MATH 557 - EXERCISES 1 : SOLUTIONS

1. (a) We have

$$f_{\mathbf{X}}(\mathbf{x};\alpha,\beta) = \left\{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right\}^n \left\{\prod_{i=1}^n x_i\right\}^{\alpha-1} \left\{\prod_{i=1}^n (1-x_i)\right\}^{\beta-1}$$

suggesting the sufficient statistic $\mathbf{T}(\mathbf{X}) = \left(\prod_{i=1}^n x_i, \prod_{i=1}^n (1-x_i)\right)^\top$ and the result follows using the Fisher-Neyman Factorization Theorem.

(b) Writing $\lambda = \log\theta$, we realize that this is the $Poisson(\log\theta)$ model. Hence by properties of the Exponential Family $T(\mathbf{X}) = \sum_{i=1}^n X_i$ is a sufficient statistic for $\log\theta$.

(c) We have

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \frac{\prod_{i=1}^n \mathbb{1}_{(\theta,2\theta)}(x_i)}{\theta^n} = \frac{\mathbb{1}_{(x_{(n)}/2,x_{(1)})}(\theta)}{\theta^n}$$

yielding $\mathbf{T}(\mathbf{X}) = \left(X_{(1)}, X_{(n)}\right)^\top$ and the result follows using the factorization theorem.

2. (a) We have

$$f_{\mathbf{X}}(\mathbf{x};\lambda) = \lambda^n \exp\left\{-\lambda\sum_{i=1}^n x_i\right\} = \lambda^n \exp\left\{-\lambda T(\mathbf{x})\right\}$$

say, so that for two points $\mathbf{x}$ and $\mathbf{y}$ the ratio

$$\frac{f_{\mathbf{X}}(\mathbf{x};\lambda)}{f_{\mathbf{X}}(\mathbf{y};\lambda)} = \exp\left\{-\lambda\left(T(\mathbf{x}) - T(\mathbf{y})\right)\right\}$$

which is a constant if and only if $T(\mathbf{x}) = T(\mathbf{y})$. Therefore $T(\mathbf{x})$ is minimal sufficient.

(b) Given $x_{(1)},\ldots,x_{(m)}$, we can construct the joint pdf for the order statistic data by noting that if $X_{(m)} = x_{(m)}$, then we have $X_{(r)} > x_{(m)}$ for the $n-m$ order statistics $X_{(r)},\ r = m+1,\ldots,n$. Thus, as the "survivor" function takes the form $1 - F_X(x;\lambda) = e^{-\lambda x}$, we have

$$f_{\mathbf{X}}(\mathbf{x};\lambda) = m!\binom{n}{m} \times \lambda^m \exp\left\{-\lambda\sum_{i=1}^m x_{(i)}\right\} \times \exp\left\{-(n-m)\lambda x_{(m)}\right\}$$

where the combinatorial term counts the number of possible arrangements of the random sample points. Thus a sufficient statistic is

$$T(\mathbf{X}) = \sum_{i=1}^m X_{(i)} + (n-m)X_{(m)}$$

by the factorization theorem.

3. We have for $t = 0,1,\ldots,$

$$f_{\mathbf{X}}(\mathbf{x};\theta) = \frac{\theta^{\sum_{i=1}^n x_i} e^{-n\theta}}{\prod_{i=1}^n x_i!}$$

and $T(\mathbf{X}) \sim Poisson(n\theta)$ from distributional results, so that

$$f_{\mathbf{X}|T}(\mathbf{x}|t) = \frac{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta} / \prod_{i=1}^{n} x_i!}{(n\theta)^t e^{-n\theta}/t!} = \frac{t!}{x_1! \ldots, x_n!} \left(\frac{1}{n}\right)^t \qquad \mathbf{x} \in A_t$$

and zero otherwise, where $A_t \equiv \{\mathbf{x} : x_1 + \cdots + x_n = t\}$.

4. (a) Suppose, for two points $\mathbf{x}$ and $\mathbf{y}$ in the parameter space, $c(\mathbf{x}, \mathbf{y})$ is a function of these two arguments. We have that

$$\begin{aligned} \frac{f_i(\mathbf{x})}{f_i(\mathbf{y})} = c(\mathbf{x}, \mathbf{y}),\ i = 0, \ldots, k \quad &\iff \quad \frac{f_i(\mathbf{x})}{f_i(\mathbf{y})} = \frac{f_0(\mathbf{x})}{f_0(\mathbf{y})},\ i = 1, \ldots, k \\ &\iff \quad \frac{f_i(\mathbf{x})}{f_0(\mathbf{x})} = \frac{f_i(\mathbf{y})}{f_0(\mathbf{y})},\ i = 1, \ldots, k \\ &\iff \quad T_i(\mathbf{x}) = T_i(\mathbf{y}) \end{aligned}$$

for the given statistic $\mathbf{T}(\mathbf{X}) = (T_1(\mathbf{X}), \ldots, T_k(\mathbf{X}))^\top = (T_i(\mathbf{X}))_{i=1,\ldots,k}^\top$. Therefore, by the theorem from 2 (a), $\mathbf{T}(\mathbf{X})$ is minimal sufficient. This result implies that for selecting a *model* from a group of models, likelihood ratios provide sufficient statistics.

(b) If $\mathbf{T}^*(\mathbf{X})$ is sufficient for θ, then by the factorization theorem there exist g and h such that

$$f_\theta(\mathbf{x}) = g\left(\mathbf{T}^*(\mathbf{x}); \theta\right) h(\mathbf{x}).$$

Thus, choosing θ_i, $i = 1, \ldots, k$, such that $f_i = f_{\theta_i}$,

$$\mathbf{T}(\mathbf{X}) = \left(\frac{f_i(\mathbf{X})}{f_0(\mathbf{X})}\right)_{i=1,\ldots,k}^\top = \left(\frac{g_i(\mathbf{T}^*(\mathbf{X}); \theta_i) h(\mathbf{X})}{g_0(\mathbf{T}^*(\mathbf{X}); \theta_0) h(\mathbf{X})}\right)_{i=1,\ldots,k}^\top = \left(\frac{g_i(\mathbf{T}^*(\mathbf{X}); \theta_i)}{g_0(\mathbf{T}^*(\mathbf{X}); \theta_0)}\right)_{i=1,\ldots,k}^\top$$

Hence for any sufficient $\mathbf{T}^*$, $\mathbf{T}$ is a function of $\mathbf{T}^*$. Therefore $\mathbf{T}$ is minimal sufficient.

(c) Let T be minimal sufficient for θ, so that if W is also sufficient for θ, there exists h such that $h(W) = T$. Let $T^* = r(T)$ where r is a 1-1 mapping. Then $T^* = r(h(W))$, so T^* is a function of every sufficient statistic.

(d) In the notation of the earlier parts, let $\mathcal{F}_1 = \{f_\beta : \beta = 1, 2\}$ and $\mathcal{F} = \{f_\beta : \beta > 0\}$, where f_β is the Exponential density with expectation β, so that

$$f_\beta(x) = \frac{1}{\beta} e^{-x/\beta} \mathbb{1}_{(0,\infty)}(x).$$

If $X_{(1)} = \min\{X_1, \ldots, X_n\}$, define

$$T(\mathbf{X}) = \frac{f_2(\mathbf{X})}{f_1(\mathbf{X})} = \frac{\frac{1}{2^n} e^{-\frac{1}{2}\sum_{i=1}^{n} X_i} \mathbb{1}_{(0,\infty)}(X_{(1)})}{e^{-\sum_{i=1}^{n} X_i} \mathbb{1}_{(0,\infty)}(X_{(1)})} = 2^{-n} \exp\left\{\frac{n\overline{X}}{2}\right\}.$$

However, since

$$f_\beta(\mathbf{x}) = \frac{1}{\beta^n} e^{-n\overline{x}/\beta} \, \mathbb{1}_{(0,\infty)}(x_{(1)}) = \frac{1}{(2\beta)^n} \frac{1}{T(\mathbf{x})^{2/\beta}} \, \mathbb{1}_{(0,\infty)}(x_{(1)})$$

$T(\mathbf{X})$ is sufficient for β by the factorization theorem. Therefore $T(\mathbf{X})$ is minimal sufficient for β by part (b). Since $T(\cdot)$ is a 1-1 transformation, $\overline{X}$ is minimal sufficient for β by part (c).

5. (a) The joint pmf takes the form of an n-sample multinomial without the combinatorial term (here we observe the X_is individually, not merely the totals in each of the categories):

$$f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\theta}) = \prod_{j=1}^{3} \theta_j^{\sum_{i=1}^{n} \mathbb{1}_{\{j\}}(x_i)} = \prod_{j=1}^{3} \theta_j^{n_j}$$

where

$$n_j = \sum_{i=1}^{n} \mathbb{1}_{\{j\}}(x_i)$$

counts the number of times X_i takes the value j, for $j = 1, 2, 3$. But $n_3 = n - n_1 - n_2$, hence

$$f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\theta}) = \theta_1^{n_1}\theta_2^{n_2}(1-\theta_1-\theta_2)^{n-n_1-n_2}$$

Thus, by the factorization Theorem, $\mathbf{T} = (N_1, N_2)^\top$ is sufficient for $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top$.

(b) When looking at N_1 and N_2, we have the traditional multinomial pmf, so

$$f_{\mathbf{T}}(t_1, t_2;\boldsymbol{\theta}) = \frac{n!}{n_1! n_2! (n-n_1-n_2)!}\theta_1^{n_1}\theta_2^{n_2}(1-\theta_1-\theta_2)_3^n$$

where $\mathbf{T}(\mathbf{x}) = (n_1, n_2)^\top$ and $n_3 = n - n_1 - n_2$.

(c) Using the using calculus approach, we form the log likelihood, partially differentiate in turn with respect to θ_1 and θ_2, equate to zero, and then solve the resulting two equations simultaneously. We have

$$l(\theta_1, \theta_2; \mathbf{x}) = c(\mathbf{n}) + n_1 \log\theta_1 + n_2 \log\theta_2 + n_3 \log(1-\theta_1-\theta_2)$$

so that

$$\begin{aligned} \frac{\partial l}{\partial \theta_1} &= \frac{n_1}{\theta_1} - \frac{n_3}{1-\theta_1-\theta_2} \\ \frac{\partial l}{\partial \theta_2} &= \frac{n_2}{\theta_2} - \frac{n_3}{1-\theta_1-\theta_2} \end{aligned}$$

Equating to zero and subtracting the second from the first equation, we obtain that

$$\frac{n_1}{\theta_1} = \frac{n_2}{\theta_2} \qquad \therefore \qquad \frac{\theta_1}{\theta_2} = \frac{n_1}{n_2}.$$

Substituting back into the first equation, we have

$$\frac{n_1}{n_3} = \frac{\theta_1}{1-\theta_1-\theta_2} = \frac{\theta_1}{1-\theta_1-n_2\theta_1/n_1}$$

Cross multiplying, we get

$$n_1 - n_1\theta_1 - n_2\theta_1 = n_3\theta_1$$

and hence

$$n_1 = (n_1+n_2+n_3)\theta_1 \qquad \therefore \qquad \theta_1 = \frac{n_1}{n_1+n_2+n_3} = \frac{n_1}{n}$$

Thus, by a similar argument for θ_2, we deduce

$$\widehat{\boldsymbol{\theta}}(\mathbf{x}) = (n_1/n, n_2/n)^\top.$$

and hence the estimator is

$$\widehat{\boldsymbol{\theta}}(\mathbf{X}) = (N_1/n, N_2/n)^\top.$$

Note that the case covered here presumes $n_1, n_2, n_3 > 0$. If this does not hold, the maximum likelihood estimator does not lie in the parameter space and is not defined.