Biostatistics 602 - Statistical Inference
Lecture 19
Likelihood Ratio Test

Hyun Min Kang

March 26th, 2013

Last Lecture

Describe the following concepts in your own words

- Hypothesis
- Null Hypothesis
- Alternative Hypothesis
- Hypothesis Testing Procedure
- Rejection Region
- Type I error
- Type II error
- Power function
- Size α test
- Level α test
- Likleihood Ratio Test

Example of Hypothesis Testing

Let $X_1, \cdots, X_n$ be changes in blood pressure after a treatment.

$$H_0 \;:\; \theta = 0$$
$$H_1 \;:\; \theta \neq 0$$

The rejection region $= \left\{ \mathbf{x} : \frac{\bar{x}}{s_{\mathbf{x}}/\sqrt{n}} > 3 \right\}$.

Decision

Truth		Accept H_0	Reject H_0
	H_0	Correct Decision	Type I error
	H_1	Type II error	Correct Decision

Power function

Definition - The power function

The power function of a hypothesis test with rejection region R is the function of θ defined by

$$\beta(\theta) = \Pr(\mathbf{X} \in R|\theta) = \Pr(\text{reject } H_0|\theta)$$

If $\theta \in \Omega_0^c$ (alternative is true), the probability of rejecting H_0 is called the power of test for this particular value of θ.

- Probability of type I error $= \beta(\theta)$ if $\theta \in \Omega_0$.
- Probability of type II error $= 1 - \beta(\theta)$ if $\theta \in \Omega_0^c$.

An ideal test should have power function satisfying $\beta(\theta) = 0$ for all $\theta \in \Omega_0$, $\beta(\theta) = 1$ for all $\theta \in \Omega_0^c$, which is typically not possible in practice.

Sizes and Levels of Tests

Size α test

A test with power function $\beta(\theta)$ is a size α test if

$$\sup_{\theta \in \Omega_0} \beta(\theta) = \alpha$$

In other words, the maximum probability of making a type I error is α.

Level α test

A test with power function $\beta(\theta)$ is a level α test if

$$\sup_{\theta \in \Omega_0} \beta(\theta) \leq \alpha$$

In other words, the maximum probability of making a type I error is equal or less than α.

Any size α test is also a level α test

Likelihood Ratio Tests (LRT)

Definition

Let $L(\theta|\mathbf{x})$ be the likelihood function of θ. The likelihood ratio test statistic for testing $H_0 : \theta \in \Omega_0$ vs. $H_1 : \theta \in \Omega_0^c$ is

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Omega_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Omega} L(\theta|\mathbf{x})} = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})}$$

where $\hat{\theta}$ is the MLE of θ over $\theta \in \Omega$, and $\hat{\theta}_0$ is the MLE of θ over $\theta \in \Omega_0$ (restricted MLE).
The *likelihood ratio test* is a test that rejects H_0 if and only if $\lambda(\mathbf{x}) \leq c$ where $0 \leq c \leq 1$.

Example of LRT

Problem

Consider $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, \sigma^2)$ where σ^2 is known.

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_0$$

For the LRT test and its power function

Solution

$$L(\theta|\mathbf{x}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\theta)^2}{2\sigma^2}\right] = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i-\theta)^2}{2\sigma^2}\right]$$

We need to find MLE of θ over $\Omega = (-\infty, \infty)$ and $\Omega_0 = (-\infty, \theta_0]$.

MLE of θ over $\Omega = (-\infty, \infty)$

To maximize $L(\theta|\mathbf{x})$, we need to maximize $\exp\left[-\frac{\sum_{i=1}^n (x_i-\theta)^2}{2\sigma^2}\right]$, or equivalently to minimize $\sum_{i=1}^n (x_i-\theta)^2$.

$$\sum_{i=1}^n (x_i-\theta)^2 = \sum_{i=1}^n (x_i^2 + \theta^2 - 2\theta x_i) = n\theta^2 - 2\theta \sum_{i=1}^n x_i + \sum_{i=1}^n x_i^2$$

The equation above minimizes when $\theta = \hat{\theta} = \frac{\sum_{i=1}^n x_i}{n} = \overline{x}$.

MLE of θ over $\Omega_0 = (-\infty, \theta_0]$

- $L(\theta|\mathbf{x})$ is maximized at $\theta = \frac{\sum_{i=1}^n x_i}{n} = \overline{x}$ if $\overline{x} \leq \theta_0$.
- However, if $\overline{x} \geq \theta_0$, $\overline{x}$ does not fall into a valid range of $\hat{\theta}_0$, and $\theta \leq \theta_0$, the likelihood function will be an increasing function. Therefore $\hat{\theta}_0 = \theta_0$.

To summarize,

$$\hat{\theta}_0 = \begin{cases} \overline{X} & \text{if } \overline{X} \leq \theta_0 \\ \theta_0 & \text{if } \overline{X} > \theta_0 \end{cases}$$

Likelihood ratio test

$$\lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})} = \begin{cases} 1 & \text{if } \overline{X} \leq \theta_0 \\ \frac{\exp\left[-\frac{\sum_{i=1}^n (x_i-\theta_0)^2}{2\sigma^2}\right]}{\exp\left[-\frac{\sum_{i=1}^n (x_i-\overline{x})^2}{2\sigma^2}\right]} & \text{if } \overline{X} > \theta_0 \end{cases}$$

$$= \begin{cases} 1 & \text{if } \overline{X} \leq \theta_0 \\ \exp\left[-\frac{n(\overline{x}-\theta_0)^2}{2\sigma^2}\right] & \text{if } \overline{X} > \theta_0 \end{cases}$$

Therefore, the likelihood test rejects the null hypothesis if and only if

$$\exp\left[-\frac{n(\overline{x}-\theta_0)^2}{2\sigma^2}\right] \leq c$$

and $\overline{x} \geq \theta_0$.

Specifying c

$$\begin{aligned} \exp\left[-\frac{n(\overline{x}-\theta_0)^2}{2\sigma^2}\right] &\leq c \\ \iff -\frac{n(\overline{x}-\theta_0)^2}{2\sigma^2} &\leq \log c \\ \iff (\overline{x}-\theta_0)^2 &\geq -\frac{2\sigma^2 \log c}{n} \\ \iff \overline{x}-\theta_0 &\geq \sqrt{-\frac{2\sigma^2 \log c}{n}} \qquad (\because \overline{x} > \theta_0) \end{aligned}$$

Specifying c (cont'd)

So, LRT rejects H_0 if and only if

$$\begin{aligned} \overline{x}-\theta_0 &\geq \sqrt{-\frac{2\sigma^2 \log c}{n}} \\ \iff \frac{\overline{x}-\theta_0}{\sigma/\sqrt{n}} &\geq \frac{\sqrt{-\frac{2\sigma^2 \log c}{n}}}{\sigma/\sqrt{n}} = c^* \end{aligned}$$

Therefore, the rejection region is

$$\left\{\mathbf{x} : \frac{\overline{x}-\theta_0}{\sigma/\sqrt{n}} \geq c^*\right\}$$

Power function

$$\begin{aligned}
\beta(\theta) &= \Pr(\text{reject } H_0) = \Pr\left(\frac{\overline{X}-\theta_0}{\sigma/\sqrt{n}} \geq c^*\right) \\
&= \Pr\left(\frac{\overline{X}-\theta+\theta-\theta_0}{\sigma/\sqrt{n}} \geq c^*\right) \\
&= \Pr\left(\frac{\overline{X}-\theta}{\sigma/\sqrt{n}} \geq \frac{\theta_0-\theta}{\sigma/\sqrt{n}} + c^*\right)
\end{aligned}$$

Since $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, \sigma^2)$, $\overline{X} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$. Therefore,

$$\begin{aligned}
\frac{\overline{X}-\theta}{\sigma/\sqrt{n}} &\sim \mathcal{N}(0,1) \\
\Longrightarrow \beta(\theta) &= \Pr\left(Z \geq \frac{\theta_0-\theta}{\sigma/\sqrt{n}} + c^*\right)
\end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$.

Making size α LRT

To make a size α test,

$$\begin{aligned}
\sup_{\theta \in \Omega_0} \beta(\theta) &= \alpha \\
\sup_{\theta \leq \theta_0} \Pr\left(Z \geq \frac{\theta_0-\theta}{\sigma/\sqrt{n}} + c^*\right) &= \alpha \\
\Pr(Z \geq c^*) &= \alpha \\
c^* &= z_\alpha
\end{aligned}$$

Note that $\Pr\left(Z \geq \frac{\theta_0-\theta}{\sigma/\sqrt{n}} + c^*\right)$ is maximized when θ is maximum (i.e. $\theta = \theta_0$).
Therefore, size α LRT test rejects H_0 if and only if $\frac{\overline{x}-\theta_0}{\sigma/\sqrt{n}} \geq z_\alpha$.

Another Example of LRT

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} f(x|\theta) = e^{-(x-\theta)}$ where $x \geq \theta$ and $-\infty < \theta < \infty$. Find a LRT testing the following one-sided hypothesis.

$$\begin{aligned}
H_0 &: \theta \leq \theta_0 \\
H_1 &: \theta > \theta_0
\end{aligned}$$

Solution

$$\begin{aligned}
L(\theta|\mathbf{x}) &= \prod_{i=1}^{n} e^{-(x_i-\theta)} I(x_i \geq \theta) \\
&= e^{-\sum x_i + n\theta} I(\theta \leq x_{(1)})
\end{aligned}$$

The likelihood function is a increasing function of θ, bounded by $\theta \leq x_{(1)}$. Therefore, when $\theta \in \Omega = \mathbb{R}$, $L(\theta|\mathbf{x})$ is maximized when $\theta = \hat{\theta} = x_{(1)}$.

Solution (cont'd)

When $\theta \in \Omega_0^c$, the likelihood is still an increasing function, but bounded by $\theta \leq \min(x_{(1)}, \theta_0)$. Therefore, the likelihood is maximized when $\theta = \hat{\theta}_0 = \min(x_{(1)}, \theta_0)$. The likelihood ratio test statistic is

$$\begin{aligned}
\lambda(\mathbf{x}) &= \begin{cases} \frac{e^{-\sum x_i + n\theta_0}}{e^{-\sum x_i + n x_{(1)}}} & \text{if } \theta_0 < x_{(1)} \\ 1 & \text{if } \theta_0 \geq x_{(1)} \end{cases} \\
&= \begin{cases} e^{n(\theta_0 - x_{(1)})} & \text{if } \theta_0 < x_{(1)} \\ 1 & \text{if } \theta_0 \geq x_{(1)} \end{cases}
\end{aligned}$$

Solution (cont'd)

The LRT rejects H_0 if and only if

$$\begin{aligned} e^{n(\theta_0 - x_{(1)})} &\leq c \qquad (\text{and } \theta_0 < x_{(1)}) \\ \Longleftrightarrow \theta_0 - x_{(1)} &\leq \frac{\log c}{n} \\ \Longleftrightarrow x_{(1)} &\geq \theta_0 - \frac{\log c}{n} \end{aligned}$$

So, LRT reject H_0 is $x_{(1)} \geq \theta_0 - \frac{\log c}{n}$ and $x_{(1)} > \theta_0$. The power function is

$$\beta(\theta) = \Pr\left(X_{(1)} \leq \theta_0 - \frac{\log c}{n} \wedge X_{(1)} > \theta_0\right)$$

To find size α test, we need to find c satisfying the condition

$$\sup_{\theta \leq \theta_0} \beta(\theta) = \alpha$$

LRT based on sufficient statistics

Theorem 8.2.4

If $T(\mathbf{X})$ is a sufficient statistic for θ, $\lambda^*(t)$ is the LRT statistic based on T, and $\lambda(\mathbf{x})$ is the LRT statistic based on $\mathbf{x}$ then

$$\lambda^*[T(\mathbf{x})] = \lambda(\mathbf{x})$$

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

Proof

By Factorization Theorem, the joint pdf of $\mathbf{x}$ can be written as

$$f(\mathbf{x}|\theta) = g(T(\mathbf{x})|\theta)h(\mathbf{x})$$

and we can choose $g(t|\theta)$ to be the pdf or pmf of $T(\mathbf{x})$.
Then, the LRT statistic based on $T(\mathbf{X})$ is defined as

$$\lambda^*(t) = \frac{\sup_{\theta\in\Omega_0} L(\theta|T(\mathbf{x}) = t)}{\sup_{\theta\in\Omega} L(\theta|T(\mathbf{x}) = t)} = \frac{\sup_{\theta\in\Omega_0} g(t|\theta)}{\sup_{\theta\in\Omega} g(t|\theta)}$$

Proof (cont'd)

LRT statistic based on $\mathbf{X}$ is

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{\sup_{\theta\in\Omega_0} L(\theta|\mathbf{x})}{\sup_{\theta\in\Omega} L(\theta|\mathbf{x})} = \frac{\sup_{\theta\in\Omega_0} f(\mathbf{x}|\theta)}{\sup_{\theta\in\Omega} f(\mathbf{x}|\theta)} \\ &= \frac{\sup_{\theta\in\Omega_0} g(T(\mathbf{x})|\theta)h(\mathbf{x})}{\sup_{\theta\in\Omega} g(T(\mathbf{x})|\theta)h(\mathbf{x})} \\ &= \frac{\sup_{\theta\in\Omega_0} g(T(\mathbf{x})|\theta)}{\sup_{\theta\in\Omega} g(T(\mathbf{x})|\theta)} = \lambda^*(T(\mathbf{x})) \end{aligned}$$

The simplified expression of $\lambda(\mathbf{x})$ should depend on $\mathbf{x}$ only through $T(\mathbf{x})$, where $T(\mathbf{x})$ is a sufficient statistic for θ.

Example

Problem

Consider $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, \sigma^2)$ where σ^2 is known.

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

Find a size α LRT.

Solution - Using sufficient statistics

$T(\mathbf{X}) = \overline{X}$ is a sufficient statistic for θ.

$$T \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$$

$$\lambda(t) = \frac{\sup_{\theta \in \Omega_0} L(\theta|t)}{\sup_{\theta \in \Omega} L(\theta|t)} = \frac{\frac{1}{2\pi\sigma^2/n} \exp\left[-\frac{(t-\theta_0)^2}{2\sigma^2/n}\right]}{\sup_{\theta \in \Omega} \frac{1}{2\pi\sigma^2/n} \exp\left[-\frac{(t-\theta)^2}{2\sigma^2/n}\right]}$$

Solution (cont'd)

The numerator is fixed, and MLE in the denominator is $\hat{\theta} = t$. Therefore the LRT statistic is

$$\lambda(t) = \exp\left[-\frac{n(t-\theta_0)^2}{2\sigma^2}\right]$$

LRT rejects H_0 if and only if

$$\lambda(t) = \exp\left[-\frac{n(t-\theta_0)^2}{2\sigma^2}\right] \leq c$$

$$\Longrightarrow \left|\frac{t-\theta_0}{\sigma/\sqrt{n}}\right| \geq \sqrt{-2\log c} = c^*$$

Solution (cont'd)

Note that

$$T = \overline{X} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$$
$$\frac{T-\theta_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

A size α test satisfies

$$\sup_{\theta \in \Omega_0} \Pr\left(\left|\frac{T-\theta}{\sigma/\sqrt{n}}\right| \geq c^*\right) = \alpha$$
$$\Pr\left(\left|\frac{T-\theta_0}{\sigma/\sqrt{n}}\right| \geq c^*\right) = \alpha$$
$$\Pr\left(|Z| \geq c^*\right) = \alpha$$
$$\Pr(Z \geq c^*) + \Pr(Z \leq -c^*) = \alpha$$
$$|Z| = \left|\frac{T-\theta}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$$

LRT with nuisance parameters

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, \sigma^2)$ where both θ and σ^2 unknown. Between $H_0 : \theta \leq \theta_0$ and $H_1 : \theta > \theta_0$.

1. Specify Ω and Ω_0
2. Find size α LRT.

Solution - Ω and Ω_0

$$\Omega = \{(\theta, \sigma^2) : \theta \in \mathbb{R}, \sigma^2 > 0\}$$
$$\Omega_0 = \{(\theta, \sigma^2) : \theta \leq \theta_0, \sigma^2 > 0\}$$

Solution - Size α LRT

$$\lambda(\mathbf{x}) = \frac{\sup_{\{(\theta,\sigma^2):\theta\leq\theta_0,\sigma^2>0\}} L(\theta,\sigma^2|\mathbf{x})}{\sup_{\{(\theta,\sigma^2):\theta\in\mathbb{R},\sigma^2>0\}} L(\theta,\sigma^2|\mathbf{x})}$$

For the denominator, the MLE of θ and σ^2 are

$$\begin{cases} \hat{\theta} = \overline{X} \\ \sigma^2 = \frac{\sum(X_i-\overline{X})^2}{n} = \frac{n-1}{n}s^2_{\mathbf{X}} \end{cases}$$

For numerator, we need to maximize $L(\theta,\sigma^2|\mathbf{x})$ over the region $\theta \leq \theta_0$ and $\sigma^2 > 0$.

$$L(\theta,\sigma^2|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i-\theta)^2}{2\sigma^2}\right]$$

Solution - Maximizing Numerator

Step 1, fix σ^2, likelihood is maximized when $\sum_{i=1}^n (x_i-\theta)^2$ is minimized over $\theta \leq \theta_0$.

$$\hat{\theta}_0 = \begin{cases} \overline{x} & \text{if } \overline{x} \leq \theta_0 \\ \theta_0 & \text{if } \overline{x} > \theta_0 \end{cases}$$

Solution - Maximizing Numerator (cont'd)

Step 2 : Now, we need to maximize likelihood (or log-likelihood) with respect to σ^2 and we substitute $\hat{\theta}_0$ for θ.

$$l(\hat{\theta},\sigma^2|\mathbf{x}) = -\frac{n}{2}\left(\log 2\pi + \log\sigma^2\right) - \frac{\sum(x_i-\hat{\theta}_0)^2}{2\sigma^2}$$

$$\frac{\partial \log l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum(x_i-\hat{\theta}_0)^2}{2(\sigma^2)^2} = 0$$

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^n (x_i-\hat{\theta}_0)^2}{n}$$

Combining the results together

$$\lambda(\mathbf{x}) = \begin{cases} 1 & \text{if } \overline{x} \leq \theta_0 \\ \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} & \text{if } \overline{x} > \theta_0 \end{cases}$$

Solution - Constructing LRT

LRT test rejects H_0 if and only if $\overline{x} > \theta_0$ and

$$\left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} \leq c$$

$$\left(\frac{\sum(x_i-\overline{x})^2/n}{\sum(x_i-\theta_0)^2/n}\right)^{n/2} \leq c$$

$$\frac{\sum(x_i-\overline{x})^2}{\sum(x_i-\theta_0)^2} \leq c^*$$

$$\frac{\sum(x_i-\overline{X})^2}{\sum(x_i-\overline{X})^2 + n(\overline{x}-\theta_0)^2} \leq c^*$$

$$\frac{1}{1+\frac{n(\overline{x}-\theta_0)^2}{\sum(x_i-\overline{x})^2}} \leq c^*$$

Solution - Constructing LRT (cont'd)

$$\frac{n(\overline{x}-\theta_0)^2}{\sum(x_i-\overline{x})^2} \geq c^{**}$$
$$\frac{\overline{x}-\theta_0}{s_{\mathbf{x}}/\sqrt{n}} \geq c^{***}$$

LRT test reject if $\frac{\overline{x}-\theta_0}{s_{\mathbf{x}}/\sqrt{n}} \geq c^{***}$

The next step is specify c to get size α test (omitted).

Unbiased Test

Definition

If a test always satisfies

$$\Pr(\text{reject } H_0 \text{ when } H_0 \text{ is false }) \geq \Pr(\text{reject } H_0 \text{ when } H_0 \text{ is true })$$

Then the test is said to be unbiased

Alternative Definition

Recall that $\beta(\theta) = \Pr(\text{reject } H_0)$. A test is unbiased if

$$\beta(\theta') \geq \beta(\theta)$$

for every $\theta' \in \Omega_0^c$ and $\theta \in \Omega_0$.

Example

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, \sigma^2)$ where σ^2 is known, testing $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$.

LRT test rejects H_0 if $\frac{\overline{x}-\theta_0}{\sigma/\sqrt{n}} > c$.

$$\begin{aligned}
\beta(\theta) &= \Pr\left(\frac{\overline{X}-\theta_0}{\sigma/\sqrt{n}} > c\right) \\
&= \Pr\left(\frac{\overline{X}-\theta+\theta-\theta_0}{\sigma/\sqrt{n}} > c\right) \\
&= \Pr\left(\frac{\overline{X}-\theta}{\sigma/\sqrt{n}} + \frac{\theta-\theta_0}{\sigma/\sqrt{n}} > c\right) \\
&= \Pr\left(\frac{\overline{X}-\theta}{\sigma/\sqrt{n}} > c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

Note that $X_i \sim \mathcal{N}(\theta, \sigma^2)$, $\overline{X} \sim \mathcal{N}(\theta, \sigma^2/n)$, and $\frac{\overline{X}-\theta}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$.

Example (cont'd)

Therefore, for $Z \sim \mathcal{N}(0,1)$

$$\beta(\theta) = \Pr\left(Z > c + \frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right)$$

Because the power function is increasing function of θ,

$$\beta(\theta') \geq \beta(\theta)$$

always holds when $\theta \leq \theta_0 < \theta'$. Therefore the LRTs are unbiased.

Summary

Today

- Examples of LRT
- LRT based on sufficient statistics
- LRT with nuisance parameters
- Unbiased Test

Next Lecture

- Uniformly Most Powerful Test