# STA 522, Spring 2022
# Introduction to Theoretical Statistics II

Lecture 9

Department of Biostatistics
University at Buffalo

## NOTES

- Exam 2 this Thursday, April 7, 5-7 pm @ Kimball 720
- Will cover everything we discuss till this afternoon

## AGENDA

- Hypothesis testing, LRT
- LRT examples

# Review: likelihood ratio test

- Recall the **likelihood function** $L(\theta \mid \underline{x}) = f(\underline{x} \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$. The **likelihood ratio test (LRT) statistic** for testing $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_0^c$ is $\lambda(\underline{x}) = \dfrac{\sup_{\Theta_0} L(\theta \mid \underline{x})}{\sup_{\Theta} L(\theta \mid \underline{x})}$.
- Note that the likelihood ratio test statistic can be viewed as

$$\lambda(\underline{x}) = \frac{L(\hat{\theta}_0 \mid \underline{x})}{L(\hat{\theta} \mid \underline{x})} = \frac{\text{restricted maximization}}{\text{unrestricted maximization}},$$

  where $\hat{\theta}$ is the MLE obtained by maximizing $L(\theta \mid \underline{x})$ over the entire parameter space $\Theta$, and $\hat{\theta}_0$ is the MLE obtained by maximizing over the restricted parameter space $\Theta_0$.
- A **likelihood ratio test (LRT)** is any test that has a rejection region of the form

$$\{\underline{x} : \lambda(\underline{x}) \leq c\},$$

  where $c \in [0, 1]$.

**Example:** Let $X_1, X_2, \ldots, X_n \sim$ iid from a (location) exponential population with pdf $f(x \mid \theta) = e^{-(x-\theta)} I_{[\theta,\infty)}(x)$, where $\theta \in \Theta = \mathbb{R}$. Suppose we wish to test $H_0 : \theta \leq a$ vs. $H_1 : \theta > a$ where $a$ is a known value (e.g. 0) supplied by the experimenter. Find the LRT rejection region.

The likelihood function for $\theta$ is:

$$L(\theta \mid \underline{x}) = \prod_{i=1}^{n} e^{-(x_i-\theta)} \; I(x_i \geq \theta) = e^{-n(\overline{x}-\theta)} \; I(x_{(1)} \geq \theta)$$

$L(\theta \mid \underline{x})$ is an increasing function in $\theta$ for $\theta \in (-\infty, x_{(1)}]$. So unrestricted MLE is $\hat{\theta} = x_{(1)}$ so that $\sup_{\theta \in \Theta} L(\theta \mid \underline{x}) = L(x_{(1)} \mid \underline{x}) = e^{-n\left(\overline{x}-x_{(1)}\right)}$.

Under $H_0$, the restricted range $\theta \in \Theta_0 = (-\infty, a]$ MLE of $\theta$ is

$$\hat{\theta}_0 = \begin{cases} x_{(1)} & \text{if } \; x_{(1)} \leq a \\ a & \text{if } \; x_{(1)} > a \end{cases}$$

Therefore, LRT is:

$$\lambda(\underline{x}) = \begin{cases} 1 & \text{if } x_{(1)} \leq a \\ e^{-n\left(x_{(1)} - a\right)} & \text{if } x_{(1)} > a \end{cases}$$

Therefore the rejection region for the LRT is:

$$\{\underline{x} : \lambda(\underline{x}) \leq c\} = \left\{\underline{x} : x_{(1)} \geq a - \frac{\log c}{n}\right\}$$

for some $0 < c < 1$.

**NOTE:** The LRT rejection region depends on the data only through $X_{(1)}$. In the normal example discussed last week, the LRT rejection region depends on data only through $\overline{X}$.

# LRT and sufficiency

**Note:** Sufficiency means that all the information about $\theta$ in $\underline{x}$ is contained in a sufficient statistic $T(\underline{x})$. Intuitively, a test based on $T$ should be as good as the test based on the complete sample $\underline{X}$. The following theorem formalizes this.

## Theorem (8.2.4)

If $T(\underline{X})$ is a sufficient statistic for $\theta$ and $\lambda^*(t)$ and $\lambda(\underline{x})$ are the LRT statistics based on $T$ and $\underline{X}$, respectively, then

$$\lambda^*(T(\underline{x})) = \lambda(\underline{x})$$

for every $\underline{x}$ in the sample space.

**Proof:** Since $T(\underline{X})$ is a sufficient statistics, therefore by the Factorization theorem, we have

$$f(\underline{x} \mid \theta) = g(T(\underline{x}) \mid \theta)\ h(\underline{x})$$

Therefore

$$\begin{aligned}
\lambda(\underline{x}) &= \frac{\sup_{\Theta_0} L(\theta \mid \underline{x})}{\sup_{\Theta} L(\theta \mid \underline{x})} \\
&= \frac{\sup_{\Theta_0} f(\underline{x} \mid \theta)}{\sup_{\Theta} f(\underline{x} \mid \theta)} \\
&= \frac{\sup_{\Theta_0} g(T(\underline{x}) \mid \theta)\ h(\underline{x})}{\sup_{\Theta} g(T(\underline{x}) \mid \theta)\ h(\underline{x})} \\
&= \frac{\sup_{\Theta_0} g(T(\underline{x}) \mid \theta)}{\sup_{\Theta} g(T(\underline{x}) \mid \theta)} \\
&= \frac{\sup_{\Theta_0} L^*(\theta \mid T(\underline{x}))}{\sup_{\Theta} L^*(\theta \mid T(\underline{x}))} \\
&= \lambda^*(T(\underline{x}))
\end{aligned}$$

This completes the proof.

**Example:** Let $X_1, X_2, \ldots, X_n \sim$ iid from a population with pdf $f(x \mid \theta) = \theta x^{\theta-1} I_{(0,1)}(x)$, $\theta > 0$. Suppose we wish to test $H_0 : \theta = 1$ vs. $H_1 : \theta \neq 1$. Find the LRT rejection region.

Note at the outset that the restricted MLE is simply $\hat{\theta}_0 = 1$.

For $\theta \in \Theta = (0, \infty)$ the likelihood function is given by

$$L(\theta \mid \underline{x}) = \theta^n \left( \prod_{i=1}^{n} x_i \right)^{(\theta-1)} \implies \log L(\theta \mid \underline{x}) = n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log x_i$$

therefore

$$\frac{\partial \ L(\theta \mid \underline{x})}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i \gtreqless 0 \iff \theta \lesseqgtr -\frac{n}{\sum_{i=1}^{n} \log x_i}$$

Therefore, the MLE of $\theta$ is $\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log X_i}$.

Therefore the LRT statistic is

$$\begin{aligned}
\lambda(\underline{x}) &= \frac{\sup_{\Theta_0} L(\theta \mid \underline{x})}{\sup_{\Theta} L(\theta \mid \underline{x})} \\
&= \exp\left[n \log \theta_0 + (\theta_0 - 1) \sum_{i=1}^{n} \log x_i - n \log \hat{\theta} - (\hat{\theta} - 1) \sum_{i=1}^{n} \log x_i\right] \\
&= \exp\left[n \log\left(\frac{\theta_0}{\hat{\theta}}\right) + (\theta_0 - \hat{\theta}) \sum_{i=1}^{n} \log x_i\right]
\end{aligned}$$

Note that $\lambda(\underline{x})$ depends on $\underline{x}$ only through $\sum_{i=1} \log x_i$.

The rejection region of the LR test is given by:

$$\left\{\underline{x} : \exp\left[n \log\left(\frac{\theta_0}{\hat{\theta}}\right) + (\theta_0 - \hat{\theta}) \sum_{i=1}^{n} \log x_i\right] \leq c\right\}$$

**Example (LRT under the presence of nuisance parameters):** Let $X_1, X_2, \ldots, X_n \sim$ iid $\mathrm{N}(\mu, \sigma^2)$ (both parameters unknown). Suppose we wish to test $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$. Find the LRT rejection region.

Note that here $\sigma^2$ is a nuisance parameter.

The unrestricted MLEs of $\mu$ and $\sigma^2$ are $\hat{\mu} = \overline{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

Under $H_0$, the restricted MLE for $\mu$ is

$$\hat{\mu}_0 = \begin{cases} \overline{X} & \text{if } \overline{X} \leq \mu_0 \\ \mu_0 & \text{if } \overline{X} > \mu_0 \end{cases}$$

The corresponding MLE of $\sigma^2$ is

$$\hat{\sigma}_0^2 = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 & \text{if } \overline{X} \leq \mu_0 \\ \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2 & \text{if } \overline{X} > \mu_0 \end{cases}$$

The LRT statistic is given by:

$$\lambda(\underline{x}) = \begin{cases} 1 & \text{if } \overline{X} \leq \mu_0 \\ \dfrac{L(\mu_0, \sigma_0^2)}{L(\hat{\mu}, \hat{\sigma}^2)} & \text{if } \overline{X} > \mu_0 \end{cases}$$

The rejection region is given by

$$\{\underline{x} : \lambda(\underline{x}) \leq c\}$$

It can be shown that (HW, exercise 8.37) the above rejection region can be equivalently expressed as ($t$-test)

$$\overline{X} > \mu_0 + c' \sqrt{\frac{S^2}{n}}$$