Miscellaneous Errors in the Chapter 6 Solutions

3.30(b) In this problem, early printings of the second edition use the beta(a,b) distribution, but later versions use the Poisson(λ) distribution. If your book uses beta(a,b), replace it by Poisson(λ). (In fact, you can**not** use Theorem 3.4.2 to calculate the mean and variance of a beta(a,b) random variable.)

6.2 There is a minor error in the solution: $e^{in\theta}$ should instead be $e^{n(n+1)\theta/2}$.

6.11(b) For dealing with the shifted exponential distribution in 6.9(b), there are two different approaches. One way to solve the problem is to show that $X_{(1)}$ is complete and sufficient, and then use Basu's lemma to prove the desired independence. Since this family of distributions is *not* an exponential family, we have to show completeness directly from the definition. We can use an argument similar to that in the text for showing $X_{(1)}$ is complete and sufficient when sampling from the Uniform(0,θ) distribution.

The solution in the manual gives a second approach. It uses the formula for the joint pdf of the order statistics given on page 234 of text. You make a multivariate transformation from $(x_{(1)},\ldots,x_{(n)})$ to $(x_{(1)},y_1,\ldots,y_{n-1})$ (this transformation has Jacobian $=1$) and then show that the resulting joint density factors, thus proving the independence.

6.17 A comment: The phrases "$T = T(X)$ is a complete statistic" and "the family of distributions of T forms a complete family" mean the same thing.

6.20(d) In spite of what the manual says, the family of distributions in part (d) **is** a one-parameter exponential family since the exponent may be factored as $-e^{-(x-\theta)} = -e^{\theta}e^{-x} = w(\theta)t(x)$. Thus, the usual argument says that a complete sufficient statistic is $T(X) = \sum_{i=1}^{n} e^{-X_i}$.

6.21(a) There is a typo in the solution. It should read: If $g(-1) = -g(1)$ and $g(0) = 0$, then $Eg(X) = 0$ for all θ, ..., There is a missing minus sign.

More Solutions

3.30(b) For the Poisson(λ) distribution, the identities become

$$\begin{aligned} E\left[\left(\frac{\partial}{\partial\lambda}\log\lambda\right)X\right] &= -\frac{\partial}{\partial\lambda}\log e^{-\lambda} \\ \operatorname{Var}\left[\left(\frac{\partial}{\partial\lambda}\log\lambda\right)X\right] &= -\frac{\partial^2}{\partial\lambda^2}\log e^{-\lambda} - E\left[\left(\frac{\partial^2}{\partial\lambda^2}\log\lambda\right)X\right] \end{aligned}$$

which simplify to

$$\begin{aligned} E\left[\frac{1}{\lambda}X\right] &= 1 \quad\Longrightarrow\quad EX = \lambda \\ \operatorname{Var}\left[\frac{1}{\lambda}X\right] &= 0 + E\left(\frac{1}{\lambda^2}X\right) \quad\Longrightarrow\quad \operatorname{Var}(X) = EX = \lambda\,. \end{aligned}$$

3.32(a)

Notation: In the special case $w_i(\theta) = \theta_i$ for all i, (3.4.1) becomes

$$f(x|\theta) = c(\theta)h(x)\exp\left(\sum_{i=1}^{k}\theta_i t_i(x)\right)$$

which is (3.4.7) with η replaced by θ. (For convenience and because I hate η, we will replace η by θ throughout this problem.)

There are two ways to prove the identities given in part (a).

Method one: Just show that the identities in Theorem 3.4.2 simplify to those in part (a) in the special case given above where $w_i(\theta) = \theta_i$ for all i. In this case $\dfrac{\partial w_i(\theta)}{\partial\theta_j} = 1$ if $i = j$, and 0 otherwise, and $\dfrac{\partial^2 w_i(\theta)}{\partial\theta_j^2} = 0$ for all i and j. Substituting these in Theorem 3.4.2 immediately leads to the formulas in part (a).

Method two: Derive the identities in (a) directly. This approach is more instructive. The proof in this special case is much less messy than the general case proved in problem 3.31.

Assume $f(x|\theta)$ is a pdf. (A similar argument applies if it is a pmf.) Differentiating both sides of

$$1 = \int_{-\infty}^{\infty} c(\theta)h(x)\exp\left(\sum_{i=1}^{k}\theta_i t_i(x)\right)\,dx$$

with respect to θ_j, taking the derivative inside the integral and using the product rule leads to

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta_j} \left\{ c(\theta) h(x) \exp\left(\sum_{i=1}^{k} \theta_i t_i(x) \right) \right\} dx \\
&= \int_{-\infty}^{\infty} \left\{ \frac{\partial c(\theta)}{\partial \theta_j} h(x) \exp\left(\sum_{i=1}^{k} \theta_i t_i(x) \right) + c(\theta) h(x) t_j(x) \exp\left(\sum_{i=1}^{k} \theta_i t_i(x) \right) \right\} dx \\
&= \int_{-\infty}^{\infty} \left\{ \frac{1}{c(\theta)} \frac{\partial c(\theta)}{\partial \theta_j} + t_j(x) \right\} f(x|\theta)\, dx = \int_{-\infty}^{\infty} \left\{ \frac{\partial}{\partial \theta_j} \log c(\theta) + t_j(x) \right\} f(x|\theta)\, dx \qquad (\dagger) \\
&= E\left(\frac{\partial}{\partial \theta_j} \log c(\theta) + t_j(X) \right)
\end{aligned}$$

Thus $Et_j(X) = -\frac{\partial}{\partial \theta_j} \log c(\theta)$. Now do the same thing over again: Differentiate the integral ($\dagger$) with respect to θ_j, take the derivative inside the integral, and use the product rule and the fact (essentially demonstrated above) that $\frac{\partial}{\partial \theta_j} f(x|\theta) = \left\{ \frac{\partial}{\partial \theta_j} \log c(\theta) + t_j(x) \right\} f(x|\theta)$ to obtain the following:

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty} \left\{ \frac{\partial^2}{\partial \theta_j^2} \log c(\theta) + \left(\frac{\partial}{\partial \theta_j} \log c(\theta) + t_j(x) \right)^2 \right\} f(x|\theta)\, dx \\
&= \frac{\partial^2}{\partial \theta_j^2} \log c(\theta) + E\left(\frac{\partial}{\partial \theta_j} \log c(\theta) + t_j(X) \right)^2 \\
&= \frac{\partial^2}{\partial \theta_j^2} \log c(\theta) + \mathrm{Var}\left(t_j(X) \right)
\end{aligned}$$

since $Et_j(X) = -\frac{\partial}{\partial \theta_j} \log c(\theta)$. Thus $\mathrm{Var}(t_j(X)) = -\frac{\partial^2}{\partial \theta_j^2} \log c(\theta)$.

3.32(b) The Gamma(α, β) pdf is

$$\begin{aligned}
f(x|\alpha, \beta) &= \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} \quad \text{for } x > 0 \\
&= x^{-1} I(x > 0) \frac{1}{\beta^\alpha \Gamma(\alpha)} \exp\left\{ \alpha \log x + \frac{-1}{\beta} x \right\} \\
&= h(x) \frac{(-\eta_2)^{\eta_1}}{\Gamma(\eta_1)} \exp\left\{ \eta_1 \log x + \eta_2 x \right\} .
\end{aligned}$$

Using the identities from part (a) with $j = 2$ we then obtain

$$\begin{aligned}
EX &= -\frac{\partial}{\partial \eta_2} \log\left(\frac{(-\eta_2)^{\eta_1}}{\Gamma(\eta_1)}\right) \\
&= -\frac{\partial}{\partial \eta_2} \left(\eta_1 \log(-\eta_2) - \log\Gamma(\eta_1)\right) \\
&= \frac{\eta_1}{(-\eta_2)} = \alpha\beta\,, \\
\mathrm{Var}(X) &= -\frac{\partial^2}{\partial \eta_2^2} \log\left(\frac{(-\eta_2)^{\eta_1}}{\Gamma(\eta_1)}\right) \\
&= \frac{\partial}{\partial \eta_2} \frac{\eta_1}{(-\eta_2)} \\
&= \frac{\eta_1}{\eta_2^2} = \alpha\beta^2\,.
\end{aligned}$$

6.16

See the discussion of the multinomial distribution in Section 4.6. Right now we need only the formula for the joint pmf given on page 180.

See also the definition of a curved exponential family in Section 3.4 on page 115.

(a) Let the cell probabilities be denoted p_1, p_2, p_3, p_4. The vector (x_1, x_2, x_3, x_4) has a multinomial distribution with m trials and cell probabilities p_1, p_2, p_3, p_4. Since $p_2 = p_3$ and $x_4 = n - x_1 - x_2 - x_3$, we may write the multinomial joint pmf as

$$\begin{aligned}
f(x_1, x_2, x_3, x_4 \mid \theta) &= \frac{m!}{x_1!x_2!x_3!x_4!}\, p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4} \\
&= \frac{m!}{x_1!x_2!x_3!x_4!}\, p_1^{x_1} p_2^{x_2+x_3} p_4^{m-x_1-(x_2+x_3)} \\
&= \frac{m!}{x_1!x_2!x_3!x_4!}\, p_4^m \left(\frac{p_1}{p_4}\right)^{x_1} \left(\frac{p_2}{p_4}\right)^{x_2+x_3} \\
&= h(x)c(\theta)\exp\left(w_1(\theta)t_1(x) + w_2(\theta)t_2(x)\right)
\end{aligned}$$

where

$$h(x) = \frac{m!}{x_1!x_2!x_3!x_4!}\,, \quad c(\theta) = p_4^m = (\theta/4)^m\,, \quad t(x) = (x_1, x_2 + x_3)\,,$$

$$w(\theta) = (\log(p_1/p_4), \log(p_2/p_4)) = \left(\log\left(\frac{2+\theta}{\theta}\right), \log\left(\frac{1-\theta}{\theta}\right)\right)$$

so that the joint pmf is a curved exponential family where the dimension of the parameter θ is $d = 1$ and the number of terms in the exponent is $k = 2$. The family is "curved" since $d < k$.

Comment: The terminology "curved exponential family" arises because when $d < k$ the set of points $\{w(\theta) : \theta \in \Theta\}$ is a curve in the natural parameter space. In this case $\{w(\theta) : 0 \le \theta \le 1\}$ is a 1-dimensional curve in R^2 which is the natural parameter space. Note that a curved exponential family never satisfies the OSC (open set condition) so that we cannot use the usual theorem to prove completeness of the natural sufficient statistic. In fact, in this case the natural sufficient statistic is not complete since $E\left(2x_1 + (x_2 + x_3) - \frac{3}{2}m\right) = 0$ for all θ, but $P\left(2x_1 + (x_2 + x_3) - \frac{3}{2}m\right) = 0) \neq 1$.

(b) and **(c)** The natural sufficient statistic $t(x) = (x_1, x_2 + x_3)$ is (of course) sufficient. To show it is minimal sufficient we apply the Lehmann-Scheffe Theorem:

$$\frac{f(x \mid \theta)}{f(y \mid \theta)} = \frac{h(x)}{h(y)} \exp\left(a_1 w_1(\theta) + a_2 w_2(\theta)\right)$$

where $a_1 = x_1 - y_1$ and $a_2 = (x_2 + x_3) - (y_2 + y_3)$. This is constant in θ if and only if

$$a_1 w_1(\theta) + a_2 w_2(\theta) \quad \text{is constant in } \theta.$$

With the expressions for $w_1(\theta)$ and $w_2(\theta)$ given above, this is "clearly" true if and only if $a_1 = a_2 = 0$ which means $(x_1, x_2 + x_3) = (y_1, y_2 + y_3)$. One direction is immediate: if $a_1 = a_2 = 0$, then obviously $a_1 w_1(\theta) + a_2 w_2(\theta) = 0$ for all θ so that it is constant in θ. Now for the other direction. Assume there exists some finite value c such that $a_1 w_1(\theta) + a_2 w_2(\theta) = c$ for all θ. Then

$$\lim_{\theta \to 1} \left(a_1 w_1(\theta) + a_2 w_2(\theta)\right) = c\,. \tag{1}$$

But $\lim_{\theta \to 1} w_1(\theta) = \log 3$ and $\lim_{\theta \to 1} w_2(\theta) = -\infty$ so that (1) implies $a_2 = 0$ (otherwise the limit would be either $+\infty$ or $-\infty$ depending on the sign of a_2). But now, since $w_1(\theta)$ is **not** constant in θ, the only way that $a_1 w_1(\theta)$ can be constant in θ is for $a_1 = 0$. Thus $a_1 = a_2 = 0$ as desired.

6.22(a)

The solution in the manual is not detailed enough.

Showing that a particular statistic $T(X)$ is **not** sufficient can be done in (at least) three different ways. The first two ways described below are relatively easy to carry out. The third approach can be difficult. If you know a minimal sufficient statistic, the second approach is the easiest.

(1): The first approach uses the following result which can be proved using the factorization criterion. (This result is one part of the Lehmann-Scheffe Theorem.)

If $T(X)$ is a sufficient statistic for θ, then for any two samples x and y:

$$\text{If } T(x) = T(y), \text{ then } \frac{f(x \mid \theta)}{f(y \mid \theta)} \text{ is constant in } \theta.$$

Thus to show that a statistic $T(X)$ is **not** sufficient, it suffices to find two samples x and y with $T(x) = T(y)$ for which the ratio above is **not** constant.

(2): The second approach uses the fact that a minimal sufficient statistic must be a function of any other sufficient statistic. So, if $S(X)$ is minimal sufficient and $S(X)$ can**not** be expressed as a function of $T(X)$, then $T(X)$ can**not** be sufficient. It is usually pretty obvious whether or not one statistic can be expressed as a function of another, but a formal proof can be obtained as follows. If you can find two samples x and y for which $T(x) = T(y)$ but $S(x) \neq S(y)$, then S can**not** be expressed as a function of T.

(3): The third approach is to use the definition of a sufficient statistic. Let $T = T(X)$. A statistic is sufficient (for θ) if $\mathcal{L}(X \mid T)$ does not depend on θ. If you can show that the conditional distribution $\mathcal{L}(X \mid T)$ **does** depend on θ, then T can**not** be sufficient.

6.22(a) by Method #2: In this problem, a minimal sufficient statistic is $\prod X_i$ (or equivalently $\sum \log X_i$). To show that $\sum X_i$ is not sufficient, it suffices to find two samples x and y for which $\sum x_i = \sum y_i$ but $\prod x_i \neq \prod y_i$. This is easy to do.