Ch 8. Hypothesis Testing

Intro

"It is a mistake to confound strangeness with mystery." Sherlock Homes (A Study in Scarlet)

Definition

1. Any statement about the unknown parameter $\boldsymbol{\theta}$ is called a *hypothesis*
2. One of the complementary hypothesis is called *Null Hypothesis* (denoted by H_0) and other is called *Alternative Hypothesis* (denoted by H_1 or H_A).

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} N(\theta_1, \sigma^2)$ - regular diet program, $Y_1, \cdots, Y_n \overset{iid}{\sim} N(\theta_2, \sigma^2)$ - Caloric restricted diet program

$$H_0 : \theta_1 = \theta_2 \quad vs \quad H_1 : \theta_1 \geq \theta_2$$

▷ Note: Θ_0 and Θ_1 are often called *Null* and *Alternative* space of parameter and the hypotheses are expressed as

$$H_0 : \boldsymbol{\theta} \in \Theta_0 \quad vs \quad H_1 : \boldsymbol{\theta} \in \Theta_1$$

Definition

A hypothesis that completely specifies the distribution of $X_1, \cdots, X_n$ is called a *simple hypothesis* otherwise it is called *composite hypothesis*.

▷ Example: $\theta_1 = \theta_2$, $\theta_1 = \theta_2 = 2$, $\theta_1 > \theta_2$.

- After observing $X_1 = x_1, \cdots, X_n = x_n$, we need to decide which hypothesis, H_0 or H_1, we will accept. Let $\mathfrak{X}$ denote the set of all possible realization of $X_1, \cdots, X_n$. Testing function (rule) plays the same role as estimator in point estimation.

Definition

1. A function $\phi : \mathfrak{X} \to [0,1]$ is called a *testing function.*
2. If a testing function takes a values in $\{0,1\}$, i.e. $\phi : \mathfrak{X} \to \{0,1\}$, it is called a *simple testing function.*

$\triangleright$ Note: The interpretation of definition 1 is that after observing $X_1 = x_1, \cdots, X_n = x_n$, reject H_0 with probability $\phi(x_1, \cdots, x_n)$ and accept H_0 with probability $1 - \phi(x_1, \cdots, x_n)$. This is called a randomized procedure.

Definition

- $R_\phi = \{\mathbf{x} : \phi(\mathbf{x}) = 1\}$ is called the *rejection region* or *critical region*
- $A_\phi = \{\mathbf{x} : \phi(\mathbf{x}) = 0\}$ is called the *acceptance region*

Definition

Let $X_1, \cdots, X_n$ have joint pdf/pmf $f(\mathbf{x}|\theta)$, $\theta \in \Theta$. Let Θ_0 be a proper subset of Θ. Define the likelihood ratio

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} f(\mathbf{x}|\theta)}{\sup_{\theta \in \Theta} f(\mathbf{x}|\theta)}.$$

Then the Likelihood Ratio Test (LRT) of size α for testing $H_0 : \theta \in \Theta_0 \quad vs \quad H_1 : \theta \in \Theta_0^c$ is

$$\phi(\mathbf{x}) = \begin{cases} 1, & \lambda(\mathbf{x}) < k, \\ \gamma, & \lambda(\mathbf{x}) = k, \\ 0, & \lambda(\mathbf{x}) > k, \end{cases}$$

where k and γ satisfy $\sup_{\theta \in \Theta_0} E_\theta[\phi(\mathbf{x})] = \alpha$.

▷ Note:

1. Let $\hat{\theta}_0$ be the MLE of θ under H_0 and $\hat{\theta}$ be the MLE of θ without any restriction. Then,

$$\lambda(\mathbf{x}) = \frac{f(\mathbf{x}|\hat{\theta}_0)}{f(\mathbf{x}|\hat{\theta})}.$$

2. $0 \leq \lambda(\mathbf{x}) \leq 1$.

▷ Example 8.2.2: $X_1, \cdots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$, σ^2 is known.

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

▷ Example 8.2.3: $X_1, \cdots, X_n \overset{iid}{\sim} f(x|\theta) = e^{-(x-\theta)}, \quad x \geq \theta.$

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

▷ Example 8.2.6: $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, μ and σ^2 are unknown.

$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu \neq \mu_0$$

Theorem

$X_1, \cdots, X_n \overset{iid}{\sim} f(x|\theta)$. *$\lambda(\mathbf{x})$ is a likelihood ratio for testing*

$$H_0 : \theta \in \Theta_0 \quad vs \quad H_1 : \theta \in \Theta_0^c$$

Then under the regularity conditions (CRLB) on $f(x|\theta)$ and H_0

$$-2\ln[\lambda(\mathbf{x})] \xrightarrow{D} \chi^2_k,$$

where $k = \#$ of free parameters for $\boldsymbol{\theta} \in \Theta$ - $\#$ of free parameters for $\boldsymbol{\theta} \in \Theta_0$. This yields the approximate size α test

$$\phi(\mathbf{x}) = \begin{cases} 1, & -2\ln[\lambda(\mathbf{x})] > \chi^2_{1-\alpha,k}, \\ \gamma, & -2\ln[\lambda(\mathbf{x})] = \chi^2_{1-\alpha,k}, \\ 0, & -2\ln[\lambda(\mathbf{x})] < \chi^2_{1-\alpha,k}. \end{cases}$$

Q: How to compare several testing function ? or How to construct a good testing functions?

- Errors in Testing

		True status of Nature	
		H_0 is true	H_1 is true
Action	Accept H_0	O.K.	Type II error
	Reject H_0	Type I error	O.K.

– Type I error: Reject H_0 when H_0 is true
– Type II error: Accept H_0 when H_0 is false

Definition

The *power function* $\beta_\phi(\theta)$ of a test $\phi(\mathbf{x})$ is the function defined as

$$\beta_\phi(\theta) = P_\theta[\phi(\mathbf{X}) = 1] = E_\theta[\phi(\mathbf{X})] = P_\theta(\mathbf{X} \in R_\phi)$$

▷ Note:

- $\sup_{\theta \in \Theta_0} \beta_\phi(\theta)$ is called the *size of the test* ϕ. Thus, any test such that $\sup_{\theta \in \Theta_0} \beta_\phi(\theta) = \alpha$ is called as a *size* α *test*.
- Test ϕ such that $\sup_{\theta \in \Theta_0} \beta_\phi(\theta) \leq \alpha$ is called a *level* α *test*.
- $\theta \in \Theta_1$, $\beta_\phi(\theta) = 1 - Pr[Type \quad II \quad error]$.

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$. σ^2 is known.

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

Consider a test function of size $\alpha = 0.10$.

$$\phi(\mathbf{x}) = \begin{cases} 1 & \bar{x} > \theta_0 + c\sigma/\sqrt{n} \\ 0 & elsewhere. \end{cases}$$

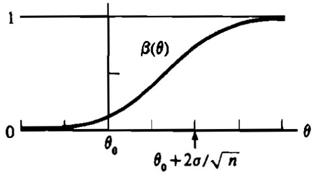
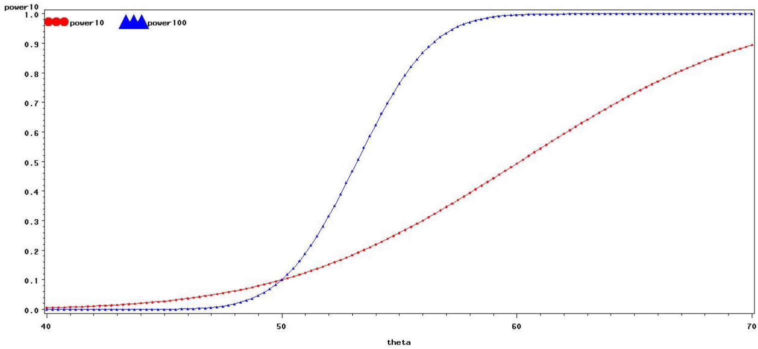


Figure 8.3.2. *Power function for Example 8.3.3*

Ch 8. Hypothesis Testing

Evaluating the test

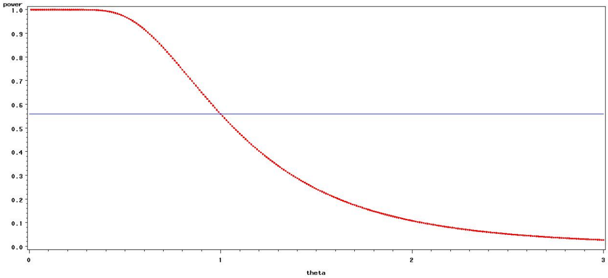


▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim}$ exponential(θ).

$$H_0 : \theta \geq 1 \quad vs \quad H_1 : \theta < 1$$

Consider a test function

$$\phi(\mathbf{x}) = \begin{cases} 1 & \bar{x} < 1 \\ 0 & elsewhere. \end{cases}$$



Definition

A test function $\phi[\mathbf{X} = (X_1, \cdots, X_n)]$ is said to be the *most powerful* test of size α for testing

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$$

if

1. $E_{\theta_0}[\phi(\mathbf{X})] = \alpha$, $[\beta_\phi(\theta_0) = \alpha.]$
2. for any other test function $\tilde{\phi}(\mathbf{X})$ with $E_{\theta_0}[\tilde{\phi}(\mathbf{X})] \leq \alpha$,

$$E_{\theta_1}[\phi(\mathbf{X})] \geq E_{\theta_1}[\tilde{\phi}(\mathbf{X})], \quad [\beta_\phi(\theta_1) \geq \beta_{\tilde{\phi}}(\theta_1)]$$

MP test has the smallest probability of type II error among all test rules with probability of type I error no bigger than α.

▷ Example: $H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$

$X = x$	0	1	2
$p(x\|\theta_0)$	0.05	0.05	0.90
$p(x\|\theta_1)$	0.90	0.08	0.02
$p(x\|\theta_1)/p(x\|\theta_0)$	18	1.6	0.022

Size $\alpha = 0.05$ tests?

Find the MP test of size 0.05? Choose the test that has the largest/smallest ratio?

Theorem (Neyman-Pearson Lemma)

$X_1, \cdots, X_n$ has a joint pdf/pmf $f(\mathbf{x}|\theta)$, $\theta \in \Theta$. Consider the testing the hypotheses,

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$$

Then, for any $0 \leq \alpha \leq 1$, there exist a MP test of size α given below;

$$\phi(\mathbf{x}) = \begin{cases} 1 & \text{if} \quad f(\mathbf{x}|\theta_1) > kf(\mathbf{x}|\theta_0), \\ \gamma & \text{if} \quad f(\mathbf{x}|\theta_1) = kf(\mathbf{x}|\theta_0), \\ 0 & \text{if} \quad f(\mathbf{x}|\theta_1) < kf(\mathbf{x}|\theta_0), \end{cases}$$

where the constants k and γ are chose to satisfy

$$E_{\theta_0}[\phi(\mathbf{X})] = \beta_\phi(\theta_0) = \alpha.$$

▷ Note:

1. The MP test ϕ reject H_0 if the likelihood ratio

$$L = \frac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)}$$

   is large.

2. In general, there may be more than one choice of k and γ that $\beta_\phi(\theta_0) = \alpha$. Then each is MP test of size α.
3. When $f(\mathbf{x}|\theta_1)/f(\mathbf{x}|\theta_0)$ has a continuous distribution under the null, H_0, $\gamma = 0$ is usually taken and considered as the MP test of size α.

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} \text{Gamma}(3, \theta)$.

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1 (> \theta_0)$$

Find the MP-test of size α.

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. ($\sigma^2$ known)

$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu = \mu_1 (> \mu_0)$$

Find the MP-test of size α.

Definition

Let $f(\mathbf{x}|\theta)$, $\theta \in \Theta$ be the joint pdf/pmf of $X_1, \cdots, X_n$. Let Θ_0 and Θ_1 be the nonempty disjoint subsets of Θ. A test rule $\phi(\mathbf{x})$ is said to be an *uniformly most powerful (UMP)* test of size α for testing

$$H_0 : \theta \in \Theta_0 \quad vs \quad H_1 : \theta \in \Theta_1$$

if

1. $\max_{\theta \in \Theta_0} E_\theta[\phi(\mathbf{X})] = \alpha$
2. for any other test $\tilde{\phi}(\mathbf{x})$ with $\max_{\theta \in \Theta_0} E_\theta[\tilde{\phi}(\mathbf{X})] \leq \alpha$, we have

$$E_\theta[\phi(\mathbf{X})] \geq E_\theta[\tilde{\phi}(\mathbf{X})]$$

for each $\theta \in \Theta_1$.

▷ Note:

1. A UMP test has the smallest probability of type II error for every $\theta \in \Theta_1$among all the test with size $\leq \alpha$.
2. Condition 2 is a really strong requirement. Unlike the simple versus simple case, UMP test may not exist for composite H_0 and for composite H_1.
3. NP lemma can be used to show that UMP test does not exist or identify the UMP test if it exists. HOW? (See next slide)

a. Fix $\theta_0 \in \Theta_0$ appropriately (usually boundary of Θ_0).
b. Choose any $\theta_1 \in \Theta_1$
c. Then find a MP test of size α, $\phi(\mathbf{x})$, for

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1.$$

If

i $\phi(\mathbf{x})$ does not depend on θ_1
ii $\max_{\theta \in \Theta_0} E_\theta[\phi(\mathbf{X})] = \alpha$

then $\phi(\mathbf{x})$ is the UMP-test of size α.

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu > \mu_0$$

$$H_0 : \mu \leq \mu_0 \quad vs \quad H_1 : \mu > \mu_0$$

Ch 8. Hypothesis Testing

Evaluating the test - UMP test

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} f(x|\lambda)$.

$$f(x|\lambda) = \lambda e^{-\lambda x}, \quad x > 0$$

$$H_0 : \lambda \leq \lambda_0 \quad vs \quad H_1 : \lambda > \lambda_0$$

Definition

Let $f(\mathbf{x}|\theta)$, $\theta \in \Theta$ be the joint pdf/pmf of $X_1, \cdots, X_n$. The family is said to have *Monotone Likelihood Ratio (MLR)* in a statistic $T(\mathbf{X})$ if, for all $\theta'' > \theta'$, $\theta'', \theta' \in \Theta$, there exist a nondecreasing function of T, g, such that

$$L = \frac{f(\mathbf{x}|\theta'')}{f(\mathbf{x}|\theta')} = g_{\theta',\ \theta''}[T(\mathbf{x})]$$

in a support of $\mathbf{x}$.

▷ Note:

- if $g_{\theta',\ \theta''}(x)$ is decreasing then $g_{\theta',\ \theta''}(-x)$ is increasing.
- if $f(\mathbf{x}|\theta'') > 0$ and $f(\mathbf{x}|\theta') = 0$ then $L = \infty$.

- Example: $X_1, \cdots X_n \stackrel{iid}{\sim} f(x|\theta)$

$$f(x|\theta) = c(\theta)h(x)\exp[w(\theta)t(x)]$$

Theorem

Let $X_1, \cdots, X_n$ have joint pdf/pmf $f(\mathbf{x}|\theta)$, $\theta \in \Theta$. Assume the family has MLR in $T(\mathbf{X})$. Then

1. *A UMP test of size α for*

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

is given by

$$\phi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) > k, \\ \gamma, & T(\mathbf{x}) = k, \\ 0, & T(\mathbf{x}) < k, \end{cases}$$

where k and γ are determined by

$$P_{\theta_0}[T(\mathbf{X}) > k] + \gamma P_{\theta_0}[T(\mathbf{X}) = k] = \alpha.$$

Theorem (-Continued)

2. *A UMP test of size α for*

$$H_0 : \theta \geq \theta_0 \quad vs \quad H_1 : \theta < \theta_0$$

is given by

$$\phi(\mathbf{x}) = \begin{cases} 1, & T(\mathbf{x}) < k, \\ \gamma, & T(\mathbf{x}) = k, \\ 0, & T(\mathbf{x}) > k, \end{cases}$$

where k and γ are determined by

$$P_{\theta_0}[T(\mathbf{X}) < k] + \gamma P_{\theta_0}[T(\mathbf{X}) = k] = \alpha.$$

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} \text{Unif}[0, \theta]$, $\theta > 0$.

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

Find a UMP test of size α.

▷ Example: $X_1, \cdots, X_n \overset{iid}{\sim} f(x|\eta)$

$$f(x|\eta) = e^{-(x-\eta)}, \quad x > \eta.$$

$$H_0 : \eta \leq \eta_0 \quad vs \quad H_1 : \eta > \eta_0$$

Find a UMP test of size α.