Ma 494 — Theoretical Statistics

Solutions for Problem Set #3 — Due February 22, 2010

Prof. Sawyer — Washington University

1. By the hint and Definition 4.6.2 p329, Y_i are gamma distributed with parameters r and $\lambda = 1/\theta$, in the notation of Section 4.6 p329. Thus by Theorem 4.6.3 p330, $E(Y_i) = r/\lambda = r\theta$ and $\mathrm{Var}(Y_i) = r/\lambda^2 = r\theta^2$.

(i) Since $E(Y_j) = r\theta$, $E\big((1/r)\overline{Y}\big) = (1/r)(1/n)\sum_{j=1}^n E(Y_j) = \big(1/(rn)\big)nr\theta = \theta$ and $(1/r)\overline{Y}$ is an unbiased estimator of θ.

(ii) Since $\log f(Y,\theta) = -\log\big((r-1)!\big) - r\log(\theta) + (r-1)\log(Y) - Y/\theta$, the j^{th} score is

$$S(Y_j,\theta) \;=\; \frac{\partial}{\partial\theta}\log f(Y_j,\theta) \;=\; -\frac{r}{\theta} \;+\; \frac{Y_j}{\theta^2} \;=\; \frac{Y_j - r\theta}{\theta^2}$$

The Fisher information $I(\theta) = r/\theta^2$ by arguing EITHER $I(\theta) = \mathrm{Var}\big(S(Y_j,\theta)\big) = \mathrm{Var}(Y/\theta^2) = \mathrm{Var}(Y)/\theta^4 = r\theta^2/\theta^4 = r/\theta^2$ OR ELSE by arguing

$$T(Y_j,\theta) \;=\; \frac{\partial^2}{\partial\theta^2}\log f(Y_j,\theta) \;=\; \frac{r}{\theta^2} \;-\; \frac{2Y_j}{\theta^3} \;=\; -\frac{2Y_j - r\theta}{\theta^3}$$

Then $I(\theta) = -E\big(T(Y_j,\theta)\big) = \big(2E(Y_j) - r\theta\big)/\theta^3 = r\theta/\theta^3 = r/\theta^2$. Thus the Cramér-Rao lower bound is $1/\Big(nI(\theta)\Big) = \theta^2/(nr)$. In comparison,

$$\mathrm{Var}\left(\frac{1}{r}\overline{Y}\right) \;=\; \frac{1}{n}\frac{1}{r^2}\,\mathrm{Var}(Y_j) \;=\; \frac{r\theta^2}{nr^2} \;=\; \frac{\theta^2}{nr}$$

Since this attains the Cramér-Rao lower bound, $(1/r)\overline{Y}$ is a minimum variance unbiased estimator of θ.

2. (a) The first step is to write $f(y,\theta) = e^{-(y-\theta)}I_{(\theta,\infty)}(y)$. Then the likelihood is

$$L(\theta,Y_1,\ldots,Y_n) \;=\; \prod_{j=1}^n e^{-(Y_j-\theta)}I_{(\theta,\infty)}(Y_j) \;=\; \exp\left(-\sum_{j=1}^n (Y_j-\theta)\right)\prod_{j=1}^n I_{(\theta,\infty)}(Y_j)$$

The product of indicator functions $\prod_{j=1}^n I_{(\theta,\infty)}(Y_j)$ equals zero unless $\theta < Y_j < \infty$ for $1 \le j \le n$, or equivalently unless $\theta < Y_{\min} < \infty$. Hence we can write the likelihood as

$$L(\theta,Y_1,\ldots,Y_n) \;=\; e^{n\theta}I_{(\theta,\infty)}(Y_{\min})\exp\left(-\sum_{j=1}^n Y_j\right) \;=\; g(\theta,Y_{\min})A(Y_1,\ldots,Y_n)$$

for $g(\theta, y) = e^{n\theta} I_{(\theta,\infty)}(y)$ and $A(y_1, \ldots, y_n) = \exp\Big(-\sum_{j=1}^n Y_j\Big)$. This implies that $Y_{\min}$ is a sufficient statistic for θ.

(b) It follows from the last displayed equation that $L(\theta, Y_1, \ldots) > 0$ if $\theta < Y_{\min}$ and $L(\theta, Y_1, \ldots) = 0$ if $Y_{\min} \le \theta$. If $Y_{\max}$ were a sufficient statistic, then

$$L(\theta, Y_1, \ldots, Y_n) \ = \ h(\theta, Y_{\max}) B(Y_1, \ldots, Y_n)$$

for functions $h(\theta, y)$ and $B(y_1, \ldots, y_n)$. Since we can find $\theta < Y_{\min}$ for any $Y = (Y_1, \ldots, Y_n) \in R^n$, it follows that $B(Y) > 0$ for all $Y \in R^n$. Thus $h(\theta, Y_{\max}) > 0$ if $\theta < Y_{\min}$ and $h(\theta, Y_{\max}) = 0$ if $Y_{\min} \le \theta$. Now assume $n = 2$, $Y_1 < Y_2$, and fix θ. Then $Y_{\min} = Y_1$, $Y_{\max} = Y_2$, and $h(\theta, Y_{\max}) = h(\theta, Y_2)$ as long as $Y_1 < Y_2$. Thus $h(\theta, Y_2) = 0$ if $Y_1 < \theta < Y_2$ but $h(\theta, Y_2) > 0$ if $\theta < Y_1 < Y_2$, which is a contradiction. Thus $Y_{\max}$ cannot be a sufficient statistic for θ.

3. If $f(x, \theta) = e^{K(x)p(\theta)+q(\theta)} A(x)$, the likelihood is

$$\begin{aligned} L(\theta, X_1, \ldots, X_n) \ &= \ \prod_{j=1}^n f(X_j, \theta) \ = \ \prod_{j=1}^n \exp\Big(K(X_j)p(\theta) + q(\theta\Big) A(X_j) \\ &= \ \exp\Bigg(\sum_{j=1}^n \Big(K(X_j)p(\theta) + q(\theta)\Big)\Bigg) \prod_{j=1}^n A(X_j) \\ &= \ \exp\Bigg(nq(\theta) + \Bigg(\sum_{j=1}^n K(X_j)\Bigg)p(\theta)\Bigg) \prod_{j=1}^n A(X_j) \\ &= \ g\Bigg(\theta, \sum_{j=1}^n K(X_j)\Bigg) B(X_1, \ldots, X_n) \end{aligned}$$

for $g(\theta, y) = e^{nq(\theta)+yp(\theta)}$ and $B(x_1, \ldots, x_n) = \prod_{j=1}^n A(x_j)$. This implies that $S(X) = \sum_{j=1}^n K(X_j)$ is a sufficient statistic for θ.

4. (a) If $f(y, \theta) = \theta/(1 + y)^{\theta+1}$, the likelihood is

$$L(\theta, Y_1, \ldots, Y_n) \ = \ \prod_{j=1}^n \frac{\theta}{(1 + Y_j)^{\theta+1}} \ = \ \theta^n \Bigg(\prod_{j=1}^n \frac{1}{1 + Y_j}\Bigg)^{\theta+1} \ = \ \frac{\theta^n}{\Big(\prod_{j=1}^n (1 + Y_j)\Big)^{\theta+1}}$$

This implies that $S(Y) = \prod_{j=1}^n (1 + Y_j)$ is a sufficient statistic for θ.

(b) Writing x in place of y,

$$f(x, \theta) = \exp\Big(\log(1 + x)(-\theta - 1) + \log\theta\Big)$$

This is of the form $f(x, \theta) = e^{K(x)p(\theta)+q(\theta)} A(x)$ for $K(x) = \log(1+x)$, $p(\theta) = -\theta-1$, and $q(\theta) = \log\theta$. Since $\sum_{j=1}^n K(X_j) = \sum_{j=1}^n \log(1 + X_j) = \log\Big(\prod_{j=1}^n (1 + X_j)\Big)$, this is consistent with part (a).

5. If $f(x,\theta) = (1/2)I_{(\theta-1,\theta+1)}(x)$, the likelihood is

$$L(\theta, X_1, \ldots, X_n) \;=\; \prod_{j=1}^{n} f(X_j, \theta) \;=\; \frac{1}{2^n}\prod_{j=1}^{n} I_{(\theta-1,\theta+1)}(X_j)$$

The likelihood is zero unless $\theta - 1 < X_j < \theta + 1$ for $1 \le j \le n$, or equivalently unless $\theta - 1 < X_{\min} \le X_{\max} < \theta + 1$. Since $X_{\min} \le X_{\max}$, this is equivalent to $X_{\max} - 1 < \theta$ and $\theta < X_{\min} + 1$. Thus

$$L(\theta, X_1, \ldots, X_n) \;=\; \frac{1}{2^n} I_{(X_{\max}-1, X_{\min}+1)}(\theta) \;=\; g(X_{\min}, X_{\max}, \theta)$$

for $g(x_1, x_2, \theta) = (1/2^n)I_{(x_2-1,x_1+1)}(\theta)$. Thus $(X_{\min}, X_{\max})$ is a vector-valued sufficient statistic for θ.

Remarks: (1) Since $\theta-1 < X_j < \theta+1$ for all j, it follows that $X_{\max} - X_{\min} < 2$ and $X_{\max} - 1 < X_{\min} + 1$. Thus the arguments in the indicator function in the last displayed equation are not a typo.

(2) Since $L(\theta, X_1, \ldots, X_n)$ is always either 0 or $1/2^n$, a statistic $T(X_1, \ldots, X_n)$ is a maximum-likelihood estimator as long as $X_{\max} - 1 < T(X_1, \ldots, X_n) < X_{\min} + 1$ for all $X = (X_1, \ldots, X_n) \in R^n$. Since any nonempty interval $I = (a, b)$ contains the average of its endpoints (that is, $(a+b)/2 \in I$), it follows that $T(X) = \big((X_{\max} - 1) + (X_{\min} + 1)\big)/2 = (X_{\min} + X_{\max})/2$ is a maximum likelihood estimator for θ.