MATH20802: STATISTICAL METHODS
SECOND SEMESTER
ANSWERS TO THE IN CLASS TEST

ANSWERS TO QUESTION 1

(i) Note that

$$\begin{aligned}
E(T(X)) &= 1 \cdot \theta^3 + 2(1-c) \cdot \theta^2(1-\theta) + 2c \cdot \theta(1-\theta) + (1-2c) \cdot \theta(1-\theta)^2 \\
&= \theta^3 + 2(1-c) \cdot \left(\theta^2 - \theta^3\right) + 2c \cdot \left(\theta - \theta^2\right) + (1-2c) \cdot \left(\theta - 2\theta^2 + \theta^3\right) \\
&= \theta^3 + 2(1-c) \cdot \theta^2 - 2(1-c) \cdot \theta^3 + 2c \cdot \theta - 2c \cdot \theta^2 + (1-2c) \cdot \theta - 2(1-2c) \cdot \theta^2 + (1-2c) \cdot \theta^3 \\
&= (2c + 1 - 2c) \cdot \theta + (2 - 2c - 2c - 2 + 4c) \cdot \theta^2 + (1 - 2 + 2c + 1 - 2c) \cdot \theta^3 \\
&= 1 \cdot \theta + 0 \cdot \theta^2 + 0 \cdot \theta^3 \\
&= \theta
\end{aligned}$$

Hence, $T(X)$ is unbiased for all c. (3 marks)

(ii) Using the fact $\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \sigma^2 \chi^2_{n-1}$, we see that

$$E\left(S^2\right) - \sigma^2 = (1/n)\sigma^2 E\left(\chi^2_{n-1}\right) - \sigma^2 = -(1/n)\sigma^2$$

and

$$Var\left(S^2\right) = \left(1/n^2\right)\sigma^4 Var\left(\chi^2_{n-1}\right) = 2\left((n-1)/n^2\right)\sigma^4.$$

So, S^2 is biased but is consistent. (2 marks)

(iii) We know that

$$E\left(\overline{X}\right) = (1/n)E\left(X_1 + X_2 + \cdots + X_n\right) = (1/n)n\mu = \mu,$$

so

$$E\left(X_1 + \overline{X}\right)/2 = (1/2)\left(E\left(X_1\right) + E\left(\overline{X}\right)\right) = \mu.$$

Note that

$$Var\left(\frac{X_1 + \overline{X}}{2}\right) = \frac{1}{4}Var\left(X_1 + \overline{X}\right) \neq \frac{1}{4}\left[Var\left(X_1\right) + Var\left(\overline{X}\right)\right]$$

since X_1 and $\overline{X}$ are not independent. But we can write

$$Var\left(\frac{X_1 + \overline{X}}{2}\right) = \frac{1}{4}Var\left(X_1 + \overline{X}\right)$$

$$
\begin{aligned}
&= \frac{1}{4} Var\left(X_1 + \frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{4} Var\left(X_1 + \frac{X_1}{n} + \frac{X_2}{n} + \cdots + \frac{X_n}{n}\right) \\
&= \frac{1}{4} Var\left(\left(1 + \frac{1}{n}\right) X_1 + \frac{X_2}{n} + \cdots + \frac{X_n}{n}\right) \\
&= \frac{1}{4} \left[Var\left(\left(1 + \frac{1}{n}\right) X_1\right) + Var\left(\frac{X_2}{n} + \cdots + \frac{X_n}{n}\right)\right] \\
&\qquad \text{since } X_1 \text{ and } X_2, \ldots, X_n \text{ are independent} \\
&= \frac{1}{4} \left[\left(1 + \frac{1}{n}\right)^2 Var\left(X_1\right) + \frac{1}{n^2} Var\left(X_2 + \cdots + X_n\right)\right] \\
&= \frac{1}{4} \left[\left(1 + \frac{1}{n}\right)^2 Var\left(X_1\right) + \frac{1}{n^2} Var\left(\sum_{i=2}^{n} X_i\right)\right] \\
&= \frac{1}{4} \left[\left(1 + \frac{1}{n}\right)^2 Var\left(X_1\right) + \frac{1}{n^2} \sum_{i=2}^{n} Var\left(X_i\right)\right] \\
&= \frac{1}{4} \left[\left(1 + \frac{1}{n}\right)^2 \sigma^2 + \frac{1}{n^2} \sum_{i=2}^{n} \sigma^2\right] \\
&= \frac{1}{4} \left[\left(1 + \frac{1}{n}\right)^2 \sigma^2 + \frac{1}{n^2} (n-1)\sigma^2\right].
\end{aligned}
$$

The limit of this as $n \to \infty$ is $\sigma^2/4$, so the estimator is not consistent. (3 marks)

(iv) We know that

$$E\left(\overline{X}\right) = (1/n)E\left(X_1 + X_2 + \cdots + X_n\right) = (1/n)n\lambda = \lambda$$

and

$$Var\left(\overline{X}\right) = \left(1/n^2\right) Var\left(X_1 + X_2 + \cdots + X_n\right) = \left(1/n^2\right) n\lambda = \lambda/n.$$

So,

$$E\left(\overline{X}^2\right) = Var\left(\overline{X}\right) + \left(E\left(\overline{X}\right)\right)^2 = \lambda/n + \lambda^2.$$

Hence, the estimator is biased. (2 marks)

ANSWERS TO QUESTION 2

(i) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x) = 2\alpha^{-2} x \exp\{-x^2/\alpha^2\}$. The log-likelihood function is

$$\log L(\alpha) = n \log 2 - 2n \log \alpha + \sum_{i=1}^{n} \log X_i - \alpha^{-2} \sum_{i=1}^{n} X_i^2.$$

The normal equation is:

$$\frac{\partial \log L(\alpha)}{\partial \alpha} = -\frac{2n}{\alpha} + \frac{2}{\alpha^3} \sum_{i=1}^{n} X_i^2 = 0.$$

The root of this equation, $\sqrt{(1/n) \sum_{i=1}^{n} X_i^2}$, is the mle of α. (2 marks)

(ii) Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal population with mean θ and standard deviation θ. The log-likelihood function is

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

The normal equation is:

$$\frac{\partial \log L(\theta)}{\partial \theta} = -\frac{n}{\theta} + \frac{1}{\theta^3} \sum_{i=1}^{n} X_i^2 - \frac{1}{\theta^2} \sum_{i=1}^{n} X_i = 0.$$

This is a quadratic equation on θ. Its two roots are:

$$-\frac{1}{2n} \sum_{i=1}^{n} X_i \pm \frac{1}{2n} \sqrt{\left(\sum_{i=1}^{n} X_i \right)^2 + 4n \sum_{i=1}^{n} X_i^2}.$$

The positive sign gives the valid mle for θ. (3 marks)

(iii) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x) = \lambda^{-2} x \exp\{-x/\lambda\}$. The log-likelihood function is

$$\log L(\lambda) = -2n \log \lambda + \sum_{i=1}^{n} \log X_i - \lambda^{-1} \sum_{i=1}^{n} X_i.$$

The normal equation is:

$$\frac{\partial \log L(\lambda)}{\partial \lambda} = -\frac{2n}{\lambda} + \frac{1}{\lambda^2} \sum_{i=1}^{n} X_i = 0.$$

The root of this equation, $(1/(2n)) \sum_{i=1}^{n} X_i$, is the mle of λ. (2 marks)

(iv) Let X_1, X_2, X_3 be a random sample from the discrete uniform distribution between 1 and N, inclusive. The likelihood function is

$$\begin{aligned} L(N) &= N^{-3} I \{1 \leq X_1 \leq N, 1 \leq X_2 \leq N, 1 \leq X_3 \leq N\} \\ &= N^{-3} I \{\max(X_1, X_2, X_3) \leq N, \min(X_1, X_2, X_3) \geq 1\}, \end{aligned}$$

a monotonic decreasing function of N for all $N \geq \max(X_1, X_2, X_3)$. So, the mle $\widehat{N} = \max(X_1, X_2, X_3)$. (3 marks)