Hypothesis Testing: UMP Tests

Tiandong Wang

Department of Statistics
Texas A&M University

Neyman-Pearson Lemma:

Consider testing $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$, where the pdf or pmf corresponding to θ_i is $f(x|\theta_i)$, $i = 0, 1$, using a test with rejection region R that satisfies

$$x \in R \quad \text{if} \quad f(x|\theta_1) > kf(x|\theta_0)$$

$$x \in R^c \quad \text{if} \quad f(x|\theta_1) < kf(x|\theta_0),$$

for some $k \geq 0$, and $\alpha = P_{\theta_0}(\mathbf{X} \in R)$, Then
a) (Sufficiency) Any test that satisfies (1) and (2) is a UMP level α test
b) (Necessary) If there is a test satisfying (1) and (2) with $k > 0$, then

- i) every UMP level α test is a size α test;
- ii) every UMP level α test satisfies (1) except on a set A satisfying $P_{\theta_0}(\mathbf{X} \in A) = P_{\theta_1}(\mathbf{X} \in A) = 0$

UMP: Binomial Example

Let $X \sim Bin(2, \theta)$. Consider testing
$H_0 : \theta = 1/2$ vs $H_1 : \theta = 3/4$

Sufficiency statistic and UMP test

Consider testing $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$. Suppose $T(\mathbf{X})$ is sufficient for θ and $g(t|\theta_i)$ is the pdf or pmf corresponding to θ_i is $f(x|\theta_i)$, $i = 0, 1,$. Then any test based on T with rejection region S is a UMP level α test if it satisfies

$$t \in S \quad \text{if} \quad g(t|\theta_1) > kg(t|\theta_0)$$

$$t \in S^c \quad \text{if} \quad g(t|\theta_1) < kg(t|\theta_0)$$

, for some $k \geq 0$, and $\alpha = P_{\theta_0}(T \in S)$,

Sufficiency statistic and UMP test: Examples

- Example: n samples iid $N(\theta, \sigma^2)$, σ^2 known. Test
  $H_0 : \theta = \theta_0$ vs $\theta = \theta_1$
  Find the UMP level α test.
- Example: n samples iid $N(\theta, \sigma^2)$, θ is known and σ^2 unknown. Test
  $H_0 : \sigma^2 = \sigma_0^2$ vs $\sigma^2 = \sigma_1^2$,
  where $\sigma_1^2 > \sigma_0^2$. Find the UMP level α test.

Monotone Likelihood Ratio (MLR)

Question: When does the UMP test exist for one-sided composite hypotheses?
Often when pdfs or pmfs have the monotone likelihood ratio property.
A family of pdfs or pmfs $\{g(t|\theta) : \theta \in \Theta\}$ for a univariate random variable T with real-valued parameter θ has a monotone likelihood ratio (MLR) if

$$g(t|\theta_2)/g(t|\theta_1) \quad \text{is an increasing function of } t$$

for every $\theta_2 > \theta_1$, on $\{t : g(t|\theta_1) > 0\}$ or $\{g(t|\theta_2) > 0\}$

Monotone Likelihood Ratio (MLR): Examples

- Normal, Poisson, Binomial all have the MLR property. (Exercise 8.25)
- If T is from an exponential family with the density

$$f(t|\theta) = h(t)c(\theta)\exp^{w(\theta)t},$$

  then the distribution of T has an MLR if $w(\theta)$ is a nondecreasing function in θ.
- If $X_1, \cdots, X_n$ iid from $N(\mu, \sigma^2)$ with σ^2 unknown, then $\sum_{i=1}^{n}(X_i - \mu)^2$ has an MLR

Note: Monotone decreasing is similarly defined.

Karlin-Rubin Theorem

Theorem

Suppose $T(\boldsymbol{X})$ is a sufficient statistic for θ and the family $\{g(t|\theta_i), \theta \in \Theta\}$ is an MLR family. Then:
(1) For testing
$H_0 : \theta \leq \theta_0$ *vs* $\theta > \theta_0$
the UMP level α test is given by rejects H_0 if and only if $T > t_0$ where $\alpha = P_{\theta_0}(T > t_0)$.

Karlin-Rubin Theorem: continue

Theorem

Suppose $T(\boldsymbol{X})$ is a sufficient statistic for θ and the family $\{g(t|\theta_i), \theta \in \Theta\}$ is an MLR family. Then:
(2) For testing
$H_0 : \theta \geq \theta_0$ *vs* $\theta < \theta_0$
the UMP level α test is given by rejects H_0 if and only if $T < t_0$ where $\alpha = P_{\theta_0}(T < t_0)$.

Karlin-Rubin Theorem: Examples

- Example 1: $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 known.
  - Find the UMP level α test for testing

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

  - Find the UMP level α test for testing

$$H_0 : \theta \geq \theta_0 \quad \text{versus} \quad H_1 : \theta < \theta_0$$

- Example 2: $X_1, \cdots, X_n \sim$ iid $N(\mu_0, \sigma^2)$ with μ_0 known and σ^2 unknown. Find the UMP level α test for testing

$$H_0 : \sigma^2 \leq \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 > \sigma_0^2$$

Nonexistence of UMP test

- For many problems, there is no UMP level α test, because the class level α test is so large that no one test dominates all the others in terms of power. Example 8.3.19 (textbook)
- Similar to UMVUE, we search a UMP test within some subset of the class of level α test, for example, the subset of all unbiased tests.

p-value

One method of reporting the hypotheses results is to report the size, α, of the test used and the decision to reject H_0 or accept H_0.

- If α is small, the decision to reject H_0 is fairly convincing
- If α is large, the decision to reject H_0 is not very convincing because the test has a large probability of incorrectly making that decision.

p-value

Two issues of this testing procedure:

- The choice of α is subjective. Different people may have differe tolerance levels α.
- The final answer does not not show the strength of decision (Is it a strong rejection or weak rejection? strong acceptance or weak acceptance?).

p-value

A p-value is the smallest possible level $\hat{\alpha}$ at which H_0 would be rejected.

- p-value is a test statistic, taking value $0 \leq p(x) \leq 1$ for the sample **x**.
- Small values of $p(\mathbf{X})$ gives evidence that H_1 is true.
- The smaller p-value, the stronger the evidence of rejecting H_0.
- A p-value is valid if, for every $\theta \in \Theta_0$ and every $0 \leq \alpha \leq 1$

$$P_\theta(p(\mathbf{X}) \leq \alpha) \leq \alpha$$

Compute p-value

Theorem

Let $W(\boldsymbol{X})$ be a test statistic such that large values of W give evidence that H_1 is true. For each sample point x, define

$$p(x) = \sup_{\theta \in \Theta_0} P_\theta(W(\boldsymbol{X}) \geq W(x))$$

Then $p(\boldsymbol{X})$ is a valid p-value.

-
- p-value testing procedure:
  - Compute p-value based on the data $x_1, \cdots, x_n$.
  - If p-value $< \alpha$, we reject H_0 at level α; otherwise accept H_0

P-value: Examples

- Example 1: $X_1, \cdots, X_n \sim$ iid $N(\theta, \sigma^2)$ with θ unknown and σ^2 unknown. Consider testing

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

  - Compute the p-value of the LRT statistic $W(X)$.
  - Assume $n = 16$ and we observed $\bar{x} = 1.5$, $s^2 = 1$. Assume $\theta_0 = 1$. Calculate the p-value. Do you reject the null hypothesis at level 0.05? at level 0.1?

- Example 2: In the above example, consider testing

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