Stat 821 Homework 3 Solution

Question 1

(a)

$$\begin{aligned} f(x) &= \prod_{i=1}^{n}[\theta x_i^{\theta-1} I_{(0,1)}(x_i)] \\ &= \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1} \prod_{i=1}^{n} I_{(0,1)}(x_i) \\ &= \left[I_{(0,1)}(\min\{x_1,\ldots,x_n\}) I_{(0,1)}(\max\{x_1,\ldots,x_n\})\right] \left[\left(\prod_{i=1}^{n} x_i\right)^{\theta-1} \theta^n\right] \end{aligned}$$

By Neyman's factorization theorem, $T = \prod_{i=1}^{n} x_i$ is the sufficient statistics for θ.

(b)

$$\begin{aligned} f(x) &= \prod_{i=1}^{n} \theta a x_i^{a-1} \exp\{-\theta x_i^a\} I_{(0,1)}(x_i) \\ &= (\theta a)^n \left(\prod_{i=1}^{n} x_i\right)^{a-1} \exp\{-\theta \sum_{i=1}^{n} x_i^a\} I_{(0,1)}(\min\{x_1,\ldots,x_n\}) \\ &= \left[a^n \left(\prod_{i=1}^{n} x_i\right)^{a-1} I_{(0,1)}(\min\{x_1,\ldots,x_n\})\right] \left[\theta^n \exp\{-\theta \sum_{i=1}^{n} x_i^a\}\right] \end{aligned}$$

By Neyman's factorization theorem, $T = \sum_{i=1}^{n} x_i^a$ is the sufficient statistics for θ.

(c)

$$\begin{aligned} f(x) &= \prod_{i=1}^{n} \frac{\theta a^\theta}{x_i^{\theta+1}} I_{(a,\infty)}(x_i) \\ &= \theta^n a^{n\theta} \frac{1}{(\prod_{i=1}^{n} x_i)^{\theta+1}} I_{(a,\infty)}(\min\{x_1,\ldots,x_n\}) \\ &= \left[I_{(a,\infty)}(\min\{x_1,\ldots,x_n\})\right] \left[\frac{\theta^n a^{n\theta}}{(\prod_{i=1}^{n} x_i)^{\theta+1}}\right] \end{aligned}$$

By Neyman's factorization theorem, $T = \prod_{i=1}^{n} x_i$ is the sufficient statistics for θ.

Question 2

(a) For any $\theta \in \Theta$, the posterior distribution of $\theta|x$ is:

$$\frac{f(x|\theta)\pi(\theta)}{\sum_{\theta_i \in \Theta} f(x|\theta_i)\pi(\theta_i)} \qquad \text{where } \pi(\theta) \text{ is any prior on } \theta$$

By the condition, we know it equals to some function $g(\theta, T(x))$, i.e.

$$\frac{f(x|\theta)\pi(\theta)}{\sum_{\theta_i \in \Theta} f(x|\theta_i)\pi(\theta_i)} = g(\theta, T(x))$$

where $g(x, T(x))$ is a function of θ and $T(x)$ only. Thus

$$f(x|\theta) = \frac{g(x, T(x))}{\pi(\theta)} \sum_{\theta_i \in \Theta} f(x|\theta_i)\pi(\theta_i)$$

By factorization theorem, T(x) is sufficient for θ.

(b) If $T(x)$ is sufficient, then $f(x|\theta)$ can be written as

$$f(x|\theta) = g(\theta, T(x))h(x)$$

Let $\pi(\theta)$ be an arbitrary prior distribution, then the posterior of θ is

$$\frac{f(x|\theta)\pi(\theta)}{\sum_{\theta_i \in \Theta} f(x|\theta_i)\pi(\theta_i)} = \frac{g(\theta, T(x))}{\sum_{\theta_i \in \Theta} g(\theta_i, T(x))\pi(\theta_i)} \pi(\theta)$$

The posterior depends on x only through $T(x)$. By factorization theorem, T(x) is sufficient for θ.

Question 3 (Problem 6.7)

(a) $\theta = \{\xi, \eta, \sigma, \tau : -\infty < \xi, \eta < \infty, 0 < \sigma, \tau\}$ The parameter space contains 4-dimensional rectangle space.

$$\begin{aligned} f(x) &= \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\xi)^2} \\ &= \frac{1}{(2\pi)^{m/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum x_i^2 + \frac{\xi}{\sigma^2}\sum x_i - \frac{m\xi^2}{2\sigma^2} - m\log\sigma\right\} \\ f(y) &= \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2\tau^2}\sum y_i^2 + \frac{\eta}{\tau^2}\sum y_i - \frac{n\eta^2}{2\tau^2} - n\log\tau\right\} \end{aligned}$$

Thus, (x,y) is distributed as $f(x)f(y)$ with parameter space $\theta = \{\xi, \eta, \sigma, \tau : -\infty < \xi, \eta < \infty, 0 < \sigma, \tau\}$. The sufficient statistics $T = \{\sum x_i, \sum x_i^2, \sum y_i, \sum y_i^2\}$ is also minimal since (x,y) belongs to exponential family.

(b) If $\sigma = \tau$, then the parameter space becomes

$$\Theta = \{\xi, \eta, \sigma : -\infty < \xi, \eta < \infty, 0 < \sigma < \infty\}$$

which contains a 3-dimensional triangle in it. The joint distribution of (x,y) is

$$\begin{aligned} f(x,y) &= \frac{1}{(2\pi)^{(m+n)/2}} \exp\Big\{-\frac{1}{2\sigma^2}\left(\sum x_i^2 + \sum y_i^2\right) \\ &\quad + \frac{\xi}{\sigma^2}\sum x_i + \frac{\eta}{\sigma^2}\sum y_i - \frac{m\xi^2 + n\eta^2}{2\sigma^2} - (m+n)\log\sigma\Big\} \end{aligned}$$

with natural parameter

$$\{\eta_1, \eta_2, \eta_3 : \eta_1 = \frac{1}{2\sigma^2}, \eta_2 = \frac{\xi}{\sigma^2}, \eta_3 = \frac{\eta}{\sigma^2}\}$$

This distribution belongs to exponential family with full rank. By theorem 6.22, $\{\sum x_i^2 + \sum y_i^2, \sum x_i, \sum y_i\}$is complete.

(c) $\xi = \eta$ and ξ, σ, η are arbitrary. The joint distribution of (x,y) is

$$\begin{aligned} f(x,y) &= \frac{1}{(2\pi)^{(m+n)/2}} \exp\Big\{-\frac{1}{2\sigma^2}\sum x_i^2 - \frac{1}{2\tau^2}\sum y_i^2 \\ &\quad + \frac{\xi}{\sigma^2}\sum x_i + \frac{\eta}{\tau^2}\sum y_i - \frac{m\xi^2}{2\sigma^2} - m\log\sigma - \frac{n\eta^2}{2\sigma^2} - n\log\tau\Big\} \end{aligned}$$

This belongs to exponential family which is full rank. Therefore, the minimal sufficient statistics is $(\sum x_i, \sum x_i^2, \sum y_i, \sum y_i^2)$.

To show that the parameter space is full rank, let

$$\eta_1 = -\frac{1}{2\sigma^2}, \eta_2 = -\frac{1}{2\tau^2}, \eta_3 = \frac{\xi}{\sigma^2}, \eta_4 = \frac{\eta}{\tau^2}$$

The natural parameter space is

$$\Theta = \{(\eta_1, \eta_2, \eta_3, \eta_4) : \eta_1 < 0, \eta_2 < 0, \frac{\eta 3}{\eta_4} = \frac{\eta_1}{\eta_2}\}$$

If we pick

$$\begin{aligned} \eta^{(1)} &= (-1/2, -1/2, 1, 1) \\ \eta^{(2)} &= (-1/4, -1/4, 2, 2) \\ \eta^{(3)} &= (-1/9, -1/3, 3, 9) \\ \eta^{(4)} &= (-1/16, -1/16, 4, 16) \\ \eta^{(5)} &= (-1/25, -1/25, 5, 5) \end{aligned}$$

then the 4×4 matrix

$$\det \begin{pmatrix} \eta^{(2)} - \eta^{(1)} \\ \eta^{(3)} - \eta^{(1)} \\ \eta^{(4)} - \eta^{(1)} \\ \eta^{(5)} - \eta^{(1)} \end{pmatrix} \neq 0$$

Thus it's full rank.

Question 4 (Problem 6.20)

(a)

$$\begin{aligned} p_\theta(x) &= \frac{1}{\sqrt{2\pi\theta^2}} \exp\left\{-\frac{1}{2\theta}\left(\sum x_i^2 - 2\theta \sum x_i + n\theta^2\right)\right\} \\ &= \frac{1}{\sqrt{2\pi\theta^2}} \exp\left\{-\frac{1}{\theta}\left(\frac{1}{2}\sum x_i^2\right) - \sum x_i + \frac{n}{2}\theta\right\} \end{aligned}$$

$T_1(x) = (\sum x_i, \sum x_i^2)$ is sufficient. Consider $T_2(x) = \sum x_i^2$ which is sufficient statistic also, but for $x \neq y \quad \exists\; T_2(x) = T_2(y)\ s.t.\ T_1(x) \neq T_1(y)$. Thus $T_1(x)$ is not minimal.

(b) To show $(\sum x_i, \sum x_i^2, \sum x_i^3)$ is minimal sufficient, we need to show that

$$\exists\, \eta^{(j)} = (\eta_1^{(j)}, \eta_2^{(j)}, \eta_3^{(j)}), \quad j = 0, 1, 2, 3$$

such that

$$\left(\frac{p(x, \eta^{(1)})}{p(x, \eta^{(0)})}, \frac{p(x, \eta^{(2)})}{p(x, \eta^{(0)})}, \frac{p(x, \eta^{(3)})}{p(x, \eta^{(0)})} \right)$$

is a one-to-one transformation of $T = (\sum x_i, \sum x_i^2, \sum x_i^3)$.

From the pdf:

$$C \exp(-n\theta^4) \exp\left(4\theta^3 \sum x_i - 6\theta^2 \sum x_i^2 + 4\theta \sum x_i^3 - \sum x_i^4\right)$$

we know the natural parameters are

$$4\theta^3 = \eta_1, -6\theta_2 = \eta_2, 4\theta = \eta_3$$

and

$$\left(\sum_{i=1}^{3} (\eta_i^{(1)} - \eta_i^{(0)}) T_i(x), \sum_{i=1}^{3} (\eta_i^{(2)} - \eta_i^{(0)}) T_i(x), \sum_{i=1}^{3} (\eta_i^{(3)} - \eta_i^{(0)}) T_i(x) \right)$$

is minimal sufficient for any $\eta^{(0)}, \eta^{(1)}, \eta^{(2)}, \eta^{(3)} \in \Theta$.

It can also be written as

$$\begin{pmatrix} \eta_1^{(1)} - \eta_1^{(0)} & \eta_2^{(1)} - \eta_2^{(0)} & \eta_3^{(1)} - \eta_3^{(0)} \\ \eta_1^{(2)} - \eta_1^{(0)} & \eta_2^{(2)} - \eta_2^{(0)} & \eta_3^{(2)} - \eta_3^{(0)} \\ \eta_1^{(3)} - \eta_1^{(0)} & \eta_2^{(3)} - \eta_2^{(0)} & \eta_3^{(3)} - \eta_3^{(0)} \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \\ T_3 \end{pmatrix}$$

Let

$$\eta^{(0)} = (0, 0, 0),\ \eta^{(1)} = (1/2, -3/2, 2),\ \eta^{(2)} = (4, -6, 4),\ \eta^{(3)} = (32, -24, 8)$$

then it's easy to see that $(\eta^{(1)} - \eta^{(0)}, \eta^{(2)} - \eta^{(0)}, \eta^{(3)} - \eta^{(0)})^t$ span E_3 and the determinant is $72 \neq 0$. Thus, (T_1, T_2, T_3) is minimal sufficient.

Question 5 (Problem 6.21)

Proof: $T = \left(\sum x_i, \sum x_i^2\right)$

$$E\left(\frac{\sum x_i}{n}\right) = E(\bar{x}) = Var(\bar{x}) + \left[E(\bar{x})\right]^2 = \frac{\xi^2}{n} + \xi^2 = \frac{n+1}{n}\xi^2$$

$$\Rightarrow \qquad E\left[\frac{n}{n+1}\left(\frac{\sum x_i}{n}\right)^2\right] = \xi^2$$

$$E\left(\sum x_i^2\right) = nE(x_i^2) = 2n\xi^2 \quad \Rightarrow \quad E\left(\frac{\sum x_i^2}{2n}\right) = \xi^2$$

$$\Rightarrow \qquad E\, f(T(x)) = 0 \quad \text{where} \quad f(T(x)) = \frac{n}{n+1}\left(\frac{\sum x_i}{n}\right)^2 - \frac{\sum x_i^2}{2n}$$

But $f(T) \neq 0 \quad a.e.$, thus $T = \left(\sum x_i, \sum x_i^2\right)$ is not complete.