BSDS (2024, Semester II): Statistics II

Mid-semester Examination

NAME :	**ROLL :**

Time: 120 minutes **Total attainable marks:** 30

1. Let $X_1, \ldots, X_n$ be a random sample from exponential distribution with pdf

$$f_\lambda(x) = \lambda \exp\{-\lambda x\}, \qquad x > 0, \quad \lambda > 0.$$

Consider the parameter of interest $\psi(\lambda) = \dfrac{\lambda}{1+\lambda}$.

(a) Find an MLE of $\psi(\lambda)$, say $T_{\mathrm{MLE},n}$. [4]

We first find the MLE of λ.

The log-likelihood function

$$l_n(\lambda \mid \underset{\sim}{x}_n) = \log\left(\lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}\right)$$

$$= n \log \lambda - \lambda \sum_{i=1}^{n} x_i$$

First order equation:

$$\frac{\partial l_n(\lambda \mid \underset{\sim}{x}_n)}{\partial \lambda} = 0 \Rightarrow \frac{n}{\lambda} - n\bar{x}_n = 0 \Rightarrow \hat{\lambda} = \frac{1}{\bar{x}_n}.$$

Second order condition:

$$\left.\frac{\partial^2 l_n(\lambda \mid \underset{\sim}{x}_n)}{\partial \lambda^2}\right|_{\lambda = \hat{\lambda}} = -\frac{n}{\hat{\lambda}^2} = -n\bar{x}_n^2 < 0$$

Thus, $\hat{\lambda}_{MLE} = 1/\bar{x}_n$.

By the invariance property of MLE, MLE of $\psi(\lambda)$ is:

$$T_{MLE,n} = \psi(\hat{\lambda}_{MLE}) = \frac{\hat{\lambda}_{MLE}}{1+\hat{\lambda}_{MLE}} = \frac{1}{\bar{x}_n + 1}.$$

(b) Show that S_n/n is consistent for $1/\lambda$. [4]

Here $S_n = X_1 + \cdots + X_n$ and $S_n/n = \bar{X}_n$.

Now, $E(\bar{x}_n) = E(x_1) = \int_0^\infty \lambda x e^{-\lambda x} dx$

$= \int_0^\infty \lambda^{-1} u e^{-u} du$ where $\lambda x = u$

$= \frac{1}{\lambda} \Gamma 2 = \frac{1}{\lambda}$

By Weak law of large numbers (WLLN) we know that when $x_1, \dots, x_n$ are IID samples with finite variance, then $\bar{X}_n \xrightarrow{P} 1/\lambda$.

Here $var(\bar{x}_n) = \frac{var(x_1)}{n}$.

Further, $var(x_1) = E(x_1^2) - 1/\lambda^2$

$= \int_0^\infty \lambda x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2}$

$= \frac{1}{\lambda^2} \int_0^\infty u^2 e^{-u} du - \frac{1}{\lambda^2}$

$= \frac{\Gamma 3}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} < \infty$

So, $\bar{x}_n$ is consistent for $1/\lambda^2$.

Is $T_{\mathrm{MLE},n}$ consistent for $\psi(\lambda)$? [4]

As $g(x) = \frac{1}{x+1}$ is a continuous function of x on $(0,\infty)$,

by continuous mapping theorem,

if $\bar{X}_n \xrightarrow{P} 1/\lambda$

then $g(\bar{x}_n) \xrightarrow{P} g(1/\lambda)$.

Further, as $g(\bar{x}_n) = T_{\mathrm{MLE},n}$, and $g(1/\lambda) = \frac{1}{1+1/\lambda} = \psi(\lambda)$,

$T_{\mathrm{MLE},n}$ is consistent for $\psi(\lambda)$.

(c) Consider another estimator of $\psi(\lambda)$ as $W_n = \dfrac{1}{n}\sum_{i=1}^{n} \exp\{-X_i\}$.

i. Show that W_n is unbiased for $\psi(\lambda)$. [4]

$$E(W_n) = \frac{1}{n}\sum_{i=1}^{n} E\left(e^{-X_i}\right) = E\left(e^{-X_1}\right)$$

Now,

$$E\left(e^{-X_1}\right) = \int_0^\infty \lambda e^{-x} e^{-\lambda x}\, dx$$

$$= \lambda \int_0^\infty e^{-(1+\lambda)x}\, dx$$

$$= \frac{\lambda}{1+\lambda} \int_0^\infty e^{-u}\, du \quad \text{where } u = (1+\lambda)x$$

$$= \frac{\lambda}{1+\lambda} = \psi(\lambda).$$

Thus, W_n is an unbiased estimator of $\psi(\lambda)$.

ii. Is W_n consistent for ~~$\psi(\theta)$?~~ $\psi(\lambda)$? [4]

For $i=1,\dots,n$, define $W_i = e^{-X_i}$. Then $W_n = \frac{1}{n}\sum_{i=1}^{n} W_i$.

By WLLN,

$$W_n \xrightarrow{P} E(W_1) = \psi(\lambda),$$

provided $\mathrm{Var}(W_1) < \infty$.

Here

$$\mathrm{Var}(W_1) = E(W_1^2) - \psi^2(\lambda)$$

$$= \lambda \int_0^\infty e^{-2x-\lambda x}\,dx - \frac{\lambda}{1+\lambda}$$

$$= \frac{\lambda}{(\lambda+2)} \int_0^\infty e^{-u}\,du - \frac{\lambda^2}{(1+\lambda)^2},$$

where $u = (\lambda+2)x$.

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

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

$$= \frac{\lambda}{(\lambda+1)^2(\lambda+2)} < \infty.$$

So, W_n is consistent for $\psi(\lambda)$.

iii. Find sequences of real numbers $\{a_n\}$ and $\{b_n\}$ such that $\dfrac{W_n - a_n}{b_n} \xrightarrow{d} X$, where $X \sim N(0,1)$ distribution. [Hint: Apply CLT] [4]

By Central Limit Theorem (CLT) we have

$$\frac{\sqrt{n}\left(\bar{W}_n - E(W_1)\right)}{\sqrt{var(W_1)}} \xrightarrow{d} X,$$

where $X \sim N(0,1)$.

By the above calculations,

$$E(W_1) = \psi(\lambda)$$

$$var(W_1) = \frac{\psi^2(\lambda)}{\lambda(\lambda+2)}$$

$$\therefore \quad \frac{\sqrt{n}\left(\bar{W}_n - \psi(\lambda)\right)}{\sqrt{\psi^2(\lambda) / \{\lambda(\lambda+2)\}}} \xrightarrow{d} X$$

So, $a_n = \psi(\lambda)$ for all n,

and $b_n = \dfrac{\psi(\lambda)}{\sqrt{n\lambda(\lambda+2)}}$, $n \in \mathbb{N}$.

2. Suppose the random measurements of a quantity is modeled as `Gamma`(α, β) with pdf

$$f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma\alpha} x^{\alpha-1} \exp\{\cancel{-\beta\alpha}\}\ \{-\beta x\}, \qquad x > 0, \quad \alpha, \beta > 0.$$

Let the true measurement is α/β. If n independent measurements are taken, find the Uniformly Minimum Variance Unbiased Estimator (UMVUE)? [5]

Let $X_1, \dots, X_n$ be the random variables indicating the independent measurements.

Then $X_i \overset{IID}{\sim}$ Gamma (α, β) ; $i = 1, \dots, n$

To find the UMVUE of α/β, we first find a complete sufficient statistic (CSS) for the family.

Observe that,

$$f_{\alpha,\beta}(\underset{\sim}{x}) = \frac{\beta^{n\alpha}}{(\Gamma\alpha)^n} \left(\prod_{i=1}^n x_i\right)^{\alpha-1} e^{-\beta\sum_{i=1}^n x_i}$$

$$= \exp\left\{ n\alpha \log\beta - n\log\Gamma\alpha + (\alpha-1)\sum_{i=1}^n \log x_i - \beta\sum_{i=1}^n x_i \right\}$$

$$= \exp\left\{ a(\alpha,\beta) + h(\underset{\sim}{x}) + T_1(\underset{\sim}{x}) K_1(\alpha,\beta) + T_2(\underset{\sim}{x}) K_2(\alpha,\beta) \right\}.$$

with $a(\alpha,\beta) = n\alpha\log\beta - n\log\Gamma\alpha$

$h(\underset{\sim}{x}) = 0$

$T_1(\underset{\sim}{x}) = \sum_{i=1}^n \log x_i$

$K_1(\alpha,\beta) = (\alpha - 1)$

$T_2(\underset{\sim}{x}) = \sum_{i=1}^n x_i$

and $K_2(\alpha,\beta) = -\beta$.

Thus, $\underset{\sim}{T}(\underset{\sim}{X}) = \left(\sum_{i=1}^n \log X_i, \sum_{i=1}^n X_i\right)'$ is jointly CSS.

Next, observe that,

$$E(X_1) = \int_0^\infty \frac{\beta^\alpha}{\Gamma\alpha} x^{\alpha} e^{-\beta x} \, dx$$

$$= \int_0^\infty \frac{1}{\beta \Gamma\alpha} u^{\alpha} e^{-u} \, du \quad \text{where } u = \beta x$$

$$= \frac{\Gamma(\alpha+1)}{\beta \Gamma\alpha} = \frac{\alpha}{\beta}.$$

Therefore, $E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\alpha}{\beta}.$

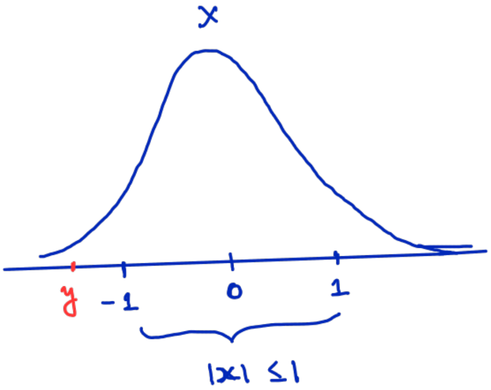
So, the statistic $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = g\left(\underset{\sim}{T}(\underset{\sim}{X})\right)$ is a function of a CSS $\left(\text{here, } g\left(\binom{x}{y}\right) = y/n\right)$, and is unbiased for α/β. So, $\overline{X}_n$ is the UMVUE of α/β.

3. Let $X \sim \texttt{normal}(0,1)$. Define $Y = -X\mathbb{I}(|X| \le 1) + X\mathbb{I}(|X| > 1)$. Find the distribution of Y. [5]

Observe that $Y = X$ when $X > 1$ or $X < -1$

and $Y = -X$ when $-1 \le X \le 1$.

We will find the CDF of Y, say F_Y.



Let $-\infty < y < 1$,

then $F_Y(y) = P(Y \le y)$

$= P(X \le y) = \Phi(y)$ — ①

Let $-1 \le y \le 1$,

then $F_Y(y) = P(Y \le -1) + P(-1 \le Y \le y)$

$= P(X \le -1) + P(-1 \le -X \le y)$

$= \Phi(-1) + P(-y \le X \le 1)$

$= \Phi(-1) + \Phi(1) - \Phi(-y)$

$= 1 - \Phi(1) + \Phi(1) - 1 + \Phi(y)$

as $\Phi(-z) = 1 - \Phi(z)$

$= \Phi(y)$ — ②

Let $y > 1$, then

$F_Y(y) = P(Y \le 1) + P(1 < Y \le y)$

$= \Phi(1) + P(1 < X \le y)$ from ②

$= \Phi(1) + \Phi(y) - \Phi(1)$

$= \Phi(y)$ — ③

Combining ①-③ we get $F_Y(\cdot) = \Phi(\cdot)$. So, $Y \sim N(0,1)$.

4. A factory produces electronic components, and the number of defective components produced by each machine per day follows a Poisson distribution with a mean of 4 defectives. If 50 machines are in operation, what is the probability that the total number of defectives produced by all machines in a day is between 180 and 200? [You may use the fact $\Phi(\sqrt{2}) = 0.9214$.] [6]

Let X_i be the random variable indicating the number of defective items produced by the ith machine in a day.

$i = 1, \cdots, 50$

We assume, $X_i \overset{IID}{\sim}$ Poisson (λ) with $\lambda = 4$.

Need to calculate the probability,

$$P\Big(180 \leq \underbrace{\sum_{i=1}^{50} X_i}_{S_n} \leq 200\Big).$$

By CLT,

$$\frac{S_n - E(S_n)}{\sqrt{var(S_n)}} \xrightarrow{d} X \text{ where } X \sim N(0,1),$$

where $E(S_n) = 50 \times 4$

$var(S_n) = 50 \times 4$

$$\therefore \quad P(180 \leq S_n \leq 200)$$

$$= P\left(\frac{180-200}{\sqrt{200}} \leq \frac{S_n - 200}{\sqrt{200}} \leq 0\right)$$

$$= P\left(-\frac{20}{10\sqrt{2}} \leq \frac{S_n - 200}{10\sqrt{2}} \leq 0\right)$$

$$= F_{Z_n}(0) - F_{Z_n}(-\sqrt{2}) \text{ where } Z_n = \frac{S_n - 200}{10\sqrt{2}}$$

$$\longrightarrow \Phi(0) - \Phi(-\sqrt{2}) \text{ by CLT}$$

$$= \tfrac{1}{2} - \{1 - \Phi(\sqrt{2})\} = 0.9214 - 0.5 = \underline{0.4214} \quad [ANS]$$