Homework#1
Name: 范又方 Fan, You-fang
Student number: 106225022

Exercise 6.3 [P300]
Let $X_1,\ldots,X_n$ be a random sample from the pdf

$$f(x|\mu,\sigma) = \frac{1}{\sigma}e^{\frac{-(x-\mu)}{\sigma}} \ , \ \mu < x < \infty , 0 < \sigma < \infty.$$

Find a two-dimensional sufficient statistic for (μ, σ).

Sol :

The joint pdf of the sample $\mathbf{X}$ is

$$f(\mathbf{x}|\mu,\sigma) = \prod_{i=1}^{n}\frac{1}{\sigma} \exp(-\frac{x_i-\mu}{\sigma}) \quad , \quad \mu < x_i < \infty , 0 < \sigma < \infty , i = 1,2,\ldots,n$$

We can use the indicator function. $I_A(x)$ is the indicator function of the set A; that is, it is equal to 1 if $x \in A$ and equal to 0 otherwise. Now $\mu < x_i < \infty$, $i = 1,2,\ldots,n$ if and only if $\min\{x_1, x_2, \ldots, x_n\} > \mu$.
So we have

$$f(\mathbf{x}|\mu,\sigma) = \prod_{i=1}^{n}\frac{1}{\sigma} \exp(-\frac{x_i-\mu}{\sigma})I_{(\mu,\infty)}(x_i)$$

$$= \frac{1}{\sigma^n} \exp(-\frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma})I_{(-\infty,\mathrm{X}_{(1)})}(\mu)$$

$$=\frac{1}{\sigma^n} \exp(\frac{n\mu}{\sigma}) \exp(-\frac{\sum_{i=1}^{n}x_i)}{\sigma})I_{(-\infty,\mathrm{X}_{(1)})}(\mu)$$

Since $f(x|\mu,\sigma) = h(x)g(T(\mathbf{X})|\mu,\sigma)$, where $h(x) = 1$ and

$g(T(\mathbf{X})|\mu,\sigma) = \frac{1}{\sigma^n} \exp(\frac{n\mu}{\sigma}) \exp(-\frac{\sum_{i=1}^{n}x_i)}{\sigma})I_{(-\infty,\mathrm{X}_{(1)})}(\mu)$, $T_1(\mathbf{X}) = \mathrm{X}_{(1)}$, $T_2(\mathbf{X}) = \sum_{i=1}^{n}\mathrm{X}_i$

Therefore, by the Factorization Theorem, $(\mathrm{X}_{(1)} , \sum_{i=1}^{n}\mathrm{X}_i)$ is a two-dimensional sufficient statistic for (μ,σ).

Exercise 6.5 [P300]

Let $X_1,\ldots,X_n$ be independent random variables with pdfs

$$f(x_i|\theta) = \begin{cases} \frac{1}{2i\theta} & -i(\theta-1) < x_i < i(\theta+1) \\ 0 & otherwise, \end{cases}$$

where $\theta > 0$. Find a two-dimensional sufficient statistic for θ.

Sol :

The joint pdf of the sample **X** is

$$f(\mathbf{x}|\theta) = \prod_{i=1}^{n} \frac{1}{2i\theta} \quad , \quad -i(\theta-1) < x_i < i(\theta+1) \ , \ \theta > 0$$

Now $-i(\theta-1) < x_i < i(\theta+1)$, using the indicator function, we can see the inequality is $-i(\theta-1) < \min x_i$ and $\max x_i < i(\theta+1)$.

$$\Rightarrow \begin{cases} -i(\theta-1) < \min x_i \\ i(\theta+1) > \max x_i \end{cases} = \begin{cases} \theta - 1 > -\min\frac{x_i}{i} \\ \theta + 1 > \max\frac{x_i}{i} \end{cases} = \begin{cases} \theta > 1 - \min\frac{x_i}{i} \\ \theta > \max\frac{x_i}{i} - 1 \end{cases}$$

So we have

$$f(\mathbf{x}|\theta) = \prod_{i=1}^{n} \frac{1}{2i\theta} I_{(-i(\theta-1),\, i(\theta+1)\ )}(x_i)$$

$$= \left(\frac{1}{2\theta}\right)^n \prod_{i=1}^{n} \frac{1}{i} I_{\left(1-\min\frac{x_i}{i},\infty\right)}(\theta) I_{\left(\max\frac{x_i}{i}-1,\infty\right)}(\theta)$$

Since $f(x|\mu,\sigma) = h(x)g(T(\mathbf{X})|\mu,\sigma)$, where $h(x) = \prod_{i=1}^{n}\frac{1}{i}$ and $g(T(\mathbf{X})|\theta) = \left(\frac{1}{2\theta}\right)^n I_{\left(1-\min\frac{x_i}{i},\infty\right)}(\theta) I_{\left(\max\frac{x_i}{i}-1,\infty\right)}(\theta)$, $T_1(\mathbf{X}) = \min\frac{X_i}{i}$, $T_2(\mathbf{X}) = \max\frac{X_i}{i}$

Therefore, by the Factorization Theorem, $(\min\frac{X_i}{i}$, $\max\frac{X_i}{i})$is a two-dimensional sufficient statistic for θ.

Exercise 6.6 [P300]

Let $X_1,\ldots,X_n$ be a random sample from a gamma(α, β) population. Find a two-dimensional sufficient statistic for (α, β).

Sol :

The joint pdf of the sample $\mathbf{X}$ is

$$f(\mathbf{x}|\alpha,\beta) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} {x_i}^{\alpha-1} \exp(-\frac{x_i}{\beta})$$

$$= \left(\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\right)^{n} (\prod_{i=1}^{n} x_i)^{\alpha-1} \exp(-\frac{\sum_{i=1}^{n} x_i}{\beta})$$

Since $f(x|\alpha,\beta) = h(x)g(T(\mathbf{X})|\alpha,\beta)$, where $h(x) = 1$ and $g(T(\mathbf{X})|\alpha,\beta) = \left(\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\right)^{n} (\prod_{i=1}^{n} x_i)^{\alpha-1} \exp(-\frac{\sum_{i=1}^{n} x_i}{\beta})$, $T_1(\mathbf{X}) = \prod_{i=1}^{n} X_i$, $T_2(\mathbf{X}) = \sum_{i=1}^{n} X_i$

Therefore, by the Factorization Theorem, $(\prod_{i=1}^{n} X_i \ , \sum_{i=1}^{n} X_i)$ is a two-dimensional sufficient statistic for (α, β).

Exercise 6.7 [P300]

Let $f(x, y|\theta_1, \theta_2, \theta_3, \theta_4)$ be the bivariate pdf for the uniform distribution on the rectangle with lower left corner (θ_1, θ_2) and upper right corner (θ_3, θ_4) in $\Re^2$. The parameters satisfy $\theta_1 < \theta_3$ and $\theta_2 < \theta_4$. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample from this pdf. Find a four-dimensional sufficient statistic for $(\theta_1, \theta_2, \theta_3, \theta_4)$.

Sol :

The joint pdf is

$$f(\mathbf{x}, \mathbf{y}|\theta) = \prod_{i=1}^{n} \frac{1}{\theta_3 - \theta_1} \frac{1}{\theta_4 - \theta_2} \quad , \quad \theta_1 \leq x_i \leq \theta_3 \ , \ \theta_2 \leq y_i \leq \theta_4$$

Using the indicator function,

$$\theta_1 \leq \min\{x_1, \ldots, x_n\} = x_{(1)} \, , \max\{x_1, \ldots, x_n\} = x_{(n)} \leq \theta_3 \, ,$$

$$\theta_2 \leq \min\{y_1, \ldots, y_n\} = y_{(1)} \, , \max\{y_1, \ldots, y_n\} = y_{(n)} \leq \theta_4$$

So we have

$$f(\mathbf{x}, \mathbf{y}|\theta) = \prod_{i=1}^{n} \frac{1}{\theta_3 - \theta_1} \frac{1}{\theta_4 - \theta_2} I_{(\theta_1, \theta_3)}(x_i) \, I_{(\theta_2, \theta_4)}(y_i)$$

$$= \left(\frac{1}{\theta_3 - \theta_1}\right)^n \left(\frac{1}{\theta_4 - \theta_2}\right)^n I_{(-\infty, x_{(1)})}(\theta_1) \, I_{(-\infty, y_{(1)})}(\theta_2) \, I_{(x_{(n)}, \infty)}(\theta_3) \, I_{(y_{(n)}, \infty)}(\theta_4)$$

Since $f(x, y|\theta) = h(x) g(T(\mathbf{X})|\theta)$, where $h(x) = 1$ and

$g(T(\mathbf{X})|\theta) =$

$$\left(\frac{1}{\theta_3 - \theta_1}\right)^n \left(\frac{1}{\theta_4 - \theta_2}\right)^n I_{(-\infty, x_{(1)})}(\theta_1) \, I_{(-\infty, y_{(1)})}(\theta_2) \, I_{(x_{(n)}, \infty)}(\theta_3) \, I_{(y_{(n)}, \infty)}(\theta_4) \, ,$$

$T_1(\mathbf{X}) = X_{(1)}$, $T_2(\mathbf{X}) = Y_{(1)}$, $T_3(\mathbf{X}) = X_{(n)}$, $T_4(\mathbf{X}) = Y_{(n)}$.

Therefore, by the Factorization Theorem, $(X_{(1)}, Y_{(1)}, X_{(n)}, Y_{(n)})$ is a four-dimensional sufficient statistic for $(\theta_1, \theta_2, \theta_3, \theta_4)$.

Exercise 6.8 [P301]

Let $X_1,\ldots,X_n$ be a random sample from a population with location pdf $f(x-\theta)$. Show that the order statistics, $T(X_1,\ldots,X_n) = (X_{(1)},\ldots,X_{(n)})$, are a sufficient statistic for θ and no further reduction is possible.

Sol :

We know $X_1,\ldots,X_n \sim f(x-\theta)$, $\theta \in \mathbb{R}$.

The joint pdf is

$$f(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} f(x_i-\theta) = \prod_{i=1}^{n} f\left(x_{(i)}-\theta\right)$$

Where $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$ are the order statistics of $X_1,\ldots,X_n$.

Since $f(x,y|\theta) = h(x)g(T(\mathbf{X})|\theta)$, where $h(x)=1$ and

$g(T(\mathbf{X})|\theta) = \prod_{i=1}^{n} f\left(x_{(i)}-\theta\right)$, $T_i(\mathbf{X}) = X_{(i)}$, $i = 1,\ldots,n$

Therefore, by the Factorization Theorem, $(X_{(1)},\ldots,X_{(n)})$ are a sufficient statistic for θ.

That no further reduction is possible without further restrictions on f. In this case, it suffices to notice the ratio

$$\frac{\prod_{i=1}^{n} f(x_{(i)}-\theta)}{\prod_{i=1}^{n} f(y_{(i)}-\theta)}$$

is general independent of θ only when $T(\mathbf{X}) = T(\mathbf{Y})$.