L can be expressed as

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

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

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

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

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

$\boxed{11/20/2024.}$

Example. $\boxed{\text{Normal w/ a known } \sigma.}$

$$L(\mu; y_1, \dots, y_n) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right)$$

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

$$= \underbrace{\frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum y_i^2\right)}_{h(y_1, \dots, y_n)} \underbrace{\exp\left(-\frac{1}{2\sigma^2}\left(-2\mu \sum y_i + n\mu^2\right)\right)}_{g(\mu, \underset{T(y_1,\dots,y_n)}{\sum_{i=1}^{n} y_i})}$$

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

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

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

$$L(\theta; y_1, \dots, y_n) = \frac{1}{\theta^n} \boxed{\mathbb{1}_{\{0 \leq \min(y_1, \dots, y_n)\}}} \boxed{\mathbb{1}_{\{\underbrace{\max(y_1, \dots, y_n)}_{T(y_1, \dots, y_n)} \leq \theta\}}}$$

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

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

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

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

Indeed, T is sufficient for θ. $\square$

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

not K

⇙ fact@ ..... ⇘ fact(not@)

These cannot coexist!

we say that we reached a contradiction!

$\Rightarrow\Leftarrow$

Our assumption of not K was wrong!

Hypothesis Testing.

Claim we're trying to $\boxed{\text{SUBSTANTIATE}}$.

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

μ_0... the null population mean (a number) (say, the historical cholesterol level on average)

$\boxed{\mu < \mu_0}$ ← Alternative Hypothesis

Assume $\boxed{\mu = \mu_0}$ ← Null Hypothesis

⋮

collect data
statistical analysis

$\boxed{p\text{-value}}$

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

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

The smaller the probability, the STRONGER THE EVIDENCE.