STA5327 HW #4 Solution

6.15 a. The parameter space consists only of the points (θ, ν) 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.

6.17 The population pmf is $f(x|\theta) = \theta(1-\theta)^{x-1} = \frac{\theta}{1-\theta}e^{\log(1-\theta)x}$, an exponential family with $t(x) = x$. Thus, $\sum_i X_i$ is a complete, sufficient statistic by Theorems 6.2.10 and 6.2.25. $\sum_i X_i - n \sim$ negative binomial(n, θ).

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.

6.19 To check if the family of distributions of X is complete, we check if $\mathrm{E}_p g(X) = 0$ for all p, implies that $g(X) \equiv 0$. For Distribution 1,

$$\mathrm{E}_p g(X) = \sum_{x=0}^{2} g(x)P(X = x) = pg(0) + 3pg(1) + (1 - 4p)g(2).$$

Note that if $g(0) = -3g(1)$ and $g(2) = 0$, then the expectation is zero for all p, but $g(x)$ need not be identically zero. Hence the family is not complete. For Distribution 2 calculate

$$\mathrm{E}_p g(X) = g(0)p + g(1)p^2 + g(2)(1 - p - p^2) = [g(1) - g(2)]p^2 + [g(0) - g(2)]p + g(2).$$

This is a polynomial of degree 2 in p. To make it zero for all p each coefficient must be zero. Thus, $g(0) = g(1) = g(2) = 0$, so the family of distributions is complete.

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 θ, 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 θ. So this sufficient statistic is not complete.

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 θ, 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.

6.22 a. The sample density is $\prod_i \theta x_i^{\theta-1} = \theta^n (\prod_i x_i)^{\theta-1}$, so $\prod_i X_i$ is sufficient for θ, not $\sum_i X_i$.

b. Because $\prod_i f(x_i|\theta) = \theta^n e^{(\theta-1)\log(\prod_i x_i)}$, $\log(\prod_i X_i)$ is complete and sufficient by Theorem 6.2.25. Because $\prod_i X_i$ is a one-to-one function of $\log(\prod_i X_i)$, $\prod_i X_i$ is also a complete sufficient statistic.

6.23 Use Theorem 6.2.13. The ratio

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\theta^{-n} I_{(x_{(n)}/2,\, x_{(1)})}(\theta)}{\theta^{-n} I_{(y_{(n)}/2,\, y_{(1)})}(\theta)}$$

is constant (in fact, one) if and only if $x_{(1)} = y_{(1)}$ and $x_{(n)} = y_{(n)}$. So $(X_{(1)}, X_{(n)})$ is a minimal sufficient statistic for θ. From Exercise 6.10, we know that if a function of the sufficient statistics is ancillary, then the sufficient statistic is not complete. The uniform$(\theta, 2\theta)$ family is a scale family, with standard pdf $f(z) \sim$ uniform$(1, 2)$. So if $Z_1, \ldots, Z_n$ is a random sample from a uniform$(1,2)$ population, then $X_1 = \theta Z_1, \dots, X_n = \theta Z_n$ is a random sample from a uniform$(\theta, 2\theta)$ population, and $X_{(1)} = \theta Z_{(1)}$ and $X_{(n)} = \theta Z_{(n)}$. So $X_{(1)}/X_{(n)} = Z_{(1)}/Z_{(n)}$, a statistic whose distribution does not depend on θ. Thus, as in Exercise 6.10, $(X_{(1)}, X_{(n)})$ is not complete.

6.24 If $\lambda = 0$, $\mathrm{E}h(X) = h(0)$. If $\lambda = 1$,

$$\mathrm{E}h(X) = e^{-1}h(0) + e^{-1}\sum_{x=1}^{\infty}\frac{h(x)}{x!}.$$

Let $h(0) = 0$ and $\sum_{x=1}^{\infty}\frac{h(x)}{x!} = 0$, so $\mathrm{E}h(X) = 0$ but $h(x) \not\equiv 0$. (For example, take $h(0) = 0$, $h(1) = 1$, $h(2) = -2$, $h(x) = 0$ for $x \geq 3$.)

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 μ, then $\int_\mu^\infty g(y)e^{-ny}\,dy = 0$ for all μ (because $ne^{n\mu} > 0$ for all μ 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 μ. This implies $g(\mu) = 0$ for all μ, so $X_{(1)}$ is complete.

b. Basu's Theorem says that if $X_{(1)}$ is a complete sufficient statistic for μ, then $X_{(1)}$ is independent of any ancillary statistic. Therefore, we need to show only that S^2 has distribution independent of μ; 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 $Z_1, \dots, Z_n$ 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 $Z_1, \dots, Z_n$, the distribution of S^2 does not depend on μ; that is, S^2 is ancillary. Therefore, by Basu's theorem, S^2 is independent of $X_{(1)}$.