Midterm Exam

Suggested Answers

Instructions: This is a closed book exam, but you may refer to one sheet of notes. You have 80 minutes for the exam. Answer as many questions as possible. Partial answers get partial credit. Please write legibly. *Good luck!*

Problem 1 (5 points). Determine whether or not the statement below is correct and give a *brief* (e.g., a bluebook page or less) justification for your answer.

Suppose $X_1, \ldots, X_n$ is a random sample from a continuous distribution with pdf $f(\cdot|\theta)$, where $\theta \in \Theta \subseteq \mathbb{R}$ is unknown. A statistic $T = T(X_1, \ldots, X_n)$ is complete if and only if

$$P_\theta[g(T) = \theta] = 1 \qquad \forall \theta \in \Theta$$

for every function $g(\cdot)$ satisfying

$$E_\theta[g(T)] = \theta \qquad \forall \theta \in \Theta.$$

The statement is incorrect. If $X_i \sim$ i.i.d. $\mathcal{N}(\theta, 1)$ and if $\Theta = \mathbb{R}$, then $\bar{X} = n^{-1} \sum_{i=1}^n X_i$ is complete and $E_\theta(\bar{X}) = \theta$, but $P_\theta(\bar{X} = \theta) = 0$ because $\bar{X} \sim \mathcal{N}(\theta, 1/n)$ is continuously distributed.

Problem 2 (45 points, each part receives equal weight). Let $X_1, \ldots, X_n$ be a random sample from a continuous distribution with mean $\sqrt{\theta}$ and pdf

$$f_X(x|\theta) = \frac{x}{j(\theta)} 1[0 \leq x \leq k(\theta)],$$

where $\theta \in \Theta = \mathbb{R}_{++}$ is an unknown parameter, $j : \Theta \to \mathbb{R}_{++}$ and $k : \Theta \to \mathbb{R}_{++}$ are some functions, and $1(\cdot)$ is the indicator function.

(a) Show that

$$j(\theta) = \frac{9}{8}\theta, \qquad k(\theta) = \frac{3}{2}\sqrt{\theta}.$$

Because $\int_{-\infty}^{\infty} f_X(x|\theta)\, dx = 1$ *and* $\int_{-\infty}^{\infty} x f_X(x|\theta)\, dx = \sqrt{\theta}$, $j(\theta)$ *and* $k(\theta)$ *must satisfy*

$$1 = \int_{-\infty}^{\infty} f_X(x|\theta)\, dx = \frac{1}{j(\theta)} \int_0^{k(\theta)} x dx = \frac{1}{j(\theta)} \frac{1}{2} x^2 \Big|_{x=0}^{k(\theta)} = \frac{k(\theta)^2}{2j(\theta)}$$

and

$$\sqrt{\theta} = \int_{-\infty}^{\infty} x f_X(x|\theta)\, dx = \frac{1}{j(\theta)} \int_0^{k(\theta)} x^2 dx = \frac{1}{j(\theta)} \frac{1}{3} x^3 \Big|_{x=0}^{k(\theta)} = \frac{k(\theta)^3}{3j(\theta)}.$$

This pair of equations rearranges as

$$j(\theta) = \frac{1}{2}k(\theta)^2, \qquad k(\theta) = \frac{3}{2}\sqrt{\theta},$$

from which the conclusion follows.

(b) Find $F_X(\cdot|\theta)$, the cdf of X.

For $x \in [0, k(\theta)]$, *we have:*

$$\int_{-\infty}^{x} f_X(r|\theta)\, dr = \frac{1}{j(\theta)} \int_0^x r dr = \frac{1}{j(\theta)} \frac{1}{2} r^2 \Big|_{r=0}^{x} = \frac{1}{2j(\theta)} x^2 = \frac{4}{9\theta} x^2.$$

As a consequence,

$$F_X(x|\theta) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{4}{9\theta} x^2 & \text{if } 0 \leq x < \frac{3}{2}\sqrt{\theta}, \\ 1 & \text{if } x \geq \frac{3}{2}\sqrt{\theta}. \end{cases}$$

(c) Derive a method moments estimator $\hat{\theta}_{MM}$ of θ. Is $\hat{\theta}_{MM}$ an unbiased estimator of θ?

Because $E_\theta(X_i) = \sqrt{\theta}$, $\theta = [E_\theta(X_i)]^2$ *and a method of moments estimator is therefore given by*

$$\hat{\theta}_{MM} = \bar{X}^2, \qquad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Because $Var_\theta(X_i) > 0$,

$$E_\theta(\hat{\theta}_{MM}) = E_\theta(\bar{X}^2) = E_\theta(\bar{X})^2 + Var_\theta(\bar{X}) = E_\theta(X_i)^2 + Var_\theta(X_i)/n > E_\theta(X_i)^2 = \theta.$$

In particular, $\hat{\theta}_{MM}$ *is a biased estimator of* θ.

(d) Find the likelihood function. Does θ admit a scalar sufficient statistic?

Defining $X_{(n)} = \max_{1\le i\le n} X_i$, *the likelihood function can be written as*

$$L\left(\theta|X_1,\ldots,X_n\right) = \prod_{i=1}^{n} f_X\left(X_i|\theta\right) = \prod_{i=1}^{n}\{\frac{X_i}{j(\theta)}\underbrace{1[0 \le X_i \le k(\theta)]}_{\equiv 1[X_i \le k(\theta)]}\} = j\left(\theta\right)^{-n}(\prod_{i=1}^{n} X_i)\cdot 1[X_{(n)} \le k(\theta)].$$

It follows from the factorization criterion that the scalar $X_{(n)}$ *is a sufficient statistic for* θ.

(e) Show that

$$\hat{\theta}_{ML} = \frac{4}{9}(\max_{1\le i\le n} X_i)^2$$

is the maximum likelihood estimator of θ.

Because

$$1[X_{(n)} \le k(\theta)] = 1[X_{(n)} \le \frac{3}{2}\sqrt{\theta}] = 1[\frac{4}{9}X_{(n)}^2 \le \theta]$$

and because $j(\theta)^{-n}$ *is a decreasing function of* θ,

$$\arg\max_{\theta\in\Theta} L\left(\theta|X_1,\ldots,X_n\right) = \frac{4}{9}X_{(n)}^2 = \hat{\theta}_{ML}.$$

(f) Show that the cdf of $\hat{\theta}_{ML}$ is given by

$$F_{ML}(x|\theta) = \begin{cases} 0 & \text{if } x < 0, \\ (x/\theta)^n & \text{if } 0 \leq x < \theta, \\ 1 & \text{if } x \geq \theta. \end{cases}$$

Clearly, $F_{ML}(x|\theta) = 0$ for $x < 0$. If $x \geq 0$, then

$$\begin{aligned} P_\theta(\hat{\theta}_{ML} &\leq x) = P_\theta(\frac{4}{9}X_{(n)}^2 \leq x) = P_\theta(X_{(n)} \leq \frac{3}{2}\sqrt{x}) = F_X(\frac{3}{2}\sqrt{x}|\theta)^n \\ &= \begin{cases} \frac{4}{9\theta}[\frac{3}{2}\sqrt{x}]^2 & \text{if } 0 \leq \frac{3}{2}\sqrt{x} < \frac{3}{2}\sqrt{\theta}, \\ 1 & \text{if } \frac{3}{2}\sqrt{x} \geq \frac{3}{2}\sqrt{\theta}, \end{cases} \\ &= \begin{cases} (x/\theta)^n & \text{if } 0 \leq x < \theta, \\ 1 & \text{if } x \geq \theta, \end{cases} \end{aligned}$$

where the penultimate equality uses part (b).

It can be shown that $\hat{\theta}_{ML}$ is complete.

(g) Find a uniform minimum variance unbiased estimator of θ.

A pdf $f_{ML}(\cdot|\theta)$ of $\hat{\theta}_{ML}$ is given by

$$f_{ML}(x|\theta) = \frac{n}{\theta^n}x^{n-1}1(0 \leq x \leq \theta).$$

As a consequence,

$$E_\theta(\hat{\theta}_{ML}) = \int_{-\infty}^{\infty} x f_{ML}(x|\theta)dx = \int_0^\theta \frac{n}{\theta^n}x^n dx = \frac{n}{\theta^n}\frac{1}{n+1}x^{n+1}\Big|_{x=0}^{\theta} = \frac{n}{n+1}\theta.$$

Therefore,

$$\hat{\theta}_{UMVU} = \frac{n+1}{n}\hat{\theta}_{ML}$$

is an unbiased estimator of θ. In fact, because $\hat{\theta}_{ML}$ is a complete sufficient statistic for θ, $\hat{\theta}_{UMVU}$ is a uniform minimum variance unbiased estimator of θ.

Let $\theta_0 > 0$ be some constant and consider the one-sided testing problem

$$H_0 : \theta \leq \theta_0 \qquad \text{vs.} \qquad H_1 : \theta > \theta_0.$$

(h) Consider a test which rejects H_0 if (and only if) $\hat{\theta}_{ML}/\theta_0 > c$, where c is some positive constant (possibly depending on θ_0). Find c such that the test has 5% size.

The size of the test is $\sup_{\theta \leq \theta_0} P_\theta(\hat{\theta}_{ML}/\theta_0 > c)$. *Using part (f),*

$$\begin{aligned} P_\theta(\hat{\theta}_{ML}/\theta_0 &> c) = 1 - P_\theta(\hat{\theta}_{ML} \leq c\theta_0) = 1 - F_{ML}(c\theta_0|\theta) = \begin{cases} 1 - (c\theta_0/\theta)^n & \text{if } 0 \leq c < \theta/\theta_0, \\ 0 & \text{if } c \geq \theta/\theta_0. \end{cases} \\ &= \max[1 - (c\theta_0/\theta)^n, 0] \end{aligned}$$

for any $c > 0$ *and any* $\theta > 0$. *Since* $P_\theta(\hat{\theta}_{ML}/\theta_0 > c)$ *is non-decreasing in* θ,

$$\sup_{\theta \leq \theta_0} P_\theta(\hat{\theta}_{ML}/\theta_0 > c) = P_{\theta_0}(\hat{\theta}_{ML}/\theta_0 > c) = \max(1 - c^n, 0),$$

so c *must satisfy*

$$c^n = 0.95 \qquad \Leftrightarrow \qquad c = \sqrt[n]{0.95}.$$

(i) Find the power function of the test derived in (h).

The power function is the function $\beta : \Theta \to [0, 1]$ *given by*

$$\beta(\theta) = P_\theta[\hat{\theta}_{ML}/\theta_0 > \sqrt[n]{0.95}] = \max[1 - (\sqrt[n]{0.95}\theta_0/\theta)^n, 0] = \max\left[1 - 0.95(\theta_0/\theta)^n, 0\right],$$

where the second equality uses a result obtained in part (h).