Midterm Exam

October 18, 2006

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 (10 points). For *two* of the three statements below, determine whether or not the statement is correct, and give a *brief* (e.g., a bluebook page or less) justification for your answer.

(a) Suppose $X \sim Ber(p)$ for some $p \in (0,1)$; that is, suppose X is discrete with pmf

$$f(x|p) = p^x (1-p)^{1-x} 1(x \in \{0,1\}),$$

where $1(\cdot)$ is the indicator function. Then $E[\log f(X|q)] \leq E[\log f(X|p)]$ for every $q \in (0,1)$.

The statement is correct. To give a constructive proof, let

$$M(q) = E[\log f(X|q)] = (1-p)\log(1-q) + p\log q.$$

Now,

$$\frac{d}{dq}M(q) = -\frac{1-p}{1-q} + \frac{p}{q} = \frac{p-q}{q(1-q)}, \qquad \frac{d^2}{dq^2}M(q) = -\frac{1-p}{(1-q)^2} - \frac{p}{q^2} < 0,$$

so $\arg\max_{q\in(0,1)} M(q) = p$, *as claimed.*

An alternative proof, which has the advantage that it generalizes easily to more complicated distributions, starts by using Jensen's inequality (applicable because $\log(\cdot)$ *is concave) to show that*

$$\begin{aligned} E[\log f(X|q)] &= E\left(\log\left[\frac{f(X|q)}{f(X|p)} \cdot f(X|p)\right]\right) = E\left[\log\frac{f(X|q)}{f(X|p)}\right] + E[\log f(X|p)] \\ &\leq \log\left(E\left[\frac{f(X|q)}{f(X|p)}\right]\right) + E[\log f(X|p)]. \end{aligned}$$

Now, because the support of $f(\cdot|p)$ *does not depend on* $p \in (0,1)$,

$$E\left[\frac{f(X|q)}{f(X|p)}\right] = \sum_{x\in\mathbb{R}} \frac{f(x|q)}{f(x|p)} f(x|p) = \sum_{x\in\mathbb{R}} f(x|q) = 1, \qquad q \in (0,1).$$

As a consequence,

$$E[\log f(X|q)] \leq \log\left(\underbrace{E\left[\frac{f(X|q)}{f(X|p)}\right]}_{=1}\right) + E[\log f(X|p)] = E[\log f(X|p)].$$

(b) Suppose $X_1, \ldots, X_n$ is a random sample with $X_i \sim Ber(p)$, where $p \in \mathcal{P} \subset (0,1)$ is unknown. Because the pmf $f(\cdot|p)$ can be written as

$$f(x|p) = 1(x \in \{0,1\})(1-p)\exp\left[x \log\left(\frac{p}{1-p}\right)\right],$$

the sufficient (for p) statistic $\sum_{i=1}^n X_i$ is complete if and only if $\mathcal{P}$ contains an open interval.

The statement is incorrect. The "if" part is correct, but the "only if" part is not. For example, if $n=1$ and $\mathcal{P} = \{p_1, p_2\}$ with $p_1 \neq p_2$, then $\sum_{i=1}^n X_i = X_1$ and the condition

$$E_p\left[g\left(\sum_{i=1}^n X_i\right)\right] = E_p[g(X_1)] = 0 \qquad \forall p \in \mathcal{P}$$

reduces to the system of linear equations

$$\begin{pmatrix} 1-p_1 & p_1 \\ 1-p_2 & p_2 \end{pmatrix}\begin{pmatrix} g(0) \\ g(1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

whose unique solution is given by $g(0) = g(1) = 0$ (implying in particular that $P_p[g(\sum_{i=1}^n X_i) = 0] = 1$ for every $p \in \mathcal{P}$). (If $n > 1$, then a similar argument shows that $\sum_{i=1}^n X_i$ is complete whenever $\mathcal{P}$ has at least $n+1$ members.)

(c) 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. Then an unbiased estimator $\hat{\theta}$ of θ (with finite variance) is a uniform minimum variance unbiased estimator of θ if and only if

$$Cov_\theta\left(\tilde{\theta} - \hat{\theta}, \hat{\theta}\right) = 0 \qquad \forall \theta \in \Theta$$

for any other unbiased estimator $\tilde{\theta}$ of θ (with finite variance).

The statement is correct. Indeed, if $Cov_\theta\left(\tilde{\theta} - \hat{\theta}, \hat{\theta}\right) = 0$, then

$$Var_\theta\left(\tilde{\theta}\right) = Var_\theta\left[\left(\tilde{\theta} - \hat{\theta}\right) + \hat{\theta}\right] = \underbrace{Var_\theta\left(\tilde{\theta} - \hat{\theta}\right)}_{\geq 0} + Var_\theta\left(\hat{\theta}\right) + \underbrace{2Cov_\theta\left(\tilde{\theta} - \hat{\theta}, \hat{\theta}\right)}_{=0} \geq Var_\theta\left(\hat{\theta}\right),$$

showing the validity of the "if" part. For the "only if" part, notice that if $\hat{\theta}$ and $\tilde{\theta}$ are unbiased estimators of θ, then $(1-\lambda)\hat{\theta} + \lambda\tilde{\theta} = \hat{\theta} + \lambda\left(\tilde{\theta} - \hat{\theta}\right)$ is an unbiased estimator of θ, implying that if $\hat{\theta}$ is a uniform minimum variance unbiased estimator of θ, then

$$Var_\theta\left[\hat{\theta} + \lambda\left(\tilde{\theta} - \hat{\theta}\right)\right] = Var_\theta\left(\hat{\theta}\right) + \lambda^2 Var_\theta\left(\tilde{\theta} - \hat{\theta}\right) + 2\lambda Cov_\theta\left(\tilde{\theta} - \hat{\theta}, \hat{\theta}\right)$$

is minimized (with respect to λ) by setting $\lambda = 0$. A necessary condition for this to occur is that

$$\left.\frac{\partial}{\partial \lambda} Var_\theta\left[\hat{\theta} + \lambda\left(\tilde{\theta} - \hat{\theta}\right)\right]\right|_{\lambda=0} = 2Cov_\theta\left(\tilde{\theta} - \hat{\theta}, \hat{\theta}\right) = 0.$$

Problem 2 (40 points, each part receives equal weight). Let $X_1, \ldots, X_n$ be a random sample from a continuous distribution with cdf

$$F_X(x|\theta) = \left[a(\theta) + b(\theta)\, x^{-1/\theta}\right] 1(x > 1),$$

where $\theta \in \Theta = (0,1)$ is an unknown parameter, $a(\cdot)$ and $b(\cdot)$ are some functions (with argument θ), and $1(\cdot)$ is the indicator function.

(a) Show that $a(\theta) = 1$ and $b(\theta) = -1$.

To solve for $a(\theta)$ and $b(\theta)$, we will use the fact that if $F_X(\cdot|\theta)$ is a cdf, then

$$\lim_{x \to \infty} F_X(x|\theta) = 1$$

and (by right-continuity)

$$\lim_{x \downarrow 1} F_X(x|\theta) = F_X(1|\theta).$$

First, because $\lim_{x\to\infty} F_X(x|\theta) = a(\theta)$, the fact that $\lim_{x\to\infty} F_X(x|\theta) = 1$ implies that $a(\theta) = 1$. Next, because $\lim_{x\downarrow 1} F_X(x|\theta) = a(\theta) + b(\theta)$ should equal $F_X(1|\theta) = 0$, we have $b(\theta) = -a(\theta) = -1$.

(b) Find $f_X(\cdot|\theta)$, "the" pdf of X. Is $\{f_x(\cdot|\theta) : \theta \in \Theta\}$ an exponential family of pdfs?

The cdf is continuously differentiable on $(1, \infty)$, with derivative

$$\frac{\partial}{\partial x} F_X(x|\theta) = \frac{\partial}{\partial x}\left[1 - x^{-1/\theta}\right] = \frac{1}{\theta} x^{-(\theta+1)/\theta}.$$

Therefore, "the" pdf of X is given by

$$f_X(x|\theta) = \frac{1}{\theta} x^{-(\theta+1)/\theta} 1(x > 1) = 1(x > 1)\frac{1}{\theta}\exp\left(-\frac{\theta+1}{\theta}\log x\right),$$

where the last expression shows that $\{f_x(\cdot|\theta) : \theta \in \Theta\}$ is an exponential family of pdfs.

(c) Let $Y_i = \log(X_i)$. Find $F_Y(\cdot|\theta)$, the cdf of Y. Also, show that a pdf of Y is given by

$$f_Y(y|\theta) = \frac{1}{\theta}\exp\left(-\frac{1}{\theta}y\right) 1(y > 0).$$

Because $\log(\cdot)$ is strictly increasing, we have:

$$\begin{aligned} F_Y(y|\theta) &= P_\theta(Y_i \le y) = P_\theta[\log(X_i) \le y] = P_\theta[X_i \le \exp(y)] = F_X[\exp(y)|\theta] \\ &= \left[1 - \exp\left(-\frac{1}{\theta}y\right)\right] 1(y > 0). \end{aligned}$$

This cdf is continuously differentiable on $(0,\infty)$, *with derivative*

$$\frac{\partial}{\partial y}F_X(y|\theta)=\frac{\partial}{\partial x}\left[1-\exp\left(-\frac{1}{\theta}y\right)\right]=\frac{1}{\theta}\exp\left(-\frac{1}{\theta}y\right).$$

Therefore, "the" pdf of Y *is given by*

$$f_Y(y|\theta)=\frac{1}{\theta}\exp\left(-\frac{1}{\theta}y\right)1(y>0).$$

(d) Show that $E(X_i)=1/(1-\theta)$ and use this fact to derive a method moments estimator $\hat{\theta}_{MM,X}$ of θ. Is $\hat{\theta}_{MM,X}$ an unbiased estimator of θ?

We have:

$$E(X_i)=\int_{-\infty}^{\infty}xf_X(x|\theta)\,dx=\int_1^{\infty}\frac{1}{\theta}x^{-1/\theta}dx=-\frac{1}{1-\theta}x^{-(1-\theta)/\theta}\Big|_{x=1}^{\infty}=\frac{1}{1-\theta}.$$

As a consequence, $\theta=1-1/E(X_i)$ *and a method moments estimator of* θ *is given by* $\hat{\theta}_{MM,X}=1-1/\bar{X}$. *By Jensen's inequality (applied to the convex function* $1/x$*),*

$$E\left(\hat{\theta}_{MM,X}\right)=1-E\left(1/\bar{X}\right)<1-1/E\left(\bar{X}\right)=1-1/E(X_i)=\theta,$$

implying in particular that $\hat{\theta}_{MM,X}$ *a biased estimator of* θ.

(e) Show that $E(Y_i)=\theta$ and use this fact to derive a method moments estimator $\hat{\theta}_{MM,Y}$ of θ. Is $\hat{\theta}_{MM,Y}$ an unbiased estimator of θ?

We have:

$$E(Y_i)=\int_0^{\infty}\frac{1}{\theta}y\exp\left(-\frac{1}{\theta}y\right)dy=-y\exp\left(-\frac{1}{\theta}y\right)\Big|_{y=0}^{\infty}+\theta\underbrace{\int_0^{\infty}\frac{1}{\theta}\exp\left(-\frac{1}{\theta}y\right)dy}_{=\int_{-\infty}^{\infty}f_Y(y|\theta)dy=1}=\theta,$$

where the second equality uses integration by parts. Therefore, $\hat{\theta}_{MM,Y}=\bar{Y}$ *is an (unbiased) method moments estimator of* θ.

(f) Find the log likelihood function and show that the maximum likelihood estimator of θ is given by

$$\hat{\theta}_{ML} = \frac{1}{n}\sum_{i=1}^{n}\log\left(X_i\right).$$

The log likelihood function is

$$\begin{aligned}\ell\left(\theta|X_1,\ldots,X_n\right) &= \sum_{i=1}^{n}\log f_X\left(X_i|\theta\right) = \sum_{i=1}^{n}\log\left[\frac{1}{\theta}\underbrace{1\left(X_i>1\right)}_{\equiv 1}\exp\left(-\frac{\theta+1}{\theta}\log X_i\right)\right] \\ &= -n\log\theta - \left(1+\frac{1}{\theta}\right)\sum_{i=1}^{n}\log X_i.\end{aligned}$$

Because

$$\begin{aligned}\frac{\partial}{\partial\theta}\ell\left(\theta|X_1,\ldots,X_n\right) &= -\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}\log X_i, \\ \frac{\partial^2}{\partial\theta^2}\ell\left(\theta|X_1,\ldots,X_n\right) &= \frac{n}{\theta^2}-\frac{2}{\theta^3}\sum_{i=1}^{n}\log X_i = -\frac{1}{\theta}\frac{\partial}{\partial\theta}\ell\left(\theta|X_1,\ldots,X_n\right) - \underbrace{\frac{1}{\theta^3}\sum_{i=1}^{n}\log X_i}_{>0},\end{aligned}$$

$\hat{\theta}_{ML} = n^{-1}\sum_{i=1}^{n}\log X_i$ *is the unique solution to the equation* $\partial\ell\left(\theta|X_1,\ldots,X_n\right)/\partial\theta = 0$ *and furthermore satisfies* $\partial^2\ell\left(\hat{\theta}_{ML}|X_1,\ldots,X_n\right)/\partial\theta < 0$. *As a consequence,* $\hat{\theta}_{ML}$ *is the maximum likelihood estimator of* θ.

(g) Show that $\hat{\theta}_{ML}$ is a complete, sufficient statistic for θ and find a uniform minimum variance unbiased estimator of θ.

It follows from (f) and the factorization criterion that $\hat{\theta}_{ML}$ *is sufficient. Moreover, because*

$$\left\{\frac{\theta+1}{\theta}:\theta\in\Theta\right\} = \left\{1+\frac{1}{\theta}:0<\theta<1\right\} = (2,\infty)$$

contains an open set, it follows from (b) and the properties of exponential families that $\hat{\theta}_{ML}$ *is complete.*

To show that $\hat{\theta}_{ML}$ *is a uniform minimum variance unbiased estimator of* θ*, it therefore suffices to show that* $\hat{\theta}_{ML}$ *is unbiased. In turn, this result follows from part (e) and the fact that* $\hat{\theta}_{ML} = \hat{\theta}_{MM,Y}$.

(h) Compute the Cramér-Rao bound (on the variance of unbiased estimators of θ). Is this bound attained by the estimator from (g)?

The Fisher information is given by

$$\mathcal{I}(\theta) = Var\left[\frac{\partial}{\partial\theta}\ell(\theta|X_1,\ldots,X_n)\right] = Var\left(-\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}\log X_i\right) = \frac{n}{\theta^4}Var(\log X_i) = \frac{n}{\theta^2},$$

where the last equality uses the fact that

$$Var(\log X_i) = Var(Y_i) = E\left(Y_i^2\right) - E(Y_i)^2 = \theta^2,$$

a result which itself follows from (e) and the fact that

$$E\left(Y_i^2\right) = \int_0^\infty \frac{1}{\theta}y^2\exp\left(-\frac{1}{\theta}y\right)dy = -y^2\exp\left(-\frac{1}{\theta}y\right)\Big|_{y=0}^{\infty} + 2\theta\underbrace{\int_0^\infty \frac{1}{\theta}y\exp\left(-\frac{1}{\theta}y\right)dy}_{=E(Y_i)=\theta} = 2\theta^2,$$

where the second equality uses integration by parts. The Cramér-Rao bound is $I(\theta)^{-1} = \theta^2/n$*, which is attained by* $\hat{\theta}_{ML}$ *because*

$$Var\left(\hat{\theta}_{ML}\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n}\log X_i\right) = \frac{Var(\log X_i)}{n} = \frac{\theta^2}{n}.$$