1/27/22

Thursday, January 27, 2022 4:11 PM

Fisher-Neyman Factorization lemma

T is sufficient for ϑ iff there are functions

g & h such that ... $L(\vartheta, x_1, \cdots, x_n) = g(T(x_1, \cdots, x_n), \vartheta) \cdot h(x_1, \cdots, x_n)$

Example, let n iid $X_i \sim \text{Poisson}(\lambda)$ be or sample

$$f_{X_i}(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad L(\lambda \mid x_1, \cdots, x_n) = \frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \cdot \cdots \cdot \frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$$

$$= \frac{e^{-n\lambda}\lambda^{x_1+\cdots+x_n}}{x_1! \cdot \cdots \cdot x_n!} = \underbrace{e^{-n\lambda}\lambda^{x_1+\cdots+x_n}}_{g(T,\lambda)} \cdot \underbrace{\frac{1}{x_1! \cdot \cdots \cdot x_n!}}_{h(x_1,\cdots,x_n)}$$

thus $T = x_1 + \cdots + x_n$ is a sufficient statistic for λ

Example: let n iid $X_i \sim N(\mu, \sigma^2 = 1)$ be our sample

* We want to find a sufficient statistic for μ.

$$L(\mu \mid x_1, \cdots, x_n) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_1-\mu)^2} \cdot \cdots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_n-\mu)^2}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2} = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 - 2x_i\mu + \mu^2}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2}\left(\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2\right)}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2}\left(n\mu^2 - 2\mu\sum_{i=1}^{n} x_i\right)} \cdot e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2}$$

$$= \underbrace{\frac{1}{(2\pi)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2}(n\mu^2 - 2\mu \sum_{i=1}^{n} x_i)}}_{g(T,\mu)} \underbrace{\cdot e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2}}_{h(X_1,\cdots,X_n)}$$

So $T=\sum_{i=1}^{n} X_i$ is a sufficient statistic for μ. We can also say that $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ is a sufficent statistic for μ.

<u>Example</u>: let n iid $X_i \sim N(\mu=0, \sigma^2)$ be our random sample.

Find a Sufficient statistic for σ^2

$$L(\sigma^2 | X_1, \cdots, X_n) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}X_1^2} \cdot \cdots \cdot \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}X_n^2}$$

$$= \underbrace{\frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(X_1^2 + \cdots + X_n^2)}}_{g(T,\sigma^2)} \cdot \underbrace{1}_{h(X_1,\cdots,X_n)}$$

so $T = \sum_{i=1}^{n} X_i^2$ is a sufficient statistic for σ^2

<u>Example</u>: let n iid $X_i \sim N(\mu, \sigma^2)$ be our sample

We want to find a sufficent statistic for (μ, σ^2)

$$L((\mu, \sigma^2) | X_1, \cdots, X_n) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2\sigma^2}(X_1-\mu)^2} \cdot \cdots \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2\sigma^2}(X_n-\mu)^2}$$

$$L((\mu,\sigma^2)\mid X_1,\dots,X_n) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2\sigma^2}(X_1-\mu)} \cdot \dots \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\dots}$$

$$= \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \cdot e^{\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2} = \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \cdot e^{\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i^2-2X_i\mu+\mu^2)}$$

$$= \underbrace{\frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \cdot e^{\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{n} X_i^2 - 2\mu\sum_{i=1}^{n} X_i + n\mu^2\right)}}_{g((T_1,T_2);(\mu,\sigma^2))} \cdot \underbrace{1}_{h(X_1,\dots,X_n)}$$

So $(T_1, T_2) = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$ is

<u>jointly</u> sufficient for (μ, σ^2)

We also could've solved this problem form a 2^{nd} approach

$$L((\mu,\sigma^2)\mid X_1,\dots,X_n) = \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \cdot e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2}$$

$$= \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \cdot e^{\frac{-1}{2\sigma^2}\underline{\sum_{i=1}^{n}\left((X_i-\overline{X})+(\overline{X}-\mu)\right)^2}}$$

$$\sum_{i=1}^{n}(X_i-\overline{X})^2 + \underbrace{\sum_{i=1}^{n} 2(X_i-\overline{X})(\overline{X}-\mu)}_{=0} + \sum_{i=1}^{n}(\overline{X}-\mu)^2$$

$$L((\mu,\sigma^2)\mid X_1,\dots,X_n) = \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \cdot e^{\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{n}(X_i-\overline{X})^2 + \sum_{i=1}^{n}(\overline{X}-\mu)^2\right)}$$

$$L((\mu,\sigma^2)\mid X_1,\dots,X_n) = \underbrace{\frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \cdot e^{\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{n}(X_i-\bar{X})^2 + \sum_{i=1}^{n}(\bar{X}-\mu)^2\right)}}_{g((T_3,T_4),(\mu,\sigma^2))} \cdot \underbrace{1}_{h(X_1,\dots,X_n)}$$

note: $\sum_{i=1}^{n}(\bar{X}-\mu)^2 = n(\bar{X}-\mu)^2$

So the pair $\left(T_3=\bar{X},\ T_4=\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$ are also jointly sufficient for (μ,σ^2)

Note here $T_4=\frac{S^2}{n-1}$ so we can also say that $(T_5=\bar{X},\ T_6=S^2)$ are also jointly sufficient for (μ,σ^2)

Note: we can usually derive one pair from another

Note: Sufficiency is about information, and when we have enough information to find our parameters

---

Example

let n iid $U_i \sim \text{uniform}[0,\theta]$ be our sample

find a sufficient statistic for θ $\left\{f(x,\theta)=\frac{1}{\theta} \text{ for } 0\le x\le\theta\right\}$

$$L(\theta\mid U_1,\dots,U_n)=\frac{1}{\theta^n} \quad \text{for } 0\le U_i\le\theta$$

$$g(T,\theta)=\frac{1}{\theta^n} \;\;\&\;\; h(X_1,\dots,X_n)=1$$

$$g(T,\vartheta) = \frac{1}{\vartheta^n} \quad \& \quad h(X_1,\dots,X_n) = 1$$

So $T = \max\{u_1,\dots,u_n\}$ is a sufficient statistic for ϑ

What makes the Fischer-Neyman lemma True?

Suppose $L(\vartheta \mid X_1,\dots,X_n) = g(s(X_1,\dots,X_n),\vartheta)\cdot h(X_1,\dots,X_n)$

$$f_{(X_1,\dots,X_n\mid S)}\left(x_1,\dots,x_n \mid S = s(x_1,\dots,x_n)\right)$$

$$= \frac{f_{(X_1,\dots,X_n)}(x_1,\dots,x_n)}{f_S(s)} = \frac{\overset{\text{constant}}{\boxed{g(s(x_1,\dots,x_n))}}\cdot h(x_1,\dots,x_n)}{\int g(s(x_1,\dots,x_n))\cdot h(x_1,\dots,x_n)\,dx_1\cdots dx_n}$$

$s(x_1,\dots,x_n) = S$

So if we can factor it implies sufficency

$h(x_1,\dots,x_n)$ doesnt depend on ϑ

Why does sufficency imply factorization?

$f_{X_1,\dots,X_n|T}(x_1,\dots,x_n|T(x_1,\dots,x_n))$ does not depend on ϑ by the definition

$$= \frac{f_{X_1,\dots,X_n}(x_1,\dots,x_n)}{f_T(T(x_1,\dots,x_n))} = h(x_1,\dots,x_n)$$

$$L(\vartheta|x_1,\dots,x_n) = \underbrace{f_T(T(x_1,\dots,x_n))}_{g(T,\vartheta)} \cdot \underbrace{h(x_1,\dots,x_n)}_{h(x_1,\dots,x_n)}$$