem in Stat3602.

$\because \{\sum_{i=1}^{n} X_i^2 \;, \; \sum_{i=1}^{n} X_i\}$ are the minimum sufficient statistics for $\theta$ and <u>not</u> complete for $\theta$.

$\therefore$ The complete statistics does not exist.

Q8) (ii) $L(\theta) \overset{iid}{=} \prod_{i=1}^{n} f(x_i) = (2\pi\theta^2)^{-\frac{n}{2}} \exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2\theta^2}\right\}$

$l(\theta) = \ln L(\theta) = -n\ln\theta - \frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2\theta^2} + C$   Note: $C \in$ Constant

$= -n\ln\theta - \frac{\sum x_i^2 - 2\sum x_i\theta + n\theta^2}{2\theta^2} + C$

$= -n\ln\theta - \frac{\sum x_i^2}{2\theta^2} + \frac{\sum x_i}{\theta} - \frac{n}{2} + C$

$\frac{dl}{d\theta} = -\frac{n}{\theta} + \frac{\sum x_i^2}{\theta^3} - \frac{\sum x_i}{\theta^2} = 0$

$\Rightarrow -n\theta^2 + \sum x_i^2 - \sum x_i\theta = 0$

$\Rightarrow \theta^2 + \bar{x}\theta - \overline{x^2} = 0$, where $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

$\therefore \theta_1 = \frac{-\bar{x} + \sqrt{\bar{x}^2 + 4\overline{x^2}}}{2}$ and $\theta_2 = \frac{-\bar{x} - \sqrt{\bar{x}^2 + 4\overline{x^2}}}{2}$

It is possible to show that the 2nd derivative of $l(\theta)$ is negative when $\theta = \theta_1$ or $\theta_2$.

$$\therefore \hat{\theta}_{MLE} = \begin{cases} \theta_1 & \text{if } l(\theta_1) > l(\theta_2) \\ \theta_1 \text{ and } \theta_2 & \text{if } l(\theta_1) = l(\theta_2) \\ \theta_2 & \text{if } l(\theta_1) < l(\theta_2) \end{cases}$$

Q8) (iii) $\ln f(x) = -\ln\theta - \frac{(x-\theta)^2}{2\theta^2} + \ln\sqrt{2\pi}$

$$= -\ln\theta - \frac{x^2 - 2\theta x + \theta^2}{2\theta^2} + \ln\sqrt{2\pi}$$

$$= -\ln\theta - \frac{x^2}{2\theta^2} + \frac{x}{\theta} - \frac{1}{2} + \ln\sqrt{2\pi}$$

$$\frac{d}{d\theta}\ln f(x) = -\frac{1}{\theta} + \frac{x^2}{\theta^3} - \frac{x}{\theta^2}$$

$$\frac{d^2}{d\theta^2}\ln f(x) = \frac{1}{\theta^2} - \frac{3x^2}{\theta^4} + \frac{2x}{\theta^3}$$

$\therefore$ $I(\theta) = -E\left[\frac{d^2}{d\theta^2}\ln f(X)\right] = -E\left[\frac{1}{\theta^2} - \frac{3X^2}{\theta^4} + \frac{2X}{\theta^3}\right]$

$$= -\frac{1}{\theta^2} + \frac{3}{\theta^4}(\theta^2+\theta^2) - \frac{2}{\theta^3}\theta$$

$$= -\frac{1}{\theta^2} + \frac{3}{\theta^4}2\theta^2 - \frac{2}{\theta^2}$$

$$= -\frac{3}{\theta^2} + \frac{6}{\theta^2}$$

$$= \frac{3}{\theta^2}$$

$\because$ $\tau(\theta) = \theta$ and $\tau'(\theta) = 1$

$\therefore$ $CRLB = \frac{[\tau'(\theta)]^2}{nI(\theta)} = \frac{1}{nI(\theta)} = \frac{\theta^2}{3n}$

$\therefore$ By the asymptotic property of MLE,

$$\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N\left(0, \frac{\theta^2}{3}\right)$$

Note: $(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N\left(0, \frac{\theta^2}{3n}\right)$ is not wrong. But this notation is not good because the variance goes to zero as $n$ goes to infinity.

Q9) Solution:

(i)

Let $X$ be number of failures until and up to the first success. Then, $X \sim \text{Geo}(p)$, $p$ is the success probability.

$$\Pr(X = x) = p(1-p)^x, \qquad \text{for } x = 0, 1, 2, \ldots, \text{ and } 0 < \theta < 1.$$

Note: This is an alternative definition of Geometric distribution and the support of $X$ is $S_X = \{0, 1, 2, \ldots\}$.

Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. Geo($p$).

Let $T = \sum_{i=1}^{n} X_i$,

$$\begin{aligned} \Pr(X_1 = x_1, \ldots, X_n = x_n) &= \prod_{i=1}^{n} \Pr(X_i = x_i) = \prod_{i=1}^{n} \left[p\,(1-p)^{x_i}\right] \\ &= p^n (1-p)^{\sum_{i=1}^{n} x_i} \times 1 = g\,(T; p) \times h(\overrightarrow{x}^{\top}), \end{aligned}$$

where $g\,(T; p) = p^n (1-p)^{\sum_{i=1}^{n} x_i}$ and $h(\overrightarrow{x}^{\top}) = 1$.

$\therefore$ By factorization theorem,

$T = \sum_{i=1}^{n} x_i$ is a sufficient statistic of $p$.

Now, let's prove the completeness of $T$.

Note: To prove completeness, you can also use Property 3.4 of Lecture 3. I have used the following method here because I want you to have a deeper understanding on the definition of completeness.

Since $X_1, X_2, \ldots, X_n$ are i.i.d. Geometric random variables, their sum must follow Negative Binomial distribution. Geo($p$), it is possible show that using mgf.

Note: You can prove this using mgf.

$\therefore T \sim \text{NB}(n, p)$.

Note: $T$ is defined as the number of failures until and up to the $n$-th successes, as we have used the alternative defintions of Negative Binomial & Geometric distributions.

Then, the pmf of $T$ is

$$\Pr(T = t) = \binom{n+t-1}{t} p^n (1-p)^t, \qquad \text{for } y = 0, 1, 2, 3, \ldots \text{ and } 0 < p < 1.$$

Let $h(T)$ be any function of $T$ s.t. $\mathrm{E}\,(h(T)) = 0$ for $\forall p \in \Theta$.

Let's prove $h(T) = 0$ for $\forall T$, i.e. $\Pr\,(h(T) = 0) = 1$.

Note: $T \sim \text{NB}(n, p)$.

$$\begin{aligned} 0 &= \mathrm{E}\,(h(T)) = \sum_{t=0}^{\infty} h(t) \Pr(T = t) = \sum_{t=0}^{\infty} h(t) \binom{n+t-1}{t} \theta^n (1-\theta)^t \\ \overset{\theta \neq 0}{\Longrightarrow} 0 &= \sum_{t=0}^{\infty} C_t u^t, \qquad \text{where } C_t = h(t) \binom{n+t-1}{t} \text{ and } u = (1-\theta). \end{aligned}$$

$\therefore$ We have $0 = \sum_{t=0}^{\infty} C_t u^t$ for $\forall u \in (0, 1)$.

**Thereom**: If $0 = \sum_{t=0}^{\infty} C_t u^t$, for all $u \in (0,1)$, then

$$C_t = 0, \forall t \in \{0,1,2,3,\ldots\}.$$

By theorem above, $C_t = 0, \forall t \in \{0,1,2,\ldots\}$.
$\because$

$$\binom{n+t-1}{t} \neq 0, \forall t \in \{0,1,2,\ldots\}.$$

$\therefore h(t) = 0, \qquad \text{for } \forall t \in \{0,1,2,3,\ldots\}.$
$\therefore$ We have the following result:

$$\mathrm{E}(h(T)) = 0, \ \forall p \in \Theta \Longrightarrow h(T) = 0, \forall T.$$

$\therefore T$ is complete for $p$.

(ii)
Now, let's use Lehmann-Scheffé Theorem to find the UMVUE of $\theta$.
First of all, noticed that $\frac{T}{n}$ is not an unbiased estimator of $\theta$ because we have used the alternative definitions of Geometric & Negative Binomial distributions.

$$\mathrm{E}(T) = n\mathrm{E}(X_i) = \frac{n(1-p)}{p},$$

and

$$\mathrm{E}\left(\frac{T}{n}\right) = \frac{(1-p)}{p}.$$

Now, let's find an unbiased estimator of $\theta$.
Define

$$I_{(0)}(X_j) = \begin{cases} 1, & \text{if } X_j = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned} \mathrm{E}\left(I_{(0)}(X_j)\right) &= 1 \Pr(X_j = 0) + 0 \Pr(X_j \neq 0) \\ &= 1 \cdot p \\ &= p. \end{aligned}$$

$\therefore I_{(0)}(X_j)$ is an unbiased estimator of $p$.
Let $g(T) = \mathrm{E}\left(I_{(0)}(X_j)|T\right)$.
Then,

$$\begin{aligned} \mathrm{E}(g(T)) &= \mathrm{E}_T\left[\mathrm{E}\left(I_{(0)}(X_j)|T\right)\right] \\ &= \mathrm{E}\left(I_{(0)}(X_j)\right) \\ &= p. \end{aligned}$$

$\therefore g(T)$ is an unbiased estimator of $p$.
$\because T$ is sufficient and complete for $p$.

$\therefore$ $g(T)$ is unique UMVUE for $p$.

Now, let's work out the explicit expression of $g(T)$.

$$\begin{aligned} g(T) &= \mathrm{E}\left(I_{(0)}(X_j|T)\right) \\ &= 1 \cdot \Pr\left(X_j = 0|T\right) + 0 \Pr\left(X_j \neq 0|T\right) \\ &= \Pr(X_j = 0|T) \qquad \ldots\ldots(*) \end{aligned}$$

Let $T' = \sum_{\forall i: i\neq j} X_i$, then from part (b), we know

$$T \sim \mathrm{NB}(n, p), \text{ and } T' \sim \mathrm{NB}(n-1, p)$$

$\because$

$$\begin{aligned} \Pr(X_j = 0|T = t) &= \frac{\Pr\left(X_j = 0 \cap T = t\right)}{\Pr(T = t)} \\ &= \frac{\Pr\left(T = t|X_j = 0\right)\Pr\left(X_j = 0\right)}{\Pr(T = t)} \\ &= \frac{\Pr\left(T' = t\right)\Pr\left(X_j = 0\right)}{\Pr(T = t)} \\ &= \frac{\binom{n-1+t-1}{t} p^{n-1}(1-p)^t p}{\binom{n+t-1}{t} p^n (1-p)^t} \\ &= \frac{(n+t-2)!}{(n-2)!t!} \div \left[\frac{(n+t-1)!}{(n-1)!t!}\right] \\ &= \frac{n-1}{n+t-1} \\ &= \frac{n-1}{t+n-1}. \end{aligned}$$

$\therefore$ By (*), $g(T) = \Pr\left(X_j = 0|T\right) = \frac{n-1}{T+n-1} = \frac{n-1}{\sum_{i=1}^n X_i + n-1}$.

$\because$ $T$ is complete and sufficient for $p$ and $g(T)$ is unbiased for $p$.

$\therefore$By Lehmann-Scheffé Theorem, $g(T) = \frac{n-1}{T+n-1} = \frac{n-1}{\sum_{i=1}^n X_i + n-1}$ is the UMVUE of $p$.

---

Note:

$\because$ $T = T' + X_j$.

$\therefore$ $\Pr\left(T = t|X_j = 0\right) = \Pr\left(T' = t\right)$.

Q10) $L(s_1, s_2) = f(x_1, \ldots, x_n) = f(x_1, \ldots, x_m)\, f(x_{m+1}, \ldots, x_n)$

$$= \prod_{i=1}^{m} \left[ \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i - s_1)^2}{2\sigma_i^2}} \right] \prod_{i=m+1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(y_i - s_2)^2}{2\sigma_i^2}} \right]$$

$$= \left( \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_i^2}} \right) \exp\left\{ -\sum_{i=1}^{m} \frac{(x_i - s_1)^2}{2\sigma_i^2} \right\} \left( \prod_{i=m+1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \right) \exp\left\{ -\sum_{i=m+1}^{n} \frac{(x_i - s_2)^2}{2\sigma_i^2} \right\}$$

$\therefore\ l(s_1, s_2) = \ln L(s_1, s_2)$

$$= -\sum_{i=1}^{m} \frac{(x_i - s_1)^2}{2\sigma_i^2} - \sum_{i=m+1}^{n} \frac{(x_i - s_2)^2}{2\sigma_i^2} + c,$$

where $c = \ln \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_i^2}} + \ln \prod_{i=m+1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}}$.

$$\frac{dl}{ds_1} = \sum \frac{2(x_i - s_1)}{2\sigma_i^2} = 0 \Rightarrow \sum \frac{x_i}{\sigma_i^2} - \sum \frac{s_1}{\sigma_i^2} = 0$$

$$\Rightarrow \sum \frac{x_i}{\sigma_i^2} = \sum \frac{1}{\sigma_i^2} s_1 \Rightarrow s_1 = \left( \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^{m} \frac{x_i}{\sigma_i^2}$$

$$\because\ \frac{d^2 l}{ds_1^2} = -\sum_{i=1}^{m} \frac{1}{\sigma_i^2} < 0$$

$\therefore\ \hat{s}_{1MLE} = \left( \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^{m} \frac{x_i}{\sigma_i^2}$, Similarly, we can show that,

$$\hat{s}_{2MLE} = \left( \sum_{i=m+1}^{n} \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=m+1}^{n} \frac{x_i}{\sigma_i^2}$$

Q10) (ii) $\because \sigma_i^2 = \frac{m}{i}, \forall i$

$$\therefore \hat{S}_1 = \left(\sum_{i=1}^{m} \frac{1}{\frac{m}{i}}\right)^{-1} \sum_{i=1}^{m} \frac{X_i}{\frac{m}{i}}$$

$$= \left(\sum_{i=1}^{m} \frac{i}{m}\right)^{-1} \sum_{i=1}^{m} \frac{iX_i}{m}$$

$$= \left(\sum_{i=1}^{m} i\right)^{-1} \sum_{i=1}^{m} iX_i$$

$$= \left[\frac{(1+m)m}{2}\right]^{-1} \sum_{i=1}^{m} iX_i$$

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

$$\therefore Var(\hat{S}_1) = \frac{4}{(1+m)^2 m^2} \sum_{i=1}^{m} i^2 Var(X_i)$$

$$= \frac{4}{(1+m)^2 m^2} \sum_{i=1}^{m} i^2 \frac{m}{i} = \frac{4}{(1+m)^2 m} \sum_{i=1}^{m} i$$

$$= \frac{4}{(1+m)^2 m} \frac{(1+m)m}{2} = \frac{2}{1+m}$$

Similarly, $\hat{S}_2 = \left(\sum_{i=m+1}^{2m} i\right)^{-1} \sum_{i=m+1}^{2m} iX_i$

$$= \left[\frac{(m+1+2m)m}{2}\right]^{-1} \sum_{i=m+1}^{2m} iX_i$$

$$= \left[\frac{2}{(3m+1)m}\right] \sum_{i=m+1}^{2m} iX_i$$

$$\therefore Var(\hat{S}_2) = \frac{4}{(3m+1)^2 m^2} \sum_{i=m+1}^{2m} i^2 Var(X_i)$$

$$= \frac{4}{(3m+1)^2 m^2} \sum_{i=m+1}^{2m} i^2 \frac{m}{i} = \frac{4}{(3m+1)^2 m} \sum_{i=m+1}^{2m} i$$

$$= \frac{4}{(3m+1)^2 m} \frac{(m+1+2m)m}{2} = \frac{2}{3m+1}$$

Q10) (ii) $\because \sigma_i^2 = \frac{m}{i}, \forall i$

$$\therefore \hat{S}_{1MLE} = \left(\sum_{i=1}^{m} \frac{1}{\frac{m}{i}}\right)^{-1} \sum_{i=1}^{m} \frac{X_i}{\frac{m}{i}}$$

$$= \left(\sum_{i=1}^{m} \frac{i}{m}\right)^{-1} \sum_{i=1}^{m} \frac{iX_i}{m}$$

$$= \left(\sum_{i=1}^{m} i\right)^{-1} \sum_{i=1}^{m} iX_i$$

$$\therefore E(\hat{S}_{1MLE}) = E\left[(\Sigma i)^{-1} \Sigma iX_i\right] = (\Sigma i)^{-1} \Sigma i E(X_i)$$

$$= S_1 (\Sigma i)^{-1} (\Sigma i) = S_1$$

Now, let's prove $Var(\hat{S}_1) \to 0$, as $n \to \infty$.

$$\hat{S}_{1MLE} = \left(\sum_{i=1}^{m} i\right)^{-1} \sum_{i=1}^{m} iX_i = \left[\frac{(1+m)m}{2}\right]^{-1} \sum_{i=1}^{m} iX_i$$

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

$$\therefore Var(\hat{S}_{1MLE}) = \frac{4}{(1+m)^2 m^2} \sum_{i=1}^{m} i^2 Var(X_i)$$

$$= \frac{4}{(1+m)^2 m^2} \sum_{i=1}^{m} i^2 \frac{m}{i} = \frac{4}{(1+m)^2 m} \sum_{i=1}^{m} i$$

$$= \frac{4}{(1+m)^2 m} \frac{(1+m)m}{2} = \frac{2}{1+m} \to 0 \quad \text{as} \quad m \to \infty$$

$\therefore \hat{S}_{1MLE}$ is consistent.

Q10) (ii) cont. For $\hat{S}_{2MLE}$

$$\hat{S}_{2MLE} = \left(\sum_{i=m+1}^{2m} \frac{1}{\sigma_i^2}\right)^{-1} \sum_{i=m+1}^{2m} \frac{X_i}{\sigma_i^2}$$

$$= \left(\sum_{i=m+1}^{2m} \frac{i}{m}\right)^{-1} \sum_{i=m+1}^{2m} \frac{iX_i}{m}$$

$$= \left(\sum_{i=m+1}^{2m} i\right)^{-1} \sum_{i=m+1}^{2m} iX_i$$

$$\therefore E(\hat{S}_{2MLE}) = \left(\sum i\right)^{-1} \sum i E(X_i)$$

$$= S_2 \left(\sum_{i=m+1}^{2m} i\right)^{-1} \sum_{i=m+1}^{2m} i = S_2 \quad \text{(unbiased)}$$

$$\hat{S}_{2MLE} = \left(\sum_{i=m+1}^{2m} i\right)^{-1} \sum_{i=m+1}^{2m} iX_i = \left[\frac{(m+1+2m)m}{2}\right]^{-1} \sum_{i=m+1}^{2m} iX_i$$

$$= \left[\frac{2}{(3m+1)m}\right] \sum_{i=m+1}^{2m} iX_i$$

$$Var(\hat{S}_{2MLE}) = \frac{4}{(3m+1)^2 m^2} \sum_{i=m+1}^{2m} i^2 Var(X_i)$$

$$= \frac{4}{(3m+1)^2 m^2} \sum_{i=m+1}^{2m} i^2 \frac{m}{i} = \frac{4}{(3m+1)^2 m} \sum_{i=m+1}^{2m} i$$

$$= \frac{4}{(3m+1)^2 m} \frac{(m+1+2m)m}{2} = \frac{2}{3m+1} \to 0 \text{ as } m \to \infty$$

$\therefore$ $\hat{S}_{2MLE}$ is consistent.