Unsupervised Learning HW1

Submitted By:

Parthiv Borgohain (pb25347)

Q1.

1) Given,

$$h(x_i;\theta) = \frac{1}{\theta\sqrt{\pi}} e^{-\frac{x_i^2}{\theta^2}}$$

$$\therefore h(x_1, x_2, \dots, x_n; \theta)$$

$$= \left(\frac{1}{\theta\sqrt{\pi}} e^{-\frac{x_1^2}{\theta^2}}\right)\left(\frac{1}{\theta\sqrt{\pi}} e^{-\frac{x_2^2}{\theta^2}}\right)\cdots\left(\frac{1}{\theta\sqrt{\pi}} e^{-\frac{x_n^2}{\theta^2}}\right)$$

$$= \left(\frac{1}{\theta\sqrt{\pi}}\right)^n e^{-\frac{1}{\theta^2}(x_1^2 + x_2^2 + \dots + x_n^2)}$$

If $T = x_1^2 + x_2^2 + \dots + x_n^2$, then

$$h(x_1, x_2, \dots, x_n; \theta) = \left(\frac{1}{\theta\sqrt{\pi}}\right)^n \left(e^{-T/\theta^2}\right) \cdot 1$$

$$\boxed{\begin{array}{l} \therefore \left(\frac{1}{\theta\sqrt{\pi}}\right)^n \left(e^{-T/\theta^2}\right) = g(T, \theta) \\ \text{and } 1 = h(x_1, x_2, \dots, x_n) \end{array}}$$

$$\boxed{\therefore \ f(x_1, x_2, \ldots, x_n; \theta) = g(T, \theta)\, h(x_1, x_2 \ldots x_n)}$$

$\therefore$ By Factorization theorem,

$\boxed{T = x_1^2 + x_2^2 + \ldots + x_n^2 \text{ is a sufficient one-dimensional statistic for } \theta}$

So, clearly via the working shown in the attached images above, we have shown via the factorization theorem that -

$T = x_1^2 + x_2^2 - \cdots + x_n^2$ is a sufficient one-dimensional statistic for θ.

Q2.

2) $h(x;\theta) = \begin{cases} \frac{1}{2\theta} & \text{for } -\theta < x_i < \theta \\ 0 & \text{otherwise} \end{cases}$

$i = 1, 2, \ldots, n$

where $\theta > 0$ is unknown

$$\therefore h(x_1, x_2, \ldots x_n; \theta)$$

$$= \left[\frac{1}{2\theta} \cdot I(-\theta < x_1 < \theta)\right] \left[\frac{1}{2\theta} \cdot I(-\theta < x_2 < \theta)\right] \cdots \left[\frac{1}{2\theta} \cdot I(-\theta < x_n < \theta)\right]$$

$$= \left(\frac{1}{2\theta}\right)^n \cdot I(-\theta < x_1 < \theta) \cdot I(-\theta < x_2 < \theta) \cdots I(-\theta < x_n < \theta)$$

$$= \left(\frac{1}{2\theta}\right)^n I\left[\max(|x_1|, |x_2|, \ldots, |x_n|) < \theta)\right]$$

$$\therefore \boxed{\left(\frac{1}{2\theta}\right)^n = h(x_1, x_2, \ldots x_n)}$$

and $\boxed{I[\max(|x_1|, |x_2|, \dots |x_n|) < \theta) = g(T, \theta)}$

where $\boxed{T = \max(|x_1|, |x_2|, \dots, |x_n|)}$

$\boxed{\therefore f(x_1, x_2, \dots, x_n; \theta) = g(T, \theta) h(x_1, x_2, \dots x_n)}$

$\therefore$ By factorization theorem,

$\boxed{T = \max(|x_1|, |x_2|, \dots, |x_n|) \text{ is a sufficient statistic for } \theta}$

$\boxed{\text{Hence, the data point with the maximum absolute value is a sufficient statistic}}$

So, clearly via the working shown in the attached images above, we have shown via the factorization theorem that -

$max(|x_1|, |x_2|, \dots, |x_n|)$ is a one-dimensional sufficient statistic for θ.

Q3.

3) Given, $f(x_i; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i - \mu)^2}{\sigma^2}}$,

for $-\infty < x_i < \infty$,

$i = 1, 2, \ldots, n$,

where $-\infty < \mu < \infty$, $\sigma > 0$

are unknown variables.

Now, $f(X_1, X_2, \ldots, X_n \mid T = Y_1, Y_2, \ldots Y_n)$

$$= \frac{f(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots Y_n)}{f(T = Y_1, Y_2, \ldots Y_n)} \rightarrow (i)$$

$\therefore$ We know that $Y_1, Y_2, \ldots, Y_n$ are $X_1, X_2, \ldots X_n$ in sorted form

$\therefore$ The joint probability of $X_1, X_2, \ldots, X_n,$
$Y_1, Y_2, \ldots, Y_n$ is the same as the probability of $X_1, X_2, \ldots, X_n$.

$$\therefore f(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n) = f(X_1, X_2, \ldots, X_n)$$

$\therefore$ From (i), we get -

$$\boxed{f(X_1, X_2, \ldots, X_n \mid T) = \frac{f(X_1, X_2, \ldots X_n)}{f(T = Y_1, Y_2, \ldots Y_n)} = 1/n!}$$

(because there are $n!$ possible permutations)

$\therefore$ $h(X_1, X_2, \ldots X_n | T)$ is independent of both $\underline{\sigma}$ and $\underline{\mu}$.

$\boxed{\therefore\ T = Y_1, Y_2, \ldots Y_n \text{ is a sufficient statistic}}$

In the above attached images, we stated that-

$$f(X_1, X_2, \ldots, X_n |\, \mathrm{T}) = \frac{f(X_1, X_2, \ldots, X_n)}{f(T = Y_1, Y_2, \ldots, Y_n)} = \frac{1}{n!}$$

As given as a hint in the question prompt, this is the case because we are trying to compute the probability of $X_1, X_2, \ldots, X_n$ occurring in a particular order given they occurred in the sequence

$\boldsymbol{Y_1}, \boldsymbol{Y_2}, \dots, \boldsymbol{Y_n}$ such that $\boldsymbol{Y_1} \leq \boldsymbol{Y_2} \leq \cdots, \leq \boldsymbol{Y_n}$**.** This probability is $\frac{\mathbf{1}}{\boldsymbol{n}!}$**.** This is because there are **n! possible permutations** and each of these permutations is equally likely. This gives us the probability of $\frac{\mathbf{1}}{\boldsymbol{n}!}$

This is independent of both **μ** and $\boldsymbol{\sigma}^2$ (the parameters). So, the statistic $\boldsymbol{T} = \boldsymbol{Y_1}, \boldsymbol{Y_2}, \dots, \boldsymbol{Y_n}$ **is sufficient** for given **μ** and $\boldsymbol{\sigma}^2$.