M378K: April 14^{th}, 2025.

More on Sufficient Statistics.

Review

If $Y_1, \ldots, Y_n \mid T$ ~ dist'n does not depend on the target parameter θ,

then, we say that T is sufficient for θ.

Theorem. The Fisher-Neyman Factorization Criterion

Let $Y_1, \ldots, Y_n$ be a random sample w/ the likelihood function

$$L(\theta; y_1, \ldots, y_n).$$

The statistic T is sufficient for θ if and only if

L can be expressed as

$$L(\theta; y_1, \ldots, y_n) = g(\theta, T(y_1, \ldots, y_n)) \cdot h(y_1, \ldots, y_n)$$

Example. Bernoulli $\quad \theta \leftrightarrow p$

$$L(p; y_1, y_2, \ldots, y_n) = p^{y_1}(1-p)^{1-y_1} \cdot \ldots \cdot p^{y_n}(1-p)^{1-y_n}$$

$$= p^{\overset{t}{\sum y_i}}(1-p)^{n-\overset{t}{\sum y_i}}$$

$$= p^t (1-p)^{n-t}$$

$$g(p,t) = p^t(1-p)^{n-t} \quad \text{and} \quad h \equiv 1$$

Example. Normal w/ a known σ. $\quad \theta \leftrightarrow \mu$

$$L(\mu; y_1, \ldots, y_n) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$$

$(y_1, \ldots, y_n)$

$$= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right)$$

$$= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_i y_i^2 - 2\mu \sum_i y_i + n\cdot\mu^2\right)\right)$$

$$= \underbrace{\frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_i y_i^2\right)}_{h(y_1,\dots,y_n)} \cdot \underbrace{\exp\left(-\frac{1}{2\sigma^2}\left(-2\mu \overbrace{\sum_i y_i}^{t} + n\mu^2\right)\right)}_{g(\mu,t)}$$

$\Rightarrow$ $T(Y_1,\dots,Y_n) = \sum_{i=1}^n Y_i$ is a sufficient statistic for μ.

Example. $\boxed{\text{Uniform.}}$

Say that $Y_1,\dots,Y_n$ is a random sample from $U(0,\theta)$ w/ θ unknown.

$$L(\theta; y_1,\dots,y_n) = \underbrace{\left(\frac{1}{\theta}\right)^n}_{g(\theta,t)} \underbrace{\boxed{\mathbb{1}_{\{0\le \min(y_1,\dots,y_n)\}}}}_{h(y_1,\dots,y_n)} \mathbb{1}_{\{\max(y_1,\dots,y_n)\le\theta\}}$$

$\max(y_1,\dots,y_n) = y_{(n)} = t$

$\Rightarrow$ We can propose $T(y_1,\dots,y_n) = \max(y_1,\dots,y_n) = y_{(n)}$ as our sufficient statistic.

Set $g(\theta,t) = \frac{1}{\theta^n}\mathbb{1}_{\{t\le\theta\}}$

and $h(y_1,\dots,y_n) = \mathbb{1}_{\{0\le\min(y_1,\dots,y_n)\}}$

Indeed, $T = Y_{(n)}$ is sufficient for θ $\square$

What if $U(\theta,\theta+1)$?

Hypothesis Testing.

Proof by Contradiction.

K... the claim we're trying to PROVE to be true

Q: What if K were not true?

Assume

$\boxed{\text{not } K}$

$\Downarrow$ .... $\Downarrow$

fact Q   fact (not Q)

These cannot coexist!

We say that we reached a **contradiction**!

$\Rightarrow\Leftarrow$

↯

Our assumption of not K was wrong!