

L can be expressed as

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

Example. Bernoulli.

$$\begin{aligned} L(p; y_1, \dots, y_n) &= p^{\sum y_i} (1-p)^{n-\sum y_i} \\ &= p^t (1-p)^{n-t} \end{aligned}$$

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

11/20/2024.

Example. Normal w/ a known σ .

$$\begin{aligned} L(\mu; y_1, \dots, y_n) &= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2\right) \\ &= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n y_i^2 - 2\mu \sum_{i=1}^n y_i + n\mu^2 \right)\right) \\ &= \underbrace{\frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum y_i^2\right)}_{h(y_1, \dots, y_n)} \cdot \underbrace{\exp\left(-\frac{1}{2\sigma^2} (-2\mu \sum y_i + n\mu^2)\right)}_{g(\mu, \sum_{i=1}^n y_i)} \\ &\quad \cdot \underbrace{T(y_1, \dots, y_n)}_{T(y_1, \dots, y_n)} \end{aligned}$$

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

Example. Uniform

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

$$L(\theta; y_1, \dots, y_n) = \frac{1}{\theta^n} \mathbf{1}_{\{0 \leq \min(y_1, \dots, y_n)\}}$$

$$\mathbf{1}_{\{\max(y_1, \dots, y_n) \leq \theta\}} \\ T(y_1, \dots, y_n)$$

\Rightarrow We can propose $T(y_1, \dots, y_n) = \max(y_1, \dots, y_n)$ as our sufficient statistic

Set $g(\theta, T) = \frac{1}{\theta^n} \cdot \mathbf{1}_{\{T \leq \theta\}}$

and

$$h(y_1, \dots, y_n) = \mathbf{1}_{\{0 \leq \min(y_1, \dots, y_n)\}}$$

Indeed, T is sufficient for θ .

$$U(\theta, \theta+1)$$

□

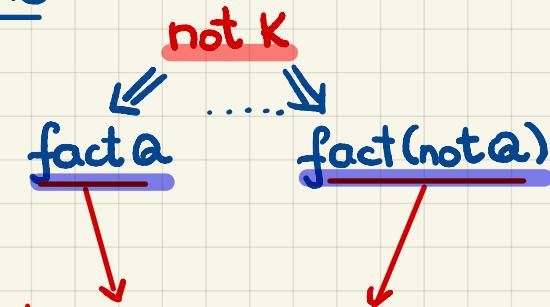
Hypothesis Testing.

Proof by Contradiction.

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

Q: What if K were not true?

Assume



These cannot coexist!

We say that we reached a contradiction!

$\Rightarrow \Leftarrow$



not K

Our assumption of was wrong!

Hypothesis Testing.

Claim we're trying to **SUBSTANTIATE**.

μ ... the population mean parameter
(say, the mean cholesterol level after treatment)

H_0 ... the null population mean (a number)
(say, the historical cholesterol level on average)

$$\mu < \mu_0$$

Alternative Hypothesis

Assume

$$\mu = \mu_0$$

Null

Hypothesis

collect data
statistical analysis

p-value

Figure out the probability of seeing the data that you saw (or something more extreme) if $\mu = \mu_0$

If this probability is "small", we have evidence against $\mu = \mu_0$

The smaller the probability, the STRONGER THE EVIDENCE.