Solution 1

The Poisson model is a $(1, 1)$ exponential family, so (by Example 36 and Theorem 38) S is a complete minimal sufficient statistic. It is a sum of independent Poisson variables, so $S \sim \text{Poiss}(n\lambda)$.

(a) The required conditional distribution is obtained by setting

$$f(y_1, \ldots, y_n \mid s) = \frac{f(y_1, \ldots, y_n) f(s \mid y_1, \ldots, y_n)}{f(s)} = \frac{\prod_{j=1}^n \lambda^{y_j} e^{-\lambda}/y_j! \times I(s = y_1 + \cdots + y_n)}{(n\lambda)^s e^{-n\lambda}/s!}$$

and this can be written as

$$f(y_1, \ldots, y_n \mid s) = \frac{s!}{\prod_{j=1}^n y_j!} \prod_{j=1}^n n^{-y_j} \times I(s = y_1 + \cdots + y_n), \quad y_1, \ldots, y_n \in \{0, 1, \ldots\},$$

which is multinomial with denominator s and probability vector $(1/n, \ldots, 1/n)$.

(b) Yes, because any function $h(S)$ that is unbiased for $\psi(\lambda)$ must be unique, by completeness of S.

(c) (i) Either strategy could be used. For the first, note that as $T = I(Y_1 = 0)$ is unbiased for $e^{-\lambda}$, we need to compute $\mathrm{E}(T \mid S)$. The multinomial distribution in (a) implies that $Y_1 \mid S = s \sim B(s, 1/n)$, so

$$\mathrm{E}(T \mid S = s) = \mathrm{E}\{I(Y_1 = 0) \mid S = s\} = \mathrm{P}(Y_1 = 0 \mid S = s) = (1 - 1/n)^s,$$

i.e., $(1 - 1/n)^S$ must be the optimal unbiased estimator. To check this, note that

$$\mathrm{E}\{(1 - 1/n)^S\} = \sum_{s=0}^{\infty} (1 - 1/n)^s (n\lambda)^s e^{-n\lambda}/s! = e^{-n\lambda} \sum_{s=0}^{\infty} (n-1)^s \lambda^s/s! = e^{-n\lambda} e^{(n-1)\lambda} = e^{-\lambda},$$

as required.

For the second strategy, we seek to solve $\mathrm{E}\{h(S)\} = e^{-\lambda}$, i.e.,

$$\sum_{s=0}^{\infty} h(s)(n\lambda)^s e^{-n\lambda}/s! = e^{-\lambda} \quad \Longrightarrow \quad \sum_{s=0}^{\infty} h(s)(n\lambda)^s/s! = e^{(n-1)\lambda},$$

and taking $h(s) = (n-1)^s/n^s$ achieves this, and gives the same estimator as the first strategy.

(ii) Here it's not clear what unbiased estimator to start from, so we use the second strategy. We need to solve $\mathrm{E}\{h(S)\} = e^{-2n\lambda}$, giving

$$\sum_{s=0}^{\infty} h(s)(n\lambda)^s e^{-n\lambda}/s! = e^{-2n\lambda} \quad \Longrightarrow \quad \sum_{s=0}^{\infty} h(s)(n\lambda)^s/s! = e^{-n\lambda},$$

so we must choose

$$h(s) = \begin{cases} -1, & s \text{ odd}, \\ 1, & s \text{ even}. \end{cases}$$

The estimator in (i) seems reasonable, but that in (ii) is not, as it can be negative when estimating a positive quantity, and ± 1 is unlikely to be close to the target (a small positive number, most likely). An obvious alternative would be $\exp(-2S)$, as $\widehat{\lambda} = S/n$ is the maximum likelihood estimator; this estimator is biased but asymptotically has the smallest variance (by the Cramèr–Rao lower bound).

Solution 2

(a) We have $\mathrm{E}(Y_j) = \theta/2$ and $\mathrm{var}(Y_j) = \theta^2/12$ so the unbiasedness of U is immediate and its variance is $\theta^2/(3n) = O(n^{-1})$.

(b) The Rao–Blackwell theorem tells us that $\mathrm{E}(U \mid M)$ will be unbiased with a smaller variance, and we can write

$$U = \frac{2}{n}(M + Y_1' + \cdots Y_{n-1}'),$$

where $Y_1', \ldots, Y_n'$ are the values of the original sample that are not the maximum. Each of these must lie in the interval $(0, M)$ and we saw in the notes that conditional on $M = m$ they are independent with $U(0, m)$ distributions, so $\mathrm{E}(Y_j' \mid M = m) = m/2$. Hence

$$\mathrm{E}(U \mid M = m) = \frac{2}{n}\{m + (n-1)m/2\} = \frac{m}{n}(2 + n - 1) = (n+1)m/n,$$

so the unbiased estimator based on M is $\tilde{\theta} = (n+1)M/n$. As $\mathrm{E}(M^r) = n\theta^r/(n+r)$, we have $\mathrm{var}(M) = n\theta^2/\{(n+2)(n+1)^2\}$, so

$$\mathrm{var}(\tilde{\theta}) = \frac{(n+1)^2}{n^2}\mathrm{var}(M) = \frac{\theta^2}{n(n+2)} = O(n^{-2}),$$

i.e., Rao–Blackwellisation has given an unbiased estimator with variance of lower order in n, which is therefore much better than U.

For an alternative argument not involving order statistics, we might write

$$\mathrm{E}(U \mid M) = \frac{2}{n}\mathrm{E}\left\{\sum_{j=1}^{n} Y_j \mid M = m\right\} = \frac{2}{n}\sum_{j=1}^{n}\mathrm{E}\left\{Y_j I(Y_j = M) + Y_j I(Y_j < M) \mid M = m\right\}$$

and then note that

$$E\left\{Y_j I(Y_j = M) \mid M = m\right\} + \mathrm{E}\left\{Y_j I(Y_j < M) \mid M = m\right\} = m\frac{1}{n} + \frac{m}{2}\frac{n-1}{n}$$

giving

$$\mathrm{E}(U \mid M) = \frac{2}{n}\sum_{j=1}^{n}\left(m\frac{1}{n} + \frac{m}{2}\frac{n-1}{n}\right) = (n+1)m/n.$$

Solution 3

(a) Replacing n in the first expression by $n-1$ allows the term $f_{Y_1,\ldots,Y_{n-1}}(y_1, \ldots, y_{n-1})$ to be written in terms of the conditional density of Y_{n-1} given $Y_1, \ldots, Y_{n-2}$, and then we simply iterate this, finishing by writing $f_{Y_1,Y_2}(y_1, y_2) = f_{Y_2|Y_1}(y_2 \mid y_1)f_{Y_1}(y_1)$.

(b) This is a Markov process because the density of Y_j given the past depends only on Y_{j-1}, so we obtain

$$\begin{aligned} f_{Y_0,\ldots,Y_n}(y_0, \ldots, y_n) &= f_{Y_0}(y_0)\prod_{j=1}^{n} f_{Y_j|Y_{j-1}}(y_j \mid y_{j-1}) \\ &= f(y_0)\prod_{j=1}^{n}\frac{\exp\{-\{y_j - \mu - \alpha(y_{j-1} - \mu)\}^2/(2\sigma^2)\}}{(2\pi\sigma^2)^{1/2}}, \end{aligned}$$

and the log likelihood is

$$-2\ell(\mu, \alpha, \sigma^2) \equiv -2\log f(y_0) + n\log\sigma^2 + \frac{1}{\sigma^2}\sum_{j=1}^{n}\{y_j - \mu - \alpha(y_{j-1} - \mu)\}^2,$$

with (i) $f(y_0) = 1$ and (ii) $-2\log f(y_0) \equiv \log\sigma^2 - \log(1-\alpha^2) + (y_0 - \mu)^2(1-\alpha^2)/\sigma^2$.

(c) The quadratic form in (b) can be expanded as

$$\frac{1}{\sigma^2}\left\{\sum_{j=1}^n y_j^2 - 2\alpha\sum_{j=1}^n y_jy_{j-1} + \alpha^2\sum_{j=1}^n y_{j-1}^2 + 2(\alpha-1)\mu\sum_{j=1}^n y_j - 2\alpha(\alpha-1)\mu\sum_{j=0}^{n-1} y_j + n(\alpha-1)^2\mu^2\right\},$$

so if (i) y_0 is constant this is a $(5,3)$ exponential family with $\theta = (\mu,\sigma^2,\alpha)$, minimal sufficient statistic

$$s(y)^{\mathrm{T}} = \left(\sum_{j=1}^n y_j^2, \sum_{j=1}^n y_{j-1}^2, \sum_{j=1}^n y_jy_{j-1}, \sum_{j=1}^n y_j, \sum_{j=0}^{n-1} y_j\right),$$

and

$$\varphi(\theta)^{\mathrm{T}} = -2(1/\theta_2, \theta_3^2/\theta_2, -2\theta_3/\theta_2, 2(\theta_3-1)\theta_1/\theta_2, 2\theta_3(\theta_3-1)\theta_1/\theta_2),\quad k(\varphi) = -\frac{n(\theta_3-1)^2\theta_1^2}{2\theta_2} - \frac{n}{2}\log\theta_2,$$

and (ii) y_0 must be added to the minimal sufficient statistic, making the model a $(6,3)$ exponential family.

Solution 4

(a) Obviously (i) $y \sim y$ (reflexivity) and (ii) $y' \sim y$ is equivalent to $y' \sim y$ (symmetry). Equally obvious is that (iii) $y' \sim y$ and $y'' \sim y'$ implies that $y'' \sim y$ (transitivity). Hence the relation $\sim$ is an equivalence relation, which implies that the equivalence classes $\mathcal{C}_s$ for $s \in \mathcal{T}$ form a partition of $\mathcal{Y}$. If $g(s)$ is a bijective function of s, then an inverse function g^{-1} exists and so $\mathcal{C}_h = \{y : g\{t(y)\} = h\} = \{y : g^{-1}[g\{t(y)\}] = g^{-1}(h)\} = \{y : t(y) = g^{-1}(h)\} = \mathcal{C}_{g^{-1}(h)}$, so the equivalence classes defined using g are the same as those defined using t.

(b) We have

$$\mathrm{P}(Y = y \mid Y \in \mathcal{C}_s;\theta) = \frac{\mathrm{P}(Y=y;\theta)\mathrm{P}(Y\in\mathcal{C}_s \mid Y = y;\theta)}{\mathrm{P}(Y\in\mathcal{C}_s;\theta)} = \frac{\mathrm{P}(Y=y;\theta)}{\mathrm{P}(Y\in\mathcal{C}_s;\theta)}I\{t(y)=s\},$$

and the factorisation theorem applies and the last expression here is, as required,

$$\frac{g\{t(y);\theta\}h(y)}{\sum_{y'\in\mathcal{C}_s} g\{t(y');\theta\}h(y')}I\{t(y)=s\} = \frac{g(s;\theta)h(y)}{\sum_{y'\in\mathcal{C}_s} g(s;\theta)h(y')} = \frac{h(y)}{\sum_{y'\in\mathcal{C}_s} h(y')}.$$

(c) The joint density is

$$\prod_{j=1}^n \frac{\theta^{y_j}}{y_j!}e^{-\theta} = \theta^{y_1+(y_2+\cdots+y_n)}\exp(-n\theta)\times h(y),$$

where $h(y) = 1/\prod y_j!$, so the factorisation theorem implies that $(T_1,T_2) = (Y_1, Y_2+\cdots+Y_n)$ is sufficient, with $\mathcal{Y} = \{(y,\ldots,y_n) : y_1,\ldots,y_n \in \{0,1,\ldots\}\}$ and

$$\mathcal{T} = \{(t_1,t_2) : t_1,t_2\in\{0,1,\ldots\}\},\quad \mathcal{C}_s = \{y\in\mathcal{Y} : y_1 = s_1, \sum_{j=2}^n y_j = s_2\},\quad s\in\mathcal{T}.$$

Hence the elements of this partition are determined by points (s_1,s_2) in the positive quadrant with values in $\{0,1,2,\ldots\}$. It is clear from the form of the joint density that $T_1+T_2 = Y_1+\cdots+Y_n$ is also sufficient, and the corresponding sufficient partition consists of the sets $\mathcal{C}'_t = \{(s_1,s_2)\in\mathcal{T} : s_1+s_2 = t\}$ for $t\in\{0,1,\ldots\}$, which lie on diagonals in the positive quadrant. This partition is minimal because if $y_1,\ldots,y_n$ and $z_1,\ldots,z_m$ are both Poisson samples, their likelihood ratio is

$$\frac{\theta^{s_z}\exp(-m\theta)/\prod z_j!}{\theta^{s_y}\exp(-n\theta)/\prod y_j!} \propto \theta^{s_z-s_y}\exp\{\theta(n-m)\},$$

and this is independent of θ iff $n = m$ and $s_z = \sum_{j=1}^m z_j = s_y = \sum_{j=1}^n y_j$. If the sample size n is fixed, the minimal sufficient statistic is $S = Y_1+\cdots+Y_n$.