Stat Inf HW2 solutions 3/25/08

6.12 a. Use Theorem 6.2.13 and write

$$\frac{f(x,n|\theta)}{f(y,n'|\theta)} = \frac{f(x|\theta,N=n)P(N=n)}{f(y|\theta,N=n')P(N=n')} = \frac{\binom{n}{x}\theta^x(1-\theta)^{n-x}p_n}{\binom{n'}{y}\theta^y(1-\theta)^{n'-y}p_{n'}} = \theta^{x-y}(1-\theta)^{n-n'-x+y}\frac{\binom{n}{x}p_n}{\binom{n'}{y}p_{n'}}.$$

The last ratio does not depend on $\theta$. The other terms are constant as a function of $\theta$ if and only if $n = n'$ and $x = y$. So $(X, N)$ is minimal sufficient for $\theta$. Because $P(N = n) = p_n$ does not depend on $\theta$, $N$ is ancillary for $\theta$. The point is that although $N$ is independent of $\theta$, the minimal sufficient statistic contains $N$ in this case. A minimal sufficient statistic may contain an ancillary statistic.

b.

$$\mathrm{E}\left(\frac{X}{N}\right) = \mathrm{E}\left(\mathrm{E}\left(\frac{X}{N}\middle|N\right)\right) = \mathrm{E}\left(\frac{1}{N}\mathrm{E}(X\mid N)\right) = \mathrm{E}\left(\frac{1}{N}N\theta\right) = \mathrm{E}(\theta) = \theta.$$

$$\mathrm{Var}\left(\frac{X}{N}\right) = \mathrm{Var}\left(\mathrm{E}\left(\frac{X}{N}\middle|N\right)\right) + \mathrm{E}\left(\mathrm{Var}\left(\frac{X}{N}\middle|N\right)\right) = \mathrm{Var}(\theta) + \mathrm{E}\left(\frac{1}{N^2}\mathrm{Var}(X\mid N)\right)$$

$$= 0 + \mathrm{E}\left(\frac{N\theta(1-\theta)}{N^2}\right) = \theta(1-\theta)\mathrm{E}\left(\frac{1}{N}\right).$$

We used the fact that $X|N \sim \text{binomial}(N, \theta)$.

6.15 a. The parameter space consists only of the points $(\theta, \nu)$ on the graph of the function $\nu = a\theta^2$. This quadratic graph is a line and does not contain a two-dimensional open set.

b. Use the same factorization as in Example 6.2.9 to show $(\bar{X}, S^2)$ is sufficient. $\mathrm{E}(S^2) = a\theta^2$ and $\mathrm{E}(\bar{X}^2) = \mathrm{Var}\bar{X} + (\mathrm{E}\bar{X})^2 = a\theta^2/n + \theta^2 = (a+n)\theta^2/n$. Therefore,

$$\mathrm{E}\left(\frac{n}{a+n}\bar{X}^2 - \frac{S^2}{a}\right) = \left(\frac{n}{a+n}\right)\left(\frac{a+n}{n}\theta^2\right) - \frac{1}{a}a\theta^2 = 0, \text{ for all } \theta.$$

Thus $g(\bar{X}, S^2) = \frac{n}{a+n}\bar{X}^2 - \frac{S^2}{a}$ has zero expectation so $(\bar{X}, S^2)$ not complete.

6.18 The distribution of $Y = \sum_i X_i$ is Poisson$(n\lambda)$. Now

$$\mathrm{E}g(Y) = \sum_{y=0}^{\infty} g(y)\frac{(n\lambda)^y e^{-n\lambda}}{y!}.$$

If the expectation exists, this is an analytic function which cannot be identically zero.

6.20 The pdfs in b), c), and e) are exponential families, so they have complete sufficient statistics from Theorem 6.2.25. For a), $Y = \max\{X_i\}$ is sufficient and

$$f(y) = \frac{2n}{\theta^{2n}}y^{2n-1}, \quad 0 < y < \theta.$$

For a function $g(y)$,

$$\mathrm{E}\,g(Y) = \int_0^\theta g(y)\frac{2n}{\theta^{2n}}y^{2n-1}\,dy = 0 \text{ for all } \theta \text{ implies } g(\theta)\frac{2n\theta^{2n-1}}{\theta^{2n}} = 0 \text{ for all } \theta$$

by taking derivatives. This can only be zero if $g(\theta) = 0$ for all $\theta$, so $Y = \max\{X_i\}$ is complete. For d), the order statistics are minimal sufficient. This is a location family. Thus, by Example 6.2.18 the range $R = X_{(n)} - X_{(1)}$ is ancillary, and its expectation does not depend on $\theta$. So this sufficient statistic is not complete.

(6.20a, use Liebniz rule for the derivative w.r.t. theta)
(correction to 6.20d, it is in the exponential family)

6.21 a. $X$ is sufficient because it is the data. To check completeness, calculate

$$\mathrm{E}g(X) = \frac{\theta}{2}g(-1) + (1-\theta)g(0) + \frac{\theta}{2}g(1).$$

If $g(-1) = g(1)$ and $g(0) = 0$, then $\mathrm{E}g(X) = 0$ for all $\theta$, but $g(x)$ need not be identically 0. So the family is not complete.

b. $|X|$ is sufficient by Theorem 6.2.6, because $f(x|\theta)$ depends on $x$ only through the value of $|x|$. The distribution of $|X|$ is Bernoulli, because $P(|X| = 0) = 1 - \theta$ and $P(|X| = 1) = \theta$. By Example 6.2.22, a binomial family (Bernoulli is a special case) is complete.

c. Yes, $f(x|\theta) = (1-\theta)(\theta/(2(1-\theta)))^{|x|} = (1-\theta)e^{|x|\log[\theta/(2(1-\theta)]}$, the form of an exponential family.

(correction to 6.21a, g(-1)=-g(1) )

6.30 a. From Exercise 6.9b, we have that $X_{(1)}$ is a minimal sufficient statistic. To check completeness compute $f_{Y_1}(y)$, where $Y_1 = X_{(1)}$. From Theorem 5.4.4 we have

$$f_{Y_1}(y) = f_X(y)\,(1-F_X(y))^{n-1} n = e^{-(y-\mu)}\left[e^{-(y-\mu)}\right]^{n-1} n = ne^{-n(y-\mu)}, \quad y > \mu.$$

Now, write $\mathrm{E}_\mu\, g(Y_1) = \int_\mu^\infty g(y)ne^{-n(y-\mu)}\,dy$. If this is zero for all $\mu$, then $\int_\mu^\infty g(y)e^{-ny}\,dy = 0$ for all $\mu$ (because $ne^{n\mu} > 0$ for all $\mu$ and does not depend on $y$). Moreover,

$$0 = \frac{d}{d\mu}\left[\int_\mu^\infty g(y)e^{-ny}\,dy\right] = -g(\mu)e^{-n\mu}$$

for all $\mu$. This implies $g(\mu) = 0$ for all $\mu$, so $X_{(1)}$ is complete.

b. Basu's Theorem says that if $X_{(1)}$ is a complete sufficient statistic for $\mu$, then $X_{(1)}$ is independent of any ancillary statistic. Therefore, we need to show only that $S^2$ has distribution independent of $\mu$; that is, $S^2$ is ancillary. Recognize that $f(x|\mu)$ is a location family. So we can write $X_i = Z_i + \mu$, where $Z_1, \ldots, Z_n$ is a random sample from $f(x|0)$. Then

$$S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2 = \frac{1}{n-1}\sum((Z_i+\mu) - (\bar{Z}+\mu))^2 = \frac{1}{n-1}\sum(Z_i - \bar{Z})^2.$$

Because $S^2$ is a function of only $Z_1, \ldots, Z_n$, the distribution of $S^2$ does not depend on $\mu$; that is, $S^2$ is ancillary. Therefore, by Basu's theorem, $S^2$ is independent of $X_{(1)}$.

7.1 For each value of $x$, the MLE $\hat{\theta}$ is the value of $\theta$ that maximizes $f(x|\theta)$. These values are in the following table.

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $\hat{\theta}$ | 1 | 1 | 2 or 3 | 3 | 3 |

At $x = 2$, $f(x|2) = f(x|3) = 1/4$ are both maxima, so both $\hat{\theta} = 2$ or $\hat{\theta} = 3$ are MLEs.

7.6 a. $f(\mathbf{x}|\theta) = \prod_i \theta x_i^{-2} I_{[\theta,\infty)}(x_i) = \left(\prod_i x_i^{-2}\right)\theta^n I_{[\theta,\infty)}(x_{(1)})$. Thus, $X_{(1)}$ is a sufficient statistic for $\theta$ by the Factorization Theorem.

b. $L(\theta|\mathbf{x}) = \theta^n\left(\prod_i x_i^{-2}\right) I_{[\theta,\infty)}(x_{(1)})$. $\theta^n$ is increasing in $\theta$. The second term does not involve $\theta$. So to maximize $L(\theta|\mathbf{x})$, we want to make $\theta$ as large as possible. But because of the indicator function, $L(\theta|\mathbf{x}) = 0$ if $\theta > x_{(1)}$. Thus, $\hat{\theta} = x_{(1)}$.

c. $\mathrm{E}X = \int_\theta^\infty \theta x^{-1}\,dx = \theta \log x|_\theta^\infty = \infty$. Thus the method of moments estimator of $\theta$ does not exist. (This is the Pareto distribution with $\alpha = \theta$, $\beta = 1$.)

7.8 a. $\mathrm{E}X^2 = \mathrm{Var}\,X + \mu^2 = \sigma^2$. Therefore $X^2$ is an unbiased estimator of $\sigma^2$.

b.

$$\begin{aligned} L(\sigma|\mathbf{x}) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)}. \quad \log L(\sigma|\mathbf{x}) = \log(2\pi)^{-1/2} - \log\sigma - x^2/(2\sigma^2). \\ \frac{\partial \log L}{\partial \sigma} &= -\frac{1}{\sigma} + \frac{x^2}{\sigma^3} \stackrel{\text{set}}{=} 0 \Rightarrow \hat{\sigma}X^2 = \hat{\sigma}^3 \;\Rightarrow \hat{\sigma} = \sqrt{X^2} = |X|. \\ \frac{\partial^2 \log L}{\partial \sigma^2} &= \frac{-3x^2\sigma^2}{\sigma^6} + \frac{1}{\sigma^2}, \text{ which is negative at } \hat{\sigma} = |x|. \end{aligned}$$

Thus, $\hat{\sigma} = |x|$ is a local maximum. Because it is the only place where the first derivative is zero, it is also a global maximum.

c. Because $\mathrm{E}X = 0$ is known, just equate $\mathrm{E}X^2 = \sigma^2 = \frac{1}{n}\sum_{i=1}^{1} X_i^2 = X^2 \Rightarrow \hat{\sigma} = |X|$.