STAT 2006A Assignment 2
Due Time and Date: 9 p.m., October 26, 2023

Question 1

Let $X_1, \dots, X_n$ be independent random variables with densities

$$f_{X_i}(x|\theta) = \begin{cases} e^{i\theta - x}, & x \geq i\theta \\ 0, & x < i\theta. \end{cases}$$

Prove that $T = \min_i (X_i/i)$ is a sufficient statistic for θ.

Note that

$$f_{X_i}(x|\theta) = \begin{cases} e^{i\theta - x}, & x \geq i\theta \\ 0, & x < i\theta \end{cases}$$
$$= \begin{cases} e^{i\theta - x}, & x/i \geq \theta \\ 0, & x/i < \theta. \end{cases}$$

Therefore,

$$f(x_1, \dots, x_n|\theta) = \prod_{i=1}^{n} e^{i\theta - x_i} I_{(\theta,\infty)}(x_i/i) = \underbrace{e^{in\theta} I_{(\theta,\infty)}\left(\min_i\left(\frac{x_i}{i}\right)\right)}_{g\left(\min_i\left(\frac{x_i}{i}\right)|\theta\right)} \underbrace{e^{-n\bar{x}}}_{h(\mathbf{x})}.$$

By factorization theorem, $T = \min_i (X_i/i)$ is a sufficient statistic for θ.

Question 2

Let $X_1, \dots, X_n$ be a random sample from the pdf

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

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

$$f(x_1, \dots, x_n|\mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma} e^{-(x-\mu)/\sigma} I_{(\mu,\infty)}(x_i) = \underbrace{\left(\frac{e^{\mu/\sigma}}{\sigma}\right) e^{-n\bar{x}/\sigma} I_{(\mu,\infty)}(x_{(1)})}_{g(\bar{x}, x_{(1)}|\mu,\sigma)} \cdot \underbrace{1}_{h(\mathbf{x})}$$

By factorization theorem, $(\bar{X}, X_{(1)})$ is a sufficient statistic for (μ, σ).

Question 3

For each of the following distributions let $X_1, \dots, X_n$ be a random sample. Find a minimal sufficient statistic for θ.

(a) $f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}, \ -\infty < x < \infty, \ -\infty < \theta < \infty.$

(b) $f(x|\theta) = e^{-(x-\theta)}, \quad \theta < x < \infty, \ -\infty < \theta < \infty.$

(a)

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{(2\pi)^{-n/2}e^{-\sum_i(x_i-\theta)^2/2}}{(2\pi)^{-n/2}e^{-\sum_i(y_i-\theta)^2/2}}$$
$$= \exp\left\{-\frac{\sum_{i=1}^n(x_i^2-2\theta x_i+\theta^2)-\sum_{i=1}^n(y_i^2-2\theta y_i+\theta^2)}{2}\right\}$$
$$= \exp\left\{-\frac{\sum_{i=1}^n(x_i^2-y_i^2)-2\theta(\bar{x}-\bar{y})}{2}\right\}$$

This is constant as a function of θ if and only if $\bar{y}=\bar{x}$. Therefore, $\bar{X}$ is a minimal sufficient statistic for θ.

(b)

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\prod_{i=1}^n e^{-(x_i-\theta)}I_{(\theta,\infty)}(x_i)}{\prod_{i=1}^n e^{-(y_i-\theta)}I_{(\theta,\infty)}(y_i)} = \frac{e^{-n\bar{x}}e^{n\theta}I_{(\theta,\infty)}(x_{(1)})}{e^{-n\bar{y}}e^{n\theta}I_{(\theta,\infty)}(y_{(1)})}$$

The ratio is a constant function of θ if and only if $x_{(1)}=y_{(1)}$. Therefore, $X_{(1)}$ is a minimal sufficient statistic.

Question 4

The random variable X takes the values $0, 1, 2$ according to one of the following distributions:

	$P(X=0)$	$P(X=1)$	$P(X=2)$	
Distribution 1	p	$3p$	$1-4p$	$0<p<\frac{1}{4}$
Distribution 2	p	p^2	$1-p-p^2$	$0<p<\frac{1}{2}$

In each case determine whether the family of distributions of X is complete.

For Distribution 1,

$$E_p g(X) = \sum_{x=0}^{2} g(x)P(X=x) = pg(0)+3pg(1)+(1-4p)g(2).$$

If $E_p g(X)=0$, it is possible that $g(2)=0$ and $g(0)=-3g(1)\neq 0$. Therefore, Distribution 1 is not complete.

For Distribution 2,

$$E_p g(X) = \sum_{x=0}^{2} g(x)P(X=x) = g(0)p+g(1)p^2+g(2)(1-p-p^2)$$
$$= [g(1)-g(2)]p^2+[g(0)-g(2)]p+g(2).$$

If $E_p g(X)=0$, we have $g(1)-g(2)=g(0)-g(2)=g(2)=0$. This implies $g(0)=g(1)=g(2)$. Hence, Distribution 2 is comlpete.

Question 5

Let $X_1, \ldots, X_n$ be a random sample from the pdf $f(x|\mu) = e^{-(x-\mu)}$, where $-\infty < \mu < x < \infty$. Use Basu's Theorem to show that $X_{(1)}$ and S^2 are independent.

First, we show that $X_{(1)}$ is a sufficient statistic:

$$f(\mathbf{x}|\theta) = \prod_{i=1}^{n} e^{-(x_i-\mu)} I_{(\mu,\infty)}(x_i) = \underbrace{e^{-n\bar{x}}}_{h(\mathbf{x})} \underbrace{e^{n\mu} I_{(\mu,\infty)}\left(x_{(1)}\right)}_{g(x_{(1)}|\mu)}.$$

Therefore, $X_{(1)}$ is a sufficient statistic.

Second, we show that the distribution of $X_{(1)}$ is complete:

$$f_{X_{(1)}}(t) = n f_X(t)\left(1 - F_X(t)\right)^{n-1} = n e^{-(t-\mu)}\left[1 - \left(1 - e^{-(t-\mu)}\right)\right]^{n-1} I_{(\mu,\infty)}(t) = n e^{-n(t-\mu)} I_{(\mu,\infty)}(t).$$

$$Eg\left(X_{(1)}\right) = \int_{-\infty}^{\infty} g(t) f_{X_{(1)}}(t)dt = \int_{-\infty}^{\infty} g(t) n e^{-n(t-\mu)} I_{(\mu,\infty)}(t)dt = \int_{\mu}^{\infty} g(t) n e^{-n(t-\mu)} dt$$

If $Eg\left(X_{(1)}\right) = 0$, we have, for all μ,

$$\begin{aligned} \int_{\mu}^{\infty} g(t) n e^{-n(t-\mu)} dt &= 0 \\ \int_{\mu}^{\infty} g(t) e^{-nt} dt &= 0 \\ \frac{d}{d\mu}\int_{\mu}^{\infty} g(t) e^{-nt} dt &= 0 \\ -g(\mu) e^{-n\mu} &= 0 \\ g(\mu) &= 0 \end{aligned}$$

Hence, $g(\cdot) = 0$ and the distribution of $X_{(1)}$ is complete.

Third, we show that S^2 is ancillary statistic.

Note that $f(x|\mu)$ is a location family. Hence, we can write $X = Z + \mu$, where Z does not depend on μ. Now,

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^{n}(Z_i + \mu - \bar{Z} - \mu)^2}{n-1} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})^2}{n-1}.$$

Hence, S^2 does not depend on μ, and it is ancillary statistic.

By Basu's theorem, $X_{(1)}$ which is complete and sufficient and is independent of S^2 which is an ancillary statistic.