Assignment 1 Answer

Question1

1. Firstly, since $\{X_1, \cdots X_{100}\}$ is a r.s. from a distribution mean μ variance $\sigma^2 = 16$

Let $\bar{X}_n$ be the sample mean when the r.s. is of size n. Then by CLT:

$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$. In this question $\sigma^2 = 16$ is known $\therefore$ we have that

$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0,16)$ when n is large.

Now: from the question, we have function $g(t) = t(t+3)$ $g'(t) = 2t+3$ exists

obviously $g'(\mu) \neq 0$. Hence by the Delta method: $\sqrt{n}(g(\bar{X}_n) - g(\mu)) \xrightarrow{d} N(0, 16g'(\mu)^2)$

i.e. $\sqrt{n}(\bar{X}_n(\bar{X}_n+3) - \mu(\mu+3)) \xrightarrow{d} N(0, 16(2\mu+3)^2)$ when n is large

i.e. $\frac{\sqrt{n}(g(\bar{X}_n) - g(\mu))}{4(2\mu+3)} \xrightarrow{d} N(0,1)$

By WLLN: $\bar{X}_n \xrightarrow{P} \mu$. by continuous mapping thm. $\frac{1}{2\bar{X}_n+3} \xrightarrow{P} \frac{1}{2\mu+3}$, $\frac{1}{2\bar{X}_n+3} \xrightarrow{d} \frac{1}{2\mu+3}$

By slutsky's Thm. $\frac{2\mu+3}{2\bar{X}_n+3} \xrightarrow{d} \frac{2\mu+3}{2\mu+3} = 1 \Rightarrow \frac{2\mu+3}{2\bar{X}_n+3} \xrightarrow{P} 1$

So we have $Y_n = \frac{\sqrt{n}(g(\bar{X}_n) - g(\mu))}{4(2\bar{X}_n+3)} = \frac{\sqrt{n}(g(\bar{X}_n) - g(\mu))}{4(2\mu+3)} \cdot \frac{2\mu+3}{2\bar{X}_n+3} \xrightarrow{d} N(0,1)$

Therefore, $1-\alpha \approx P(-z_{\frac{\alpha}{2}} \leq Y_n \leq z_{\frac{\alpha}{2}})$

$= P\left(\bar{X}_n(\bar{X}_n+3) - \frac{4(2\bar{X}_n+3)}{\sqrt{n}} z_{\frac{\alpha}{2}} \leq \mu(\mu+3) \leq \bar{X}_n(\bar{X}_n+3) + \frac{4(2\bar{X}_n+3)}{\sqrt{n}} \cdot z_{\frac{\alpha}{2}}\right)$

The term inside the probability can represent the approximate $100(1-\alpha)\%$ random interval

Now, when $n=100$, $\bar{x}=15$. (replace $\bar{X}_n$) $\alpha = 0.05$. we have the 95% appro. C.I.

for $\mu(\mu+3)$: $\left[15(15+3) - \frac{4(2\times15+3)}{10} \cdot 1.96,\ 15\times(15+3) + \frac{4\times(2\times15+3)}{10} \cdot 1.96\right]$

Question2

2. Let the population be X with mean $\mu = \theta > 0$, variance σ^2. μ, σ are unknown.

$g'(\mu)\cdot\sigma = 1 \Rightarrow g'(\theta) = \frac{1}{\theta} \Rightarrow$ we assume $g(t) = \ln t$. $\therefore g'(t) = \frac{1}{u}$ exists. $g'(\mu)\cdot\sigma = 1$

Now: let $\{X_1, X_2, \cdots X_n\}$ be the r.s. and $\bar{X}_n$ be the sample mean.

By CLT: we have $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\mu} \xrightarrow{d} N(0,1)$ i.e. $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \mu^2)$ as n is large.

By Delta method $\sqrt{n}(g(\bar{X}_n) - g(\mu)) \xrightarrow{d} N(0,1)$ i.e. $\sqrt{n}(\ln\bar{X}_n - \ln\mu) \xrightarrow{d} N(0,1)$

Therefore $1-\alpha \approx P(-z_{\frac{\alpha}{2}} \leq \sqrt{n}(\ln\bar{X}_n - \ln\mu) \leq z_{\frac{\alpha}{2}})$

$$= P(\ln\bar{X}_n - \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}} \leq \ln\mu \leq \ln\bar{X}_n + \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}})$$

$$= P(\exp(\ln\bar{X}_n - \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}}) \leq \mu \leq \exp(\ln\bar{X}_n + \frac{z_{\frac{\alpha}{2}}}{\sqrt{n}}))$$

The term inside the probability can represent the approximate $100(1-\alpha)\%$ random interval

$\therefore$ For 95% approximated c.I. for μ, replace $\bar{X}_n$ by $\bar{x} = 34.23$, $n=50$. $z_{\frac{\alpha}{2}} = 1.96$

We get $[25.9434, 45.1635]$

Question 3

V1:

3. We find density of Cauchy i.e. $t(1)$ distribution first:

By the lecture notes: when $df=1$ $f_T(t) = \frac{\Gamma(1)}{\sqrt{\pi}\Gamma(\frac{1}{2})} \cdot \frac{1}{1+t^2} = \frac{1}{\pi} \cdot \frac{1}{1+t^2}$.

And: $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ $f_Y(y) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}}$

$\because$ X and Y are independent ($X \perp Y$), The joint density: $f_{X,Y}(x,y) = \frac{1}{2\pi} \cdot e^{-\frac{x^2+y^2}{2}}$

We want to find the density of $\frac{X}{Y}$, we can do transformation. $\begin{cases} W = h(X,Y) = \frac{X}{Y} \\ V = g(X,Y) = Y \end{cases}$

So the inverse transform is $\begin{cases} X = h^{-1}(W,V) = WV \\ Y = g^{-1}(W,V) = V \end{cases}$. We can get the Jacobian

$$J = \begin{vmatrix} \frac{\partial h^{-1}}{\partial W} & \frac{\partial h^{-1}}{\partial V} \\ \frac{\partial g^{-1}}{\partial W} & \frac{\partial g^{-1}}{\partial V} \end{vmatrix} = \begin{vmatrix} V & W \\ 0 & 1 \end{vmatrix} = V$$

Then, by what was introduced in previous probability courses, we have the joint density of W, V: $f_{W,V}(w,v) = |J| \cdot f_{X,Y}(h^{-1}(w,v), g^{-1}(w,v)) = |v| \cdot \frac{1}{2\pi} e^{-\frac{w^2v^2+v^2}{2}}$

Our goal is the density of $\frac{X}{Y} = W$ $\therefore$ We integrate to have

$$f_W(w) = \int_{-\infty}^{\infty} f_{W,V}(w,v)\,dv = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{w^2v^2+v^2}{2}} |v|\,dv = \frac{1}{2\pi}\int_{-\infty}^{\infty} |v| \cdot e^{-\frac{v^2(w^2+1)}{2}}\,dv$$

Consider $\int_0^{\infty} v e^{-\frac{(w^2+1)v^2}{2}}\,dv = \frac{1}{2}\int_0^{\infty} e^{-\frac{(w^2+1)}{2}t}\,dt = \frac{1}{2}\left[-\frac{2}{w^2+1} e^{-\frac{w^2+1}{2}t}\right]_0^{\infty} = \frac{1}{w^2+1}$

Similarly $\int_{-\infty}^{0} -v e^{-\frac{(w^2+1)v^2}{2}}\,dv = \frac{1}{w^2+1}$ Hence: $f_W(w) = \frac{1}{2\pi} \cdot \frac{2}{w^2+1} = \frac{1}{\pi} \cdot \frac{1}{w^2+1}$

$\therefore W = \frac{X}{Y} \sim t(1)$

V2

Q3.

We know $Y^2 \sim \chi^2(1) \Rightarrow \frac{X}{|Y|} \sim t(1)$

Need to prove $\frac{X}{|Y|} \overset{d}{=} \frac{X}{Y}$

$P(\frac{X}{Y} \le u) = P(X \le uY, Y>0) + P(X \ge uY, Y<0)$

$P(\frac{X}{|Y|} \le u) = P(X \le u|Y|)$

$= P(X \le uY, Y>0) + \underline{P(X \le -uY, Y<0)}$

N(0,1) is symmetric, so it equals $P(-X \le -uY, Y<0)$

$= P(X \ge uY, Y<0)$

So they are equal.

Question4

$$4.\ S_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

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

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

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

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

$$(n-1)^2(S_{n-1}^4) = \left(\sum_{i=1}^{n}(X_i-\mu)^2 - n(\bar{X}-\mu)^2\right)^2$$

$$= \left(\sum_{i=1}^{n}(X_i-\mu)^2\right)^2 + n^2(\bar{X}-\mu)^4 - 2\left(\sum_{i=1}^{n}(X_i-\mu)^2\right)(n\bar{X}-\mu)^2$$

$$\therefore E[(n-1)^2(S_{n-1}^4)] = E\left[\sum_{ij}(X_i-\mu)^2(X_j-\mu)^2\right] + n^2E(\bar{X}-\mu)^4 - 2nE\left[\sum_{i=1}^{n}(X_i-\mu)^2(\bar{X}-\mu)^2\right]$$

As $E(X_i-\mu)=0$, $E(X_j-\mu)^2=\sigma^2$, by simplication,

$$E[(n-1)^2(S_{n-1}^4)] = n(n-1)\sigma^4 + n\mu_4 + 2\mu_4 + 2(n-1)\sigma^4 - \frac{1}{n}(\mu_4+3(n-1)\sigma^4)$$

$$(n-1)^2E(S_{n-1}^4) = n(n-1)\sigma^4 + n\mu_4 + 2\mu_4 + 2(n-1)\sigma^4 - \frac{1}{n}\mu_4 - \frac{3}{n}(n-1)\sigma^4$$

$$= (n-1)\sigma^4(n+2-\frac{n}{3}) + \mu_4(2+n-\frac{1}{n})$$

$$= (n-1)\sigma^4(\frac{2n+6}{3}) + \mu_4(\frac{n^2-1+2n}{n})$$

by simplification: $E(S^4_{n-1}) = \frac{n^2-2n+3}{n(n-1)}\sigma^4 + \frac{1}{n}\mu_4$

$$\therefore Var\, S_n^2 = E(S^4_{n-1}) - E(S^2_{n-1})^2 = \frac{n^2-2n+3}{n(n-1)}\sigma^4 + \frac{1}{n}\mu_4 - \sigma^4$$

$$= \frac{1}{n}\left(\frac{n^2-2n+3}{n-1}\sigma^4 - n\sigma^4 + \mu_4\right)$$

$$= \frac{1}{n}\left(\frac{n^2-2n+3-n^2+n}{n-1}\sigma^4 + \mu_4\right)$$

$$= \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right)$$

Question 5

V1

5. We reorder the r.s. $\{X_1, X_2, \dots X_n\}$ to an order statistics $\{X_{(1)}, X_{(2)}, \dots X_{(n)}\}$, where $X_{(1)} \le X_{(2)} \le \dots \le X_{(n)}$.

Since r.s. $\{X_1, X_2, \dots X_n\}$ froms a uniform distribution over $[0, \theta]$,

$$f_{X_{(i)}}(u) = \frac{n!}{(i-1)!(n-i)!} f_X(u)[F_X(u)]^{i-1}[1-F_X(u)]^{n-i} = \frac{n!}{(i-1)!(n-i)!} \cdot \frac{1}{\theta} \cdot \left(\frac{x}{\theta}\right)^{i-1} \cdot \left(\frac{\theta - x}{\theta}\right)^{n-i}$$

$\frac{X_{(i)}}{\theta} \sim Beta(i, n-i+1)$. $Var\left(\frac{X_{(i)}}{\theta}\right) = \frac{i(n-i+1)}{(n+1)^2(n+2)} \Rightarrow Var(X_{(i)}) = \frac{i(n-i+1)\theta^2}{(n+1)^2(n+2)}$

$$\frac{d}{di} Var(X_{(i)}) = \frac{\theta^2}{(n+1)^2(n+2)}(n+1-2i)$$

Therefore, $Var(X_{(i)})$ has the global maximum where $\frac{d}{di} Var(X_{(i)}) = 0$, $i = \frac{n+1}{2}$, and has no global minimum since it is concave. The relative minimum is at $i=1$ and $i=n$, where $\frac{d}{di} Var(X_{(i)}) = \frac{\theta^2 n}{(n+1)^2(n+2)}$.

The order statistics with the smallest variance is $X_{(1)} = \min(X_1, \dots X_n)$ and $X_{(n)} = \max(X_1, \dots X_n)$.

V2

5. p.d.f of X_i = $f_i(x) = \frac{n!}{(i-1)!(n-i)!} \cdot \frac{x^{i-1} \cdot (\theta-x)^{n-i}}{\theta^n}$

$Var(X_i) = E(X_i^2) - (E(X_i))^2 = \frac{n!}{(i-1)!(n-i)!} \frac{x^{i-1} \cdot (\theta-x)^{n-i}}{\theta^n}$

$$
\begin{aligned}
Var(X_i) &= E(X_i^2) - (E(X_i))^2 \\
&= \frac{n!}{(i-1)!(n-i)!} \int_0^\theta x^2 \frac{x^{i-1}(\theta-x)^{n-i}}{\theta^n} dx - \left[\frac{n!}{(i-1)!(n-1)!} \left(\frac{x^{i-1}(\theta-x)^{n-1}}{\theta^n}\right)\right]^2 dx \\
&\overset{y=\frac{x}{\theta}}{=} \frac{n!}{(i-1)!(n-i)!} \int_0^1 \frac{(\theta y)^{i+1}(\theta-\theta y)^{n-i}}{\theta^n} \theta dy - \left[\frac{n!}{(i-1)!(n-i)!} \int_0^1 \frac{(\theta y)^i(\theta-\theta y)^{n-i}}{\theta^n} \cdot \theta dy\right]^2 \\
&= \frac{n!}{(i-1)!(n-i)!} \theta^2 \int_0^1 y^{i+1}(1-y)^{n-1} dy - \left[\frac{n!}{(i-1)!(n-i)!} \theta \int_0^1 y^2 (1-y)^{n-i} dy\right]^2 \\
&= \frac{n!}{(i-1)!(n-i)!} \theta^2 B(i+2, n-i+1) - \left[\frac{n!}{(i-1)!(n-i)!} \theta B(i+1, n-i+1)\right]^2 \\
&= \frac{i(i+1)}{(n+1)(n+2)} \theta^2 \frac{(i+1)!(n-i)!}{(n+2)!} - \left[\frac{n!}{(i-1)!(n-i)!} \theta \cdot \frac{i!(n-i)!}{(n+1)!}\right]^2 \\
&= \frac{i(i+1)}{(n+1)(n+2)} \theta^2 - \left(\frac{2}{n+1} \theta\right)^2 \\
&= \frac{i}{n+1} \theta^2 \left[\frac{i+1}{n+2} - \frac{i}{n+1}\right] \\
&= \frac{i(n-i+1)\theta^2}{(n+1)^2(n+2)}
\end{aligned}
$$

$\therefore$ when $i=1$ or n, $Var(X_i)$ is the smallest.

V3

5 Question 5

For order statistics of X,

$$\begin{aligned}
f_{X_i}(u) &= \frac{n!}{(i-1)!(n-i)!} f_X(u)[F_X(u)]^{i-1}[1-F_X(u)]^{n-i} \\
&= \frac{n!}{(i-1)!(n-i)!}\frac{1}{\theta}[\frac{u}{\theta}]^{i-1}[1-\frac{u}{\theta}]^{n-i} \\
&= \frac{n!}{(i-1)!(n-i)!\theta}[\frac{u}{\theta}]^{i-1}[1-\frac{u}{\theta}]^{n-i} \\
&= \frac{n!}{(i-1)!(n-i)!\theta^n}[u]^{i-1}[\theta-u]^{n-i} \\
E(X_i) &= \int_0^\theta u f_{X_i}(u)du \\
&= \int_0^\theta u\frac{n!}{(i-1)!(n-i)!\theta^n}[u]^{i-1}[\theta-u]^{n-i}du \\
&= \frac{n!}{(i-1)!(n-i)!\theta^n}\int_0^\theta [u]^i[\theta-u]^{n-i}du \\
\int_0^\theta [u]^i[\theta-u]^{n-i}du &= \theta^{n+1}B(i+1, n-i+1) \\
&= \theta^{n+1}\frac{i!(n-i)!}{(n+1)!} \\
E(X_i) &= \frac{n!}{(i-1)!(n-i)!\theta^n}\theta^{n+1}\frac{i!(n-i)!}{(n+1)!} \\
&= \frac{\theta}{1}\frac{i}{n+1} \\
&= \frac{i\theta}{n+1}
\end{aligned}$$

Similar to E(X_i),

$$\begin{aligned}
E(X_i^2) &= \int_0^\theta u^2 f_{X_i}(u)du \\
&= \frac{n!}{(i-1)!(n-i)!\theta^n} \int_0^\theta [u]^{i+1}[\theta-u]^{n-i}du \\
\int_0^\theta [u]^{i+1}[\theta-u]^{n-i}du &= \theta^{n+2}B(i+2, n-i+1) \\
&= \theta^{n+2}\frac{(i+1)!(n-i)!}{(n+2)!} \\
E(X_i^2) &= \frac{n!}{(i-1)!(n-i)!\theta^n}\theta^{n+2}\frac{(i+1)!(n-i)!}{(n+2)!} \\
&= \frac{\theta^2}{1}\frac{i(i+1)}{(n+1)(n+2)} \\
&= \frac{i(i+1)\theta^2}{(n+1)(n+2)} \\
Var(X_i) &= E[X_i^2] - E[X_i]^2 \\
&= \frac{i(i+1)\theta^2}{(n+1)(n+2)} - (\frac{i\theta}{n+1})^2 \\
&= \frac{(n+1)i(i+1)\theta^2 - (n+2)i^2\theta^2}{(n+1)^2(n+2)} \\
&= \frac{\theta^2}{(n+1)^2(n+2)}(ni^2 + ni + i^2 + i - ni^2 - 2i^2) \\
&= \frac{\theta^2}{(n+1)^2(n+2)}(-i^2 + (n+1)i)
\end{aligned}$$

From $Var(X_i)$, $(-i^2)$ indicate that it only contains maximum point.
After calculated $\frac{dVar(X_i)}{dt} = 0$, the peak locate at $i = \frac{n+1}{2}$, while $1 \leq i \leq n$ and $Var(X_i)$ is quadratic with symmetric property.
$\therefore$ i minimize at $i = 1$ or n.
$\therefore$ X_1 and X_n has smallest variance.