STAT 543 Homework 4

Yifan Zhu

March 1, 2017

1.

$$\begin{aligned} f(x_1, x_2, \ldots, x_n|\theta) &= \prod_{i=1}^{n} \mathrm{e}^{i\theta - x_i} \cdot \mathbf{1}\{x_i \geq i\theta\} \\ &= \prod_{i=1}^{n} \mathrm{e}^{i\theta - x_i} \cdot \mathbf{1}\left\{\frac{x_i}{i} \geq \theta\right\} \\ &= \mathrm{e}^{-\sum_{i=1}^{n} x_i} \mathrm{e}^{\frac{n(n+1)\theta}{2}} \mathbf{1}\left\{\theta \leq \min_i \left(\frac{x_i}{i}\right)\right\} \end{aligned}$$

By Factorization Theorem, $T = \min_i \left(\frac{X_i}{i}\right)$ is sufficient statistic.

2.

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) &= \prod_{i=1}^{n} f(x_i|\mu, \sigma) \\ &= \prod_{i=1}^{n} \frac{1}{\sigma} \mathrm{e}^{-(x_i-\mu)/\sigma} \cdot \mathbf{1}\{x_i > \mu\} \\ &= \frac{1}{\sigma^n} \mathrm{e}^{-\frac{1}{\sigma}\sum_{i=1}^{n} x_i} \mathrm{e}^{\frac{n\mu}{\sigma}} \cdot \mathbf{1}\{\mu < x_{(1)}\} \end{aligned}$$

By Factorization Theorem, $\boldsymbol{S} = (\sum_{i=1}^{n} X_i, X_{(1)})$ is sufficient statistic for (μ, σ).

3. (b)

$$\begin{aligned} \frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)} &= \frac{\prod_{i=1}^{n} \mathrm{e}^{-(x_i-\theta)} \mathbf{1}\{\theta < x_i\}}{\prod_{i=1}^{n} \mathrm{e}^{-(y_i-\theta)} \mathbf{1}\{\theta < y_i\}} \\ &= \mathrm{e}^{\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i} \frac{\mathbf{1}\{\theta < x_{(1)}\}}{\mathbf{1}\{\theta < y_{(1)}\}} \end{aligned}$$

It is a constant as a function of θ if and only if $x_{(1)} = y_{(1)}$. Thus X_1 is minimal sufficient for θ.

(e)

$$\frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)} = \mathrm{e}^{\sum_{i=1}^{n} |y_i-\theta| - \sum_{i=1}^{n} |x_i-\theta|}$$

Let $x_{(1)}, \ldots, x_{(n)}$ and $y_{(1)}, \ldots, y_{(n)}$ be order statistics of $\{x_i\}_{i=1}^{n}$ and $\{y_i\}_{i=1}^{n}$. Then denote

$x_{(k(\theta))} \le \theta < x_{(k(\theta)+1)}$, $y_{(l(\theta))} \le \theta < y_{(l(\theta)+1)}$, $x(0) = y(0) = -\infty$, $x_{(n+1)} = y_{(n+1)} = \infty$. Then

$$\begin{aligned}
\sum_{i=1}^{n} |y_i - \theta| - \sum_{i=1}^{n} |x_i - \theta| &= -\sum_{i=1}^{k(\theta)} (\theta - x_{(i)}) - \sum_{i=k(\theta)+1}^{n} (x_{(i)} - \theta) + \sum_{i=1}^{l(\theta)} (\theta - y_{(i)}) + \sum_{i=k(\theta)+1}^{n} (y_{(i)} - \theta) \\
&= -k(\theta)\theta + (n - k(\theta))\theta - \sum_{i=k(\theta)+1}^{n} x_{(i)} + \sum_{i=1}^{k(\theta)} x_{(i)} \\
&\quad + l(\theta)\theta - (n - l(\theta))\theta + \sum_{i=l(\theta)+1}^{n} y_{(i)} - \sum_{i=1}^{l(\theta)} y_{(i)} \\
&= 2(l(\theta) - k(\theta))\theta + \left(\sum_{i=l(\theta)+1}^{n} y_{(i)} - \sum_{i=1}^{l(\theta)} y_{(i)} - \sum_{i=k(\theta)+1}^{n} x_{(i)} + \sum_{i=1}^{k(\theta)} x_{(i)} \right)
\end{aligned}$$

When θ is in an interval where no x_i or y_i exist, the right term in the parentheses above is a constant. Then we need $k(\theta) - l(\theta) = 0$ to make expression above a constant in this interval. This means for any θ, we should have the same number of sample points in x_i and y_i ahead of θ. This is equivalent to $x_{(i)} = y_{(i)}$ for all $i = 1, 2, \ldots, n$. Hence the minimal sufficient statistic is

$$\boldsymbol{S} = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$$

4. $T(\boldsymbol{X}) = (X_{(1)}, X_{(n)})$, and define

$$u(T(\boldsymbol{X})) = u(X_{(1)}, X_{(n)}) = X_{(n)} - X_{(1)} - \frac{n-1}{n+1}$$

We know

$$X_{(n)} - X_{(1)} \sim \text{Beta}(n-1, 2)$$

Then

$$E(X_{(n)} - X_{(1)}) = \frac{n-1}{n+1} - \frac{n-1}{n+1} = 0$$

However

$$P(u(T(\boldsymbol{X})) = 0) = P\left(X_{(n)} - X_{(1)} = \frac{n-1}{n+1} \right) = 0$$

Thus $(X_{(1)}, X_{(n)})$ is not complete.

5. (a)

$$E_p(T(X_1, X_2, \ldots, X_{n+1})) = P_p\left(\sum_{i=1}^{n} X_i > X_{n+1} \right) \times 1 = h(p)$$

(b)

$$f(x) = p^x (1-p)^{1-x} = \left(\frac{p}{1-p} \right)^x (1-p) = (1-p) \mathrm{e}^{x \log \frac{p}{1-p}}$$

$\{\log \frac{p}{1-p} : p \in (0,1)\} \supseteq (0,1)$, then $\sum_{i=1}^{n+1} X_i$ is sufficient and complete.

As T is unbiased, then $E_p(T | \sum_{i=1}^{n+1} X_i)$ is UMVUE.

$$E_p(T | \sum_{i=1}^{n} X_i = s) = \frac{P_p(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = s)}{P_p(\sum_{i=1}^{n+1} X_i = s)}$$

$$P_p(\sum_{i=1}^{n+1} X_i = s) = \binom{n+1}{s} p^s (1-p)^{n+1-s}$$

$$\begin{aligned}
&P_p(\sum_{i=1}^{n} > X_{n+1}, \sum_{i=1}^{n+1} X_i = s)\\
=&P_p(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = s | X_{n+1} = 0) P_p(X_{n+1} = 0)\\
&+ P_p(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = s | X_{n+1} = 1) P_p(X_{n+1} = 1)\\
=&P_p(\sum_{i=1}^{n} X_i > 0, \sum_{i=1}^{n} X_i = s | X_{n+1} = 0)(1-p) + P_p(\sum_{i=1}^{n} X_i > 1, \sum_{i=1}^{n} X_i = s-1 | X_{n+1} = 1)p\\
=&P_p(\sum_{i=1}^{n} X_i > 0, \sum_{i=1}^{n} X_i = s)(1-p) + P_p(\sum_{i=1}^{n} X_i > 1, \sum_{i=1}^{n} X_i = s-1)p\\
=&\begin{cases} 0 & s = 0 \\ \binom{n}{s} p^s (1-p)^{n-s}(1-p) + 0 & s = 1, 2 \\ \binom{n}{s} p^s (1-p)^{n-s}(1-p) + \binom{n}{s-1} p^{s-1}(1-p)^{n-s+1} p \end{cases}\\
=&\begin{cases} 0 & s = 0 \\ \binom{n}{s} p^s (1-p)^{n+1-s} + 0 & s = 1, 2 \\ \binom{n}{s} p^s (1-p)^{n+1-s} + \binom{n}{s-1} p^s (1-p)^{n+1-s} & 2 < s \le n+1 \end{cases}
\end{aligned}$$

Hence

$$T^*(\sum_{i=1}^{n+1} X_i = s) = \begin{cases} 0 & s = 0 \\ \frac{\binom{n}{s} p^s (1-p)^{n+1-s}}{\binom{n+1}{s} p^s (1-p)^{n+1-s}} = \frac{\binom{n}{s}}{\binom{n+1}{s}} = \frac{n+1-s}{n+1} & s = 1, 2 \\ \frac{\binom{n}{s} p^s (1-p)^{n+1-s} + \binom{n}{s-1} p^s (1-p)^{n+1-s}}{\binom{n+1}{s} p^s (1-p)^{n+1-s}} = \frac{\binom{n}{s} + \binom{n}{s-1}}{\binom{n+1}{s}} = 1 & 2 < s \le n+1 \end{cases}$$

6.

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \mathrm{e}^{-x/\beta}$$

$\{-1/\beta : \beta > 0\} \supseteq (-2, -1)$. Thus $\sum_{i=1}^{n} X_i$ is sufficient and complete for β. And $\sum_{i=1}^{n} X_i \sim \Gamma(n\alpha, \beta)$. Let $T(\sum_{i=1}^{n} X_i) = \frac{n\alpha - 1}{\sum_{i=1}^{n} X_i}$ when $n\alpha - 1 > 0$. Then

$$E(\frac{n\alpha - 1}{\sum_{i=1}^{n} X_i}) = (n\alpha - 1)E(\frac{1}{\sum_{i=1}^{n} X_i}) = (n\alpha - 1)\int_0^\infty \frac{1}{t} \frac{1}{\Gamma(n\alpha)\beta^{n\alpha}} t^{n\alpha} \mathrm{e}^{-t/\beta} \mathrm{d}t = (n\alpha - 1)\frac{1}{(n\alpha - 1)\beta} = \frac{1}{\beta}$$

Hence $\frac{n\alpha - 1}{\sum_{i=1}^{n} X_i}$ is UMVUE of $1/\beta$.