Lecture 3 - Sufficiency

Setup: $X_n \overset{iid}{\sim} f_\theta$ and $\theta \in \Theta$.

T captures all the info about θ in the sample $\underset{\sim}{X}$

Defn: A statistic $T = T(\underset{\sim}{X})$ is <u>sufficient</u> for θ if

$$f_{\underset{\sim}{X} \mid T = t}(\underset{\sim}{x}) \text{ is "free" of } \theta.$$

RV, RV, $\mathbb{R}$, $\underset{\sim}{x} \in \mathbb{R}^N$

conditional on knowing $T = t$

dist. formula doesn't depend on θ

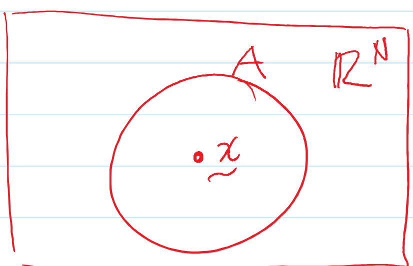
Ex. $\underset{\sim}{x} \in \mathbb{R}^N$, $A \subset \mathbb{R}^N$

$$P(\underset{\sim}{X} = \underset{\sim}{x} \text{ and } \underbrace{\underset{\sim}{X} \in A}_{\text{statement about } \underset{\sim}{X}})$$

<u>Two options:</u>

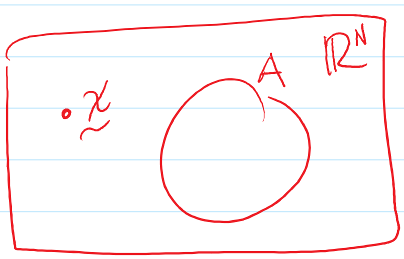
$\underset{\sim}{x} \in A$

#1



$$P(\underset{\sim}{X} = \underset{\sim}{x} \text{ and } \underset{\sim}{X} \in A) = P(\underset{\sim}{X} = \underset{\sim}{x}) \quad \text{if } \underset{\sim}{x} \in A$$

#2



$$P(\underset{\sim}{X} = \underset{\sim}{x} \text{ and } \underset{\sim}{X} \in A) = 0 \quad \text{if } \underset{\sim}{x} \notin A$$

intersection is empty

We could simply summarize as

$$P(\underset{\sim}{X} = \underset{\sim}{x}, \underset{\sim}{X} \in A) = P(\underset{\sim}{X} = \underset{\sim}{x})\, \mathbb{I}(\underset{\sim}{x} \in A)$$

or

or

$$\mathbb{P}(\underset{\sim}{X} = \underset{\sim}{x},\ \underbrace{T(\underset{\sim}{X}) = t}_{\underset{\sim}{X} \in A}) = \mathbb{P}(\underset{\sim}{X} = \underset{\sim}{x})\,\mathbb{1}(T(\underset{\sim}{x}) = t)$$

Ex. Let $X_1, X_2, X_3 \overset{iid}{\sim} \text{Beroulli}(\theta)$ $\quad \theta \in [0, 1]$

Let $T = X_1 + X_2 + X_3 \sim \text{Bin}(3, \theta)$

Is T sufficient for θ?

$$f_{X|Y=y} = \frac{f(x,y)}{f(y)}$$

$$f(\underset{\sim}{x} \mid T = t) = \frac{f_{\underset{\sim}{X},T}(\underset{\sim}{x}, t)}{f_T(t)}$$

$$= \frac{P(\underset{\sim}{X} = \underset{\sim}{x},\ T = t)}{P(T = t)}$$

$$= \frac{P(\underset{\sim}{X} = \underset{\sim}{x})\,\mathbb{1}(T = t)}{P(T = t)}$$

$$= \frac{\cancel{\theta^{\overset{t}{\sum x_n}}}\ \cancel{(1-\theta)^{3-\overset{t}{\sum x_n}}}\ \mathbb{1}(\overset{\sum x_n}{T} = t)}{\binom{3}{t}\cancel{\theta^t}\ \cancel{(1-\theta)^{3-t}}}$$

$$f(\underset{\sim}{x}) = \mathbb{P}(\underset{\sim}{X} = \underset{\sim}{x})$$
$$= \prod_{n=1}^{3} \theta^{x_n}(1-\theta)^{1-x_n}$$
$$= \theta^{\sum x_n}(1-\theta)^{3-\sum x_n}$$

$T \sim \text{Bin}(3, \theta)$

$$f(t) = \mathbb{P}(T = t)$$
$$= \binom{3}{t}\theta^t(1-\theta)^{3-t}$$

$t = \sum x_n$

$$= 1 \Big/ \binom{3}{t}\ \mathbb{1}(\textstyle\sum x_n = t)$$

no θ!

So T is sufficient for θ.

Ex. let $X_n \overset{iid}{\sim} f_\theta$

N-dimensional statistic

$$T = (X_{(1)}, X_{(2)}, X_{(3)}, \ldots, X_{(N)})$$

recall

$$f_T(\underset{\mathbb{R}^N}{t}) = n! \prod_{n=1}^{N} f(t_n) \qquad t = (t_1, \ldots, t_N)$$

Q: is T sufficient for θ?

only need to consider where $T(\underline{X}) = t$

$$f_{\underline{X}|T=t}(\underline{x}) = \frac{f_{\underline{X},T}(\underline{x}, t)}{f_T(t)}$$

$$= \frac{f_{\underline{X}}(\underline{x})}{f_T(t)} = \frac{\cancel{\prod_{n=1}^{N} f(x_n)}}{n! \cancel{\prod_{n=1}^{N} f(t_n)}} = \frac{1}{n!}$$

recall: t_n are a re-ordering of x_n

So since this is free of θ, T is sufficient.

Ex. $X_n \overset{iid}{\sim} N(\mu, 1)$

let $T = \bar{X}$. We know $\bar{X} \sim N(\mu, 1/N)$.

Q: is $\bar{X}$ sufficient for μ?

need to consider $\bar{x} = t$

$$f(\underline{x} \mid T=t) = \cdots = \frac{f_{\underline{X}}(\underline{x})}{f_T(t)}$$

$T = \bar{X} \sim N(\mu, 1/N)$

$$\frac{\prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_n-\mu)^2\right)}{\frac{1}{\sqrt{2\pi/N}} \exp\left(-\frac{N}{2}(\bar{x}-\mu)^2\right)}$$

$\prod_n e^{a_n} = e^{\sum a_n}$

$$= \frac{(2\pi)^{-N/2} \exp\left(-\frac{1}{2}\sum_{n=1}^{N}(x_n-\mu)^2\right)}{(2\pi/N)^{-1/2} \exp\left(-\frac{N}{2}(\bar{x}-\mu)^2\right)}$$

$$\sum_n (x_n-\mu)^2 = \sum_n (x_n^2 - 2\mu x_n + \mu^2)$$

$\bar{x} = \frac{1}{N}\sum x_n$

$$= \sum x_n^2 - 2\mu \underbrace{\sum x_n}_{N\bar{x}} + N\mu^2$$

$$N(\bar{x}-\mu)^2 = N\bar{x}^2 - N2\mu\bar{x} + N\mu^2$$

$$\propto \frac{\exp\left(-\frac{1}{2}\left(\sum x_n^2 \cancel{- 2\mu N\bar{x}} \cancel{+ N\mu^2}\right)\right)}{\exp\left(-\frac{1}{2}\left(N\bar{x}^2 \cancel{- 2\mu N\bar{x}} \cancel{+ N\mu^2}\right)\right)}$$

doesn't depend on μ.

So $T = \bar{X}$ is sufficient for μ.

So $T = \bar{X}$ is sufficient for μ.

Factorization Theorem:

T is sufficient for θ

iff

there is a function $g(\theta, T)$ and $h(\underset{\sim}{x})$

so that

$$\boxed{f(\underset{\sim}{x}) = g(\theta, T)\, h(\underset{\sim}{x})}$$

depends on θ and $\underset{\sim}{x}$ only through $T = T(\underset{\sim}{x})$

depends on $\underset{\sim}{x}$ not θ

Ex. $X_n \overset{iid}{\sim} N(\mu, 1)$

Show $\bar{X}$ sufficient for μ.

$$f(\underset{\sim}{x}) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_n - \mu)^2\right)$$

$$= (2\pi)^{-N/2} \exp\left(-\frac{1}{2} \sum_n (x_n - \mu)^2\right)$$

$$= (2\pi)^{-N/2} \exp\left(-\frac{1}{2}\left(\sum_n x_n^2 - 2\mu N \bar{x} + N\mu^2\right)\right)$$

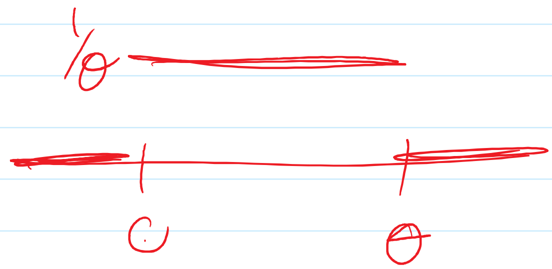
$$= \underbrace{(2\pi)^{-N/2} \exp\left(-\frac{1}{2} \sum x_n^2\right)} \underbrace{\exp\left(-\frac{1}{2}\left(-2\mu N \bar{x} + N\mu^2\right)\right)}$$

$$\underbrace{(2\pi)^{-N/2} \exp\left(-\tfrac{1}{2}\textstyle\sum x_n^2\right)}_{h(\underset{\sim}{x})} \underbrace{\exp\left(-\tfrac{1}{2}\left(-2\mu N \bar{x} + N\mu^2\right)\right)}_{g(\mu, \bar{X})}$$

$$g(\mu, t) = \exp\left(-\tfrac{1}{2}\left(-2\mu N t + N\mu^2\right)\right)$$

So since I can factor in this way, $\bar{X}$ is sufficient for θ.

Ex. Let $\boxed{X_n \overset{iid}{\sim} U(0, \theta)}$



Can I find a sufficient stat for θ?

$\mathbb{1}(A)\mathbb{1}(B) = \mathbb{1}(A \text{ and } B)$

$$f(\underset{\sim}{x}) = \prod_{n=1}^{N} \frac{1}{\theta} \mathbb{1}(0 < x_n < \theta)$$

$$= \theta^{-N} \underbrace{\prod_{n=1}^{N} \mathbb{1}(0 < x_n < \theta)}_{\text{all have to be true}}$$

all x_ns btwn 0 and θ.

$\Updownarrow$

min > 0

max $< \theta$

$$= \theta^{-N} \mathbb{1}(X_{(1)} > 0) \mathbb{1}(X_{(N)} < \theta)$$

$$= g(\theta, T)\, h(\underset{\sim}{x})$$

$h(\underset{\sim}{x}) = \mathbb{1}(X_{(1)} > 0)$

$$g(\theta, T) = \theta^{-N} \mathbb{1}(X_{(N)} < \theta)$$

$$g(\theta, T) = \theta^{-N} \mathbb{1}(X_{(N)} < \theta)$$

T

$$= \theta^{-N} \mathbb{1}(T < \theta) \quad \text{when } T = X_{(N)}$$

So $T = X_{(N)}$ is sufficient for θ.

Ex. $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$. $\lambda > 0$

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Sufficient stat for λ.

$$f(\underset{\sim}{x}) = \prod_n f(x_n) = \prod_n \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} = \left(\prod_n \frac{1}{x_n!}\right) \lambda^{N \frac{\sum x_n}{N}} e^{-N\lambda}$$

$$= h(\underset{\sim}{x})\, g(\lambda, T)$$

$$h(\underset{\sim}{x}) = \left(\prod_n \frac{1}{x_n!}\right) \; ; \; g(\lambda, t) = \lambda^t e^{-N\lambda}$$

$$T = \frac{\sum x_n}{N} = \bar{X}$$

So $\bar{X}$ is sufficient for λ.

Notice $\mathbb{E}[X_1] = \lambda$

and so $\mathbb{E}[\bar{X}] = \lambda$

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{N} \sum_n X_n\right] = \frac{1}{N} \sum_n \mathbb{E}[X_n] = \frac{1}{N} \sum_n \lambda = \lambda$$

Theorem: Exponential Families and Sufficiency

Theorem: Exponential Families and Sufficiency

If we have $X_n \overset{iid}{\sim} f_\theta$ and f_θ is an exponential family, then

$$f_\theta(\underset{\sim}{X}) = h(\underset{\sim}{X})\, c(\theta) \exp(T(\underset{\sim}{X})\, w(\theta)).$$

then T is sufficient for θ.

Pf. let $g(\theta, t) = c(\theta) \exp(t\, w(\theta))$

then

$$f_\theta(\underset{\sim}{x}) = h(\underset{\sim}{x})\, g(\theta, T)$$

So T is suffient for θ.

Ex. $X_n \overset{iid}{\sim} \mathrm{Exp}(\lambda)$ $\quad f(x) = \lambda e^{-\lambda x} \mathbb{I}(x>0)$

then we saw

$$f_\theta(\underset{\sim}{x}) = \prod_n \lambda e^{-\lambda x_n} \mathbb{I}(x_n > 0)$$

$$= \lambda^N e^{-\lambda \overset{T}{\sum x_n}} \prod_n \mathbb{I}(x_n > 0)$$
