Hypothesis Testing: Likelihood Ratio Tests (LRT)

Tiandong Wang

Department of Statistics
Texas A&M University

Likelihood Ratio Tests (LRT)

Let $L(\theta|\mathbf{x})$ be the likelihood function of θ. The likelihood ratio test statistic for testing $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_0^c$ is

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta\in\Theta_0} L(\theta|\mathbf{x})}{\sup_{\theta\in\Theta} L(\theta|\mathbf{x})} = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})},$$

where $\hat{\theta}_0$ is the MLE of θ in Θ_0 (restricted maximization); $\hat{\theta}$ is the MLE of θ in the full set Θ (unrestricted maximization);
A likelihood ratio test (LRT) is a test that has a rejection region

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

where c is any number satisfying $0 \leq c \leq 1$.

Comments on LRT

- The numerator in $\lambda(\mathbf{x})$ is the maximum probability of the observed sample $\mathbf{x}$ computed over parameters in H_0. The denominator is the maximum probability of the observed $\mathbf{x}$ over all possible parameters.
- $\hat{\theta}_0$ is the value in Θ_0 which makes the observation of data most likely; $\hat{\theta}$ is the value in Θ which makes the observation of data most likely
- If $\lambda(\mathbf{x})$ is small, it implies that there are some parameter points H_1 for which the observed sample is much more likely than for any parameter in H_0. So the LRT suggests we reject H_0 and accept H_1.
- The LRT statistic $\lambda(\mathbf{x})$ is a function of $\mathbf{x}$ not a function of θ
- $0 \leq \lambda(\mathbf{x}) \leq 1$

About the cut-off value c

- Different choices of $c \in [0, 1]$ give different tests and rejection regions.
- The smaller c, the smaller Type I error; The larger the c, the smaller Type II error.
- We will discuss the ideal choice of c later.

After finding an expression for $\lambda(\mathbf{x})$, we should get the simplest expression for R.

LRT: Example 1

(Normal One-sided LRT) $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 known. Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

- (i) Find the LRT and its power function.
- (ii) Comment on the decision rules given by different c's.

Let $X_1, \cdots, X_n$ be a random sample from a location-exponential family

$$f(x|\theta) = \exp^{-(x-\theta)}, \quad \text{if}, \quad x \geq \theta, -\infty < \theta < \infty$$

Test

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

Find the LRT and its power function.

LRT based on Sufficient Statistics

If T is sufficient for θ, then we can construct the LRT based on T and the likelihood function $L^*(\theta; t) = g(t; \theta)$. Since $T(\mathbf{x})$ contains all the information about θ in $\mathbf{x}$, the test based on T should be as good as the test based on $\mathbf{x}$. In fact, the tests are equivalent.

Theorem

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

$$\lambda^*(t) = \lambda(\boldsymbol{x})$$

for every x in the sample space.

Comment: The simplified expression for $\lambda(\mathbf{x})$ should depend on $\mathbf{x}$ only through $T(\mathbf{x})$ if $T(\mathbf{X})$ is a sufficient statistic for θ

LRT based on Sufficient Statistics: Examples

Examples:

- (Normal Two-sided LRT) $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 known. Consider testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0$$

  Find the LRT and its power function.

- (Normal with Nuisance Parameters) $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 unknown. Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

  (i) Specify Θ and Θ_0; (ii) Find the LRT and its power function.

Choose c for LRT

Choose c such that Type I error probability of LRT is bounded above by α

$$\sup_{\theta \in \Theta_0} P(\lambda(\mathbf{x}) \leq c) = \alpha$$

Example: n samples iid $N(\theta, \sigma^2)$, σ^2 known. Test $H_0 : \theta \leq \theta_0$ vs $\theta > \theta_0$

- (1) Find the size α LRT test.
- (2) Find size 0.05 test and 0.01 test.

Choose c for LRT: Examples

- $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 known. Consider testing
$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0$$
Find the size α of LRT test.
- $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 unknown. Consider testing
$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0$$
Show that the LRT test that rejects H_0 if
$$|\bar{\mathbf{X}} - \theta_0| > t_{n-1,\alpha/2}\sqrt{S^2/n}$$
is a test of size α
- iid location-exponential dist. Consider testing
$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$
Find the size α LRT test.

Sample Size Calculation

For fixed n, it is usually impossible to make both types of error probabilities arbitrarily small. But if we can increase n it is possible to achieve the desired power level.
Example: $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 known. Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

The LRT test rejects H_0 if $(\bar{X} - \theta_0)/(\sigma/\sqrt{n}) > C$ has the power function $\pi(\theta) = 1 - \Phi(C + \dfrac{\theta_0 - \theta}{\sigma/\sqrt{n}})$

- (1) The maximum Type I error is

$$\sup_{\theta \leq \theta_0} \pi(\theta) = \pi(\theta_0) = 1 - \Phi(C),$$

  no matter what n is. To make the test have size α, we choose $C = z_\alpha$.

Sample Size Calculation: Normal Example

Example: $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 known. Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

- (2) After C is chosen, it is possible to increase $\pi(\theta)$ for $\theta > \theta_0$ by increasing n. Thus we can minimize Type II error (Recall that Type I error is already under control). Draw the picture of $\pi(\theta)$ for small and large n. [Note: this is generally impossible when n is fixed].
- (3) Assume $C = z_\alpha$. What n should we choose such that the maximum Type II error is 0.2 if $\theta \geq \theta_0 + \sigma$?
- (4) Compute n if we choose $\alpha = 0.05$ in (3)

Unbiased Test

Idea:

- Sometimes we would like a test to be more likely to reject H_0 if $\theta \in \Theta_0^c$ than if $\theta \in \Theta_0$, i.e., $\mathbb{P}_\theta$(reject H_0 when H_0 is false) $\geq \mathbb{P}_\theta$(reject H_0 when H_0 is true).
- A test with such a property is *unbiased*.

Recall $\pi(\theta) = \mathbb{P}_\theta(\text{reject} \quad H_0)$.

Definition

A test with power function $\pi(\theta)$ is unbiased if

$$\pi(\theta') \geq \pi(\theta''), \text{for every} \quad \theta \in \Theta_0^c \quad \text{and} \quad \theta'' \in \Theta_0$$

In most problems, there are many unbiased tests.

Unbiased Test: Example

Let $X \sim Bin(5, \theta)$. Consider testing

$$H_0 : \theta \leq 1/2 \quad \text{versus} \quad H_1 : \theta > 1/2$$

with the procedure:

$$\text{reject} \quad H_0 \quad \text{if} \quad X = 5.$$

Show that the test is unbiased.

Unbiased Test: Example

Let $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 known. Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

- (1) Construct the LRT.
- (2) Graph the power function, and show the LRT is unbiased.
- (3) If we wish to have a maximum Type I error probability of 0.1 and to have a maximum Type II error probability of 0.2 if $\theta > \theta_0 + \sigma$, how to choose c and n?

- A good class of hypothesis tests are those with a small probability (say, less than α) of Type I error.
- A desired test in a good class would also have small Type II error, or, a large power function for $\theta \in \Theta_0^c$.

Uniformly Most Powerful (UMP) Tests

Definition

Let $\mathcal{C}$ be a class of tests for $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_0^c$
A test in class $\mathcal{C}$ with power function $\pi(\theta)$, is uniformly most powerful in class $\mathcal{C}$ (UMP) if

$$\pi(\theta) \geq \pi'(\theta), \forall \theta \in \Theta_0^c,$$

for every $\pi'(\theta)$ which is a power function of another test in $\mathcal{C}$.

If we consider $\mathcal{C} = \{\text{all the level } \alpha \text{ tests}\}$. The UMP test in this class is called a **UMP level α test**. It is the best test in the class $\mathcal{C}$, or the most powerful level α test.

Uniformly Most Powerful (UMP) Tests: Interpretation

The power function $\pi(\theta)$ of the UMP level α test satisfies:

$$\pi(\theta) \geq \pi'(\theta), \qquad \forall \theta \in \Theta_0^c,$$

where $\sup_{\theta \in \Theta_0} \pi(\theta) \leq \alpha$, $\sup_{\theta \in \Theta_0} \pi'(\theta) \leq \alpha$.

Uniformly Most Powerful (UMP): Test function

For each testing procedure, define a *test* function on the sample space

$$\phi(x) = \begin{cases} 1 & \text{if} \quad x \in R \\ 0 & \text{if} \quad x \notin R \end{cases}$$

Note $\phi(x) = \mathbf{1}_{\{x \in R\}}$. The expected value of ϕ is the power function

$$E_\theta[\phi(X)] = P_\theta(X \in R) = \pi(\theta)$$

When do UMP tests exist and how to find it?

For simple hypotheses, $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$ the UMP level α test always exists.

Important tool: **Neyman-Pearson Lemma**