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 $(X, Y)'$ is a (bivariate) random vector with $E(Y^2) < \infty$ and suppose $m(\cdot)$ is a function satisfying $E[m(X)^2] < \infty$. If

$$E[Yg(X)] = E[m(X)g(X)]$$

for every function $g(\cdot)$ satisfying $E[g(X)^2] < \infty$, then

$$E[|Y - m(X)|^2] \leq E[|Y - h(X)|^2]$$

for every function $h(\cdot)$ satisfying $E[h(X)^2] < \infty$.

The statement is correct. Indeed,

$$\begin{aligned}
E[|Y - h(X)|^2] &= E[|Y - m(X) + m(X) - h(X)|^2] \\
&= E[|Y - m(X)|^2] + E[|m(X) - h(X)|^2] + 2E[\{Y - m(X)\}\{m(X) - h(X)\}] \\
&= E[|Y - m(X)|^2] + E[|m(X) - h(X)|^2] \\
&\geq E[|Y - m(X)|^2],
\end{aligned}$$

where the third equality uses

$$E[Yg(X)] = E[m(X)g(X)], \qquad g = m - h.$$

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

$$f_X(x|\theta) = c(\theta)\exp(x)1(x \leq \log\theta),$$

where $\theta \in \Theta = (0, \infty)$ is an unknown parameter, $1(\cdot)$ is the indicator function, and $c(\cdot)$ is some function.

(a) Show that

$$c(\theta) = \frac{1}{\theta}.$$

The function $c(\cdot)$ *is such that* $\int_{-\infty}^{\infty} f_X(x|\theta) = 1$, *so*

$$c(\theta) = \frac{1}{\int_{-\infty}^{\infty}\exp(x)1(x \leq \log\theta)dx} = \frac{1}{\theta},$$

where the second equality uses

$$\int_{-\infty}^{\infty}\exp(x)1(x \leq \log\theta)dx = \int_{-\infty}^{\log\theta}\exp(x)dx = \exp(x)|_{x=-\infty}^{\log\theta} = \theta.$$

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

For $x < \log\theta$, *we have:*

$$\int_{-\infty}^{x} f_X(r|\theta)\,dr = \frac{1}{\theta}\int_{-\infty}^{x}\exp(r)dr = \frac{1}{\theta}\exp(r)|_{r=0}^{x} = \frac{1}{\theta}\exp(x).$$

As a consequence,

$$F_X(x|\theta) = \begin{cases} \frac{1}{\theta}\exp(x) & \text{if } x < \log\theta, \\ 1 & \text{if } x \geq \log\theta. \end{cases}$$

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

Using integration by parts, we have:

$$E_\theta(X_i) = \int_{-\infty}^{\infty} x f_X(x|\theta)dx = \frac{1}{\theta}\int_{-\infty}^{\log\theta} x\exp(x)dx = \frac{1}{\theta}\, x\exp(x)|_{x=-\infty}^{\log\theta} - \frac{1}{\theta}\int_{-\infty}^{\log\theta} \exp(x)dx = \log\theta - 1.$$

In other words,

$$\theta = \exp[E_\theta(X_i) + 1],$$

so

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

Because $\exp(\cdot + 1)$ *is strictly convex, it follows from Jensen's inequality that*

$$E_\theta(\hat{\theta}_{MM}) = E_\theta[\exp(\bar{X} + 1)] > \exp[E_\theta(\bar{X} + 1)] = \exp[E_\theta(X_i + 1)] = \theta.$$

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*

$$L(\theta|X_1,\ldots,X_n) = \prod_{i=1}^{n} f_X(X_i|\theta) = \prod_{i=1}^{n}\{\frac{1}{\theta}\exp(X_i)1(X_i \le \log\theta)\} = \theta^{-n}\exp(\sum_{i=1}^{n} X_i)1(X_{(n)} \le \log\theta).$$

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

(e) Show that

$$\hat{\theta}_{ML} = \exp(\max_{1\le i\le n} X_i)$$

is the maximum likelihood estimator of θ.

Because $L(\theta|X_1,\ldots,X_n) = 0$ *for* $\theta < \exp(X_{(n)})$ *and because* θ^{-n} *is a decreasing function of* θ,

$$\arg\max_{\theta\in\Theta} L(\theta|X_1,\ldots,X_n) = \exp(X_{(n)}) = \hat{\theta}_{ML}.$$

(f) Find $F_{ML}(\cdot|\theta)$, the cdf of $\hat{\theta}_{ML}$.

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

$$P_\theta(\hat{\theta}_{ML} \leq x) = P_\theta[\exp(X_{(n)}) \leq x] = P_\theta(X_{(n)} \leq \log x) = F_X(\log x|\theta)^n = \begin{cases} (x/\theta)^n & \text{if } 0 \leq x < \theta, \\ 1 & \text{if } x \geq \theta, \end{cases}$$

where the last equality uses part (b).

In other words,

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

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 two-sided testing problem

$$H_0 : \theta = \theta_0 \qquad \text{vs.} \qquad H_1 : \theta \neq \theta_0.$$

(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.

Suppose $c \in [0, \theta_0]$. Then

$$\begin{aligned} P_\theta(|\hat{\theta}_{ML} - \theta_0| &> c) = P_\theta(\hat{\theta}_{ML} > \theta_0 + c) + P_\theta(\hat{\theta}_{ML} < \theta_0 - c) = 1 - F_{ML}(\theta_0 + c|\theta) + F_{ML}(\theta_0 - c|\theta) \\ &= 1 - \min\left(\frac{\theta_0 + c}{\theta}, 1\right)^n + \min\left(\frac{\theta_0 - c}{\theta}, 1\right)^n, \end{aligned}$$

where the last equality uses part (f). In particular,

$$P_{\theta_0}(|\hat{\theta}_{ML} - \theta_0| > c) = \left(\frac{\theta_0 - c}{\theta_0}\right)^n,$$

so the test has 5% size if c satisfies

$$\left(\frac{\theta_0 - c}{\theta_0}\right)^n = 0.05 \qquad \Leftrightarrow \qquad c = \theta_0(1 - \sqrt[n]{0.05}).$$

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

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

$$\beta(\theta) = P_\theta[|\hat{\theta}_{ML} - \theta_0| > \theta_0(1 - \sqrt[n]{0.05})] = 1 - \min\left(\frac{\theta_0(2 - \sqrt[n]{0.05})}{\theta}, 1\right)^n + \min\left(\frac{\theta_0 \sqrt[n]{0.05})}{\theta}, 1\right)^n.$$