Practice Midterm II

STA 131B
Winter 2016
University of California, Davis

Exam Rules: This exam is closed book and closed notes. Use of calculators, cell phones or other communication devices (including smartwatch) is not allowed. You must show all of your work to receive credit. You will have 50 minutes to complete the exam.

Note: The practice exam is longer than the actual exam so you will have more to practice.

Name : ______________________

ID : ______________________

Signature : ______________________

1. Let $X_1, \ldots, X_n$ be a random sample from a distribution with p.d.f.

$$f(x|\theta_1, \theta_2) = \frac{1}{\theta_2} e^{-(x-\theta_1)/\theta_2},$$

for $x \geq \theta_1$, $-\infty < \theta_1 < \infty$, and $\theta_2 > 0$.

a) Find jointly sufficient statistics (T_1, T_2) where θ_1 and θ_2 are both unknown.

Solution:

$$f_n(\mathbf{x}|\theta_1) = \frac{1}{\theta_2^n} \exp\{-\frac{1}{\theta_2} \sum_{i=1}^{n} X_i + n\frac{\theta_1}{\theta_2}\} 1_{[\theta_1, \infty)}(\min_{i=1,\ldots,n}(X_i)).$$

So $(T_1, T_2) = (\sum_{i=1}^n X_i, \min_{i=1,\ldots,n}(X_i))$ is jointly sufficient for (θ_1, θ_2) by factorization theorem.

b) If θ_2 is known, find a sufficient statistic for θ_1.

Solution:

$$f_n(\mathbf{x}|\theta_1,\theta_2) = \frac{1}{\theta_2^n}\exp\{-\frac{1}{\theta_2}\sum_{i=1}^{n} X_i\}\exp\{\frac{n}{\theta_2}\theta_1\}1_{[\theta_1,\infty)}(\min_{i=1,\dots,n}(X_i)).$$

So $T = \min_{i=1,\dots,n}(X_i)$ is sufficient for θ_1 by factorization theorem.

c) Is the M.L.E. for θ_1 minimal sufficient (assuming that θ_2 is known)? Justify your answer.

Solution: Write $f_n(\mathbf{x}|\theta_1)$ without factors not involving θ_1:

$$f_n(\mathbf{x}|\theta_1) \propto \exp\{\frac{n}{\theta_2}\theta_1\}, \quad \text{if } \theta_1 \le \min_{i=1,\dots,n}(X_i).$$

Note $\exp\{a\theta_1\}$ is increasing in θ_1 for all $a > 0$ (check this by plotting a graph or take the first derivative, and then take $a = \frac{n}{\theta_2}$). Due to the restriction $\theta_1 \le \min_{i=1,\dots,n}(X_i)$, the MLE $\hat{\theta}_1 = \min_{i=1,\dots,n}(X_i)$. By book theorem 7.8.3, since the MLE $\hat{\theta}_1$ is sufficient (shown in b.), it is minimal sufficient.

d) Does the Fisher information for θ_1 exist? Explain briefly.

Solution: The Fisher information for θ_1 does not exist because the support $[\theta_1,\infty)$ of the distribution of X_i depends on θ_1.

2. Suppose $X_1, \ldots, X_n$ form a random sample from a distribution with known mean μ and unknown variance σ^2.

   (a) Show that the following estimator is an unbiased estimator of σ^2:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Solution:

$$\begin{aligned} E(\hat{\sigma}^2) &= E(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2) \\ &= \frac{1}{n}\sum_{i=1}^{n}E(X_i - \mu)^2 \\ &= E(X_1 - \mu)^2 \quad \text{since the samples are iid} \\ &= \sigma^2. \end{aligned}$$

So $\hat{\sigma}^2$ is unbiased for σ^2.

   (b) Show that the Fisher information for $\theta = \sigma^2$ based on a random sample from $N(\mu, \sigma^2)$ is $\frac{n}{2\sigma^4}$.

Solution: Write the density and its second derivative for a single sample

$$\begin{aligned} f(x|\theta) &= \frac{1}{\sqrt{2\pi\theta}}\exp\{-\frac{(x-\mu)^2}{2\theta}\} \\ \log(f(x|\theta)) &= -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\theta) - \frac{(x-\mu)^2}{2\theta} \\ \frac{\partial^2}{\partial\theta^2}\log(f(x|\theta)) &= \frac{1}{2\theta^2} - \frac{(x-\mu)^2}{\theta^3}. \end{aligned}$$

By theorem (8.8.3) we have $I(\theta) = -E(\frac{\partial^2}{\partial\theta^2}\log(f(x|\theta))) = \frac{1}{2\theta^2} = \frac{1}{2\sigma^4}$. Therefore based on a random sample $I_n(\theta) = nI(\theta) = \frac{n}{2\sigma^4}$.

   (c) Based on the results in (b), what can you say about the efficiency of the estimator in (a)? Explain clearly but you may use the fact that the fourth central moment $E(X-\mu)^4$ of $N(\mu, \sigma^2)$ is $3\sigma^4$.

Solution: $\hat{\sigma}^2$ has mean σ^2 and

$$\begin{aligned}
\operatorname{var}(\hat{\sigma}^2) &= \operatorname{var}\Big(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\Big) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{var}((X_i-\mu)^2), \quad \text{since the samples are independent} \\
&= \frac{1}{n}\operatorname{var}((X_1-\mu)^2), \quad \text{since the samples are identical} \\
&= \frac{1}{n}\left[E[(X_1-\mu)^4]-(E[(X_1-\mu)^2])^2\right] \\
&= \frac{1}{n}(3\sigma^4-\sigma^4) \\
&= \frac{2\sigma^4}{n} = 1/I_n(\theta).
\end{aligned}$$

Therefore the Cramer-Rao lower bound (8.8.14) is reached and thus $\hat{\sigma}^2$ is efficient.

3. Suppose we draw a sample $X_1, \ldots, X_n$ of size n from the distribution $N(\mu_1, \sigma_1^2)$ and a sample $Y_1, \ldots, Y_m$ of size m from the distribution $N(\mu_2, \sigma_2^2)$. Assume $\sigma_1^2 = 4\sigma_2^2$ and $\mu_1 = \mu_2 =: \mu$. We aim to estimate $\theta = \mu$ and use the estimator $\theta_\alpha = \alpha \bar{X}_n + (1-\alpha)\bar{Y}_m$, where $\bar{X}$ denotes the sample means.

   (a) Obtain the bias of θ_α for all α.

   Solution:

$$
\begin{aligned}
E(\theta_\alpha) &= E(\alpha \bar{X}_n + (1-\alpha)\bar{Y}_m) \\
&= \alpha E(\bar{X}_n) + (1-\alpha) E(\bar{Y}_m) \\
&= \alpha \mu_1 + (1-\alpha)\mu_2 \\
&= \mu_1 = \mu
\end{aligned}
$$

   for all α, since $\mu_1 = \mu_2$. So θ_α is unbiased for μ.

   (b) Obtain the MSE of θ_α for all α.

   Solution: $\mathrm{MSE}(\theta_\alpha) = \mathrm{bias}(\theta_\alpha)^2 + \mathrm{var}(\theta_\alpha) = \mathrm{var}(\theta_\alpha)$ by (a). Then

$$
\begin{aligned}
\mathrm{var}(\theta_\alpha) &= \mathrm{var}(\alpha \bar{X}_n + (1-\alpha)\bar{Y}_m) \\
&= \alpha^2 \,\mathrm{var}(\bar{X}_n) + (1-\alpha)^2 \,\mathrm{var}(\bar{Y}_m) \\
&= \alpha^2 \frac{\sigma_1^2}{n} + (1-\alpha)^2 \frac{\sigma_2^2}{m}.
\end{aligned}
$$

   (c) For what value of α is the MSE minimized? What is the value of the MSE at the minimum?

   Solution:

$$
\frac{\partial}{\partial \alpha} \mathrm{MSE}(\theta_\alpha) = 2\alpha \frac{\sigma_1^2}{n} - 2(1-\alpha)\frac{\sigma_2^2}{m}.
$$

   Set $\frac{\partial}{\partial \alpha} \mathrm{MSE}(\theta_\alpha) = 0$ we have

$$
\alpha^* = \frac{\sigma_2^2/m}{\sigma_2^2/m + \sigma_1^2/n}.
$$

   Check that α^* is a minimizer of $\mathrm{MSE}(\theta_\alpha)$ because $\frac{\partial^2}{\partial \alpha^2} \mathrm{MSE}(\theta_\alpha) = 2\frac{\sigma_1^2}{n} + 2\frac{\sigma_2^2}{m} > 0$. The MSE at the minimum is

$$
\mathrm{MSE}(\theta_{\alpha^*}) = \frac{(\sigma_1^2/n) \cdot (\sigma_2^2/m)}{\sigma_1^2/n + \sigma_2^2/m}.
$$

(d) How does this MSE compare to that of the estimator that is obtained when you pool the two samples into one and take the sample average as estimator?

Solution: The pooled estimator is

$$\frac{n\bar{X}_n + m\bar{Y}_m}{m+n} = \frac{n}{m+n}\bar{X}_n + \frac{m}{m+n}\bar{Y}_m = \theta_{\alpha_1},$$

where $\alpha_1 = \frac{n}{m+n}$. By (c), since θ_{α^*} minimized $\mathrm{MSE}(\theta_\alpha)$ over all α, we have $\mathrm{MSE}(\theta_{\alpha^*}) \leq \mathrm{MSE}(\theta_{\alpha_1})$. Since $\mathrm{MSE}(\theta_\alpha)$ is quadratic in α, the equality holds if and only if $\alpha^* = \alpha_1$, ie $\frac{n}{m+n} = \frac{\sigma_2^2/m}{\sigma_2^2/m+\sigma_1^2/n}$.

(e) Now assume $\mu_1 = 2\mu_2$, $n = 2m$ and that the target is the parameter $\theta = \frac{\mu_1+\mu_2}{2}$. Redo (b) and (c) under these assumptions, for the estimator θ_α of θ.

Solution:

$$\begin{aligned}\mathrm{bias}(\theta_\alpha) &= E(\theta_\alpha) - \theta \\ &= E(\alpha\bar{X}_n + (1-\alpha)\bar{Y}_m) - \theta \\ &= \alpha\mu_1 + (1-\alpha)\mu_2 - \theta \\ &= (\alpha - \frac{1}{2})\mu_2,\end{aligned}$$

since $\mu_1 = 2\mu_2$ and $\theta = \frac{\mu_1+\mu_2}{2} = 3\mu_2/2$. The variance is as in (b), except $n = 2m$:

$$\mathrm{var}(\theta_\alpha) = \alpha^2\frac{\sigma_1^2}{2m} + (1-\alpha)^2\frac{\sigma_2^2}{m}.$$

So

$$\begin{aligned}\mathrm{MSE}(\theta_\alpha) &= \ \mathrm{bias}(\theta_\alpha)^2 + \ \mathrm{var}(\theta_\alpha) \\ &= (\alpha - 1/2)^2\mu_2^2 + \alpha^2\frac{\sigma_1^2}{2m} + (1-\alpha)^2\frac{\sigma_2^2}{m}\end{aligned}$$

$$\frac{\partial}{\partial\alpha}\mathrm{MSE}(\theta_\alpha) = 2(\alpha-1/2)\mu_2^2 + 2\alpha\frac{\sigma_1^2}{2m} - 2(1-\alpha)\frac{\sigma_2^2}{m}.$$

Set $\frac{\partial}{\partial\alpha}\mathrm{MSE}(\theta_\alpha) = 0$ we have

$$\alpha^{**} = \frac{\mu_2^2/2 + \sigma_2^2/m}{\mu_2^2 + \sigma_1^2/2m + \sigma_2^2/m}.$$

Check that α^{**} is a minimizer of $\mathrm{MSE}(\theta_\alpha)$ because $\frac{\partial^2}{\partial\alpha^2}\mathrm{MSE}(\theta_\alpha) = 2\mu_2^2 + \sigma_1^2/m + 2\sigma_2^2/m > 0$. The MSE at the minimum is

$$\mathrm{MSE}(\theta_{\alpha^{**}}) = \frac{4(\mu_2^2 + \sigma_1^2/2m + \sigma_2^2/m)(\mu_2^2/4 + \sigma_2^2/m) - (\mu_2^2 + 2\sigma_2^2/m)^2}{4(\mu_2^2 + \sigma_1^2/2m + \sigma_2^2/m)}.$$

4. Suppose $X_1, \ldots, X_n$ form a random sample from a distribution with p.d.f. $f(x|\theta) = \theta x^{\theta-1}$, for $0 < x < 1$ and $\theta > 0$.

a) Find the M.L.E. of θ.

Solution:

$$f_n(\mathbf{x}|\theta) = \theta^n \exp\{(\theta - 1)\sum_{i=1}^{n} \log(X_i)\}$$

$$\log(f_n(\mathbf{x}|\theta)) = n\log(\theta) + (\theta - 1)\sum_{i=1}^{n} \log(X_i)$$

$$\frac{\partial}{\partial\theta}\log(f_n(\mathbf{x}|\theta)) = \frac{n}{\theta} + \sum_{i=1}^{n} \log(X_i).$$

Set $\frac{\partial}{\partial\theta}\log(f_n(\mathbf{x}|\theta)) = 0$ we have MLE $\hat{\theta} = \frac{-n}{\sum_{i=1}^{n}\log(X_i)}$. This is a maximum since

$$\frac{\partial^2}{\partial\theta^2}\log(f_n(\mathbf{x}|\theta)) = \frac{-n}{\theta^2} < 0,$$

for all $\theta > 0$.

b) Is the M.L.E. in (a) minimal sufficient? Justify your answer.

Solution: Yes. $\hat{\theta}$ is sufficient because it is a one to one function of $\sum_{i=1}^{n}\log(X_i)$, which is sufficient. Then apply theorem 7.8.3 to conclude the MLE $\hat{\theta}$ is minimal sufficient.

c) Is the sample mean admissible for estimating θ. Explain briefly.

Solution: No. The sample mean $\bar{X}$ is not a function of sufficient statistic $\sum_{i=1}^n \log(X_i)$. By Rao-Blackwell theorem (book theorem 7.9.1) $\delta^*(X) = E(\bar{X}|\sum_{i=1}^n \log(X_i))$ dominates $\bar{X}$. Therefore $\bar{X}$ is inadmissible.

d) Use that fact that $E(\log X_i) = \frac{1}{\theta}$ to find an UMVUE of $\frac{1}{\theta}$.

Solution: Use $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n \log(X_i)$. Then

$$E(\hat{\theta}) = E(\frac{1}{n}\sum_{i=1}^n \log(X_i)) = 1/n\sum_{i=1}^n E(\log(X_i)) = 1/\theta$$

by the hint, and thus $\hat{\theta}$ is unbiased for $\frac{1}{\theta}$. Also,

$$E(\log(X_i)^2) = \int_0^1 \log(x)^2\theta x^{\theta-1}dx = \int_0^\infty y^2\theta\exp\{-\theta y\} = 2/\theta^2,$$

where the second inequality is by transformation $y = -\log(x)$, and the third can be obtained by the moment formula for an exponential distribution with parameter θ. So $\mathrm{var}(\log(X_i)) = E(\log(X_i)^2) - [E(\log(X_i))]^2 = 1/\theta^2$. Therefore

$$\mathrm{var}(\hat{\theta}) = 1/n^2\sum_{i=1}^n \mathrm{var}(\log(X_i)) = 1/n\theta^2.$$

Meanwhile, the Fisher information for a single observation is

$$I(\theta) = -E[\frac{\partial^2}{\partial\theta^2}\log(\theta X^{\theta-1})] = \frac{1}{\theta^2}.$$

So the Cramer-Rao lower bound is

$$\frac{(\frac{\partial}{\partial\theta}\frac{1}{\theta})^2}{nI(\theta)} = \frac{1/\theta^4}{n/\theta^2} = \frac{1}{n\theta^2}.$$

Notice that our estimator $\hat{\theta}$ reaches the Cramer-Rao lower bound, so by the information inequality theorem (theorem 8.8.3) it has smaller variance than any other unbiased estimator. Therefore $\hat{\theta}$ is a UMVUE of $\frac{1}{\theta}$.