Lecture - Testing of hypothesis

Dr. Arabin Kumar Dey

Associate Professor
Department of Mathematics
Indian Institute of Technology Guwahati

March 30, 2024

Table of Contents

Table of contents

- Suppose $\{X_1, X_2, \ldots, X_n\}$ is a sample observed from F_θ where $\theta \in \Theta$. For example X_i's may be a sample from $N(\theta, 1)$, where $\theta \in R$ is not known. One may be interested in testing the validity of a statement like is the mean θ equal to zero?
- Or, in a coin tossing experiment one may be interested in testing whether the unknown probability of heads p, is equal to a specified value $p_0 \in (0, 1)$. Many other real-life situations of finance involve situations where a statement about the parameter is of interest and it is required to test if the statement is trite or false. Such a statement is known as a hypotheses.

Table of contents

Definition

- **A hypothesis is a statement about the underlying true parameter**.
- Suppose Θ is the underlying parameter space. We are given two statements: $\theta \in \Theta_0$ and $\theta \in \Theta_1 = \Theta - \Theta_0$. Based on the observed sample $X_1, \cdots, X_n$, one needs to decide on the true statement. We denote the first statement as

$$H_0 : \theta \in \Theta_0$$

.

This is known as the **null hypotheses**.

- The other statement is denoted as

$$H_1 : \theta \in \Theta_1$$

,

and is known as **alternative hypothesis**.

- If the parameter set Θ_0 (or Θ_1) consists of only one point, then H_0 (or H_1) is known as **simple hypotheses**.
- Otherwise, the hypothesis is known as a **composite hypothesis**.

Example

If $\sigma^2 = \sigma_0^2$, is known, then $H_0 : \mu = \mu_0$ is a simple null hypothesis, but $H_0 : \mu = \mu_0$ is composite null hypotheses, when σ^2 is unknown. Also, $H_1 : \mu \neq \mu_0$ is composite. But, $H_0 : \mu = \mu_0; \sigma^2 = \sigma_0^2$ will be a simple null hypothesis.

	Decision	
	Retain H_0	Reject H_0
H_0 true	✓	Type I error (false positive)
H_1 true	Type II error (false negative)	✓

- Choose a test statistic $W = W(X_1, \cdots, X_n)$
- Choose a rejection region R.
- If $W \in R$ we reject H_0 otherwise we retain H_0.

Example 1: $X_1, \cdots, X_n \sim Bernoulli(p)$.

$$H_0 : p = \frac{1}{2}, \quad H_1 : p \neq \frac{1}{2}.$$

Let $W = n^{-1} \sum_{i=1}^{n} X_i$. Let $R = \{x^n : |w(x^n) - 1/2| > \delta\}$. So we reject H_0 if $|W - 1/2| > \delta$

We will consider the following tests:

- Neyman-Pearson Test
- Wald test
- Likelihood Ration Test (LRT)
- the score test

Before we discuss these methods, we first need to talk about how we evaluate tests.

Evaluating Tests

Suppose we reject H_0 when $X^n = (X_1, cdots, X_n) \in R$. Define the power function by

$$\beta(\theta) = P_\theta(X^n \in R)$$

We want $\beta(\theta)$ to be small when $\theta \in \Theta_0$ we want $\beta(\theta)$ to be large when $\theta \in \Theta_1$.
The general strategy is:

- Fix $\alpha \in [0, 1]$.
- Now try to maximise $\beta(\theta)$ for $\theta \in \Theta_1$, subject to $\beta(\theta) \leq \alpha$ for $\theta \in \Theta_0$.

We need the following definitions. A test is **size** α if

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$$

A test is **level** α if

$$\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$$

A size α test and a level α test are almost same. **Sometimes we want a size α test and we cannot construct a test with exact size α but we can construct one with a smaller error rate.**

Example 2: $X_1, \cdots, X_n \sim N\left(\theta, \sigma^2\right)$ with σ^2 known. Suppose

$$H_0 : \theta = \theta_0, \quad H_1 : \theta > \theta_0.$$

This is called a **one-sided alternative**. Suppose, we reject H_0 if $W > c$ where

$$W = \frac{\overline{X_n} - \theta_0}{\sigma/\sqrt{n}}$$

.

Then,

$$\begin{aligned}
\beta(\theta) &= P_\theta\left(\frac{\overline{X_n} - \theta_0}{\sigma/\sqrt{n}} > c\right) \\
&= P_\theta\left(\frac{\overline{X_n} - \theta_0}{\sigma/\sqrt{n}} > c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= P\left(Z > c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= 1 - \underline{\overline{\Phi}}\left(c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

where $\underline{\overline{\Phi}}$ is the cdf of a standard Normal. Now

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \beta(\theta_0) = 1 - \underline{\overline{\Phi}}(c)$$

To get a size α test, Set $1 - \underline{\overline{\Phi}}(c) = \alpha$ so that

$$c = z_\alpha$$

where $z_\alpha = \underline{\overline{\Phi}}^{-1}(1-\alpha)$. Our test is : reject H_0 when

$$W = \frac{\overline{X_n} - \theta_0}{\sigma/\sqrt{n}} > z_\alpha$$

Example 3: $X_1, \cdots, X_n \sim N\left(\theta, \sigma^2\right)$ with σ^2 known. Suppose

$$H_0 : \theta = \theta_0, \quad H_1 : \theta \neq \theta_0.$$

This is called a two-sided alternative. we will reject H_0 if $|W| > c$ where W is defined as before.

$$
\begin{aligned}
\beta(\theta) &= P_\theta(W < -c) + P_\theta(W > c) \\
&= P_\theta\left(\frac{\overline{X_n} - \theta_0}{\sigma/\sqrt{n}} < -c\right) + P_\theta\left(\frac{\overline{X_n} - \theta_0}{\sigma/\sqrt{n}} > c\right) \\
&= P\left(Z < -c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) + P\left(Z > c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= \underline{\overline{\Phi}}\left(-c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) + 1 - \underline{\overline{\Phi}}\left(c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= \underline{\overline{\Phi}}\left(-c - \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) + \underline{\overline{\Phi}}\left(-c - \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right)
\end{aligned}
$$

since $\underline{\overline{\Phi}}(-x) = 1 - \underline{\overline{\Phi}}(x)$. The size is

$$\beta(\theta_0) = 2\underline{\overline{\Phi}}(-c)$$

To get a size α test we set $2\underline{\overline{\Phi}}(-c) = \alpha$ so that $c = -\underline{\overline{\Phi}}^{-1}(\alpha/2) = \underline{\overline{\Phi}}^{-1}(1-\alpha/2) = z_{\alpha/2}$.
The test is : reject H_0 when

$$|W| = \left|\frac{\overline{X_n} - \theta_0}{\sigma/\sqrt{n}}\right| > z_{\alpha/2}$$

–

Outline

1. Motivation

2. Basic Overview
   - Neyman Pearson Lemma

Summary of Neymann Pearson Lemma (Not exact Form)

A MP test of size $\alpha \in [0,1]$ for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ is given by

$$\phi(\underline{x}) = \begin{cases} 1 & \text{if } f_1(\underline{x}) > kf_0(\underline{x}) \\ \gamma & \text{if } f_1(\underline{x}) = kf_0(\underline{x}) \\ 0 & \text{if } f_1(\underline{x}) < kf_0(\underline{x}). \end{cases}$$

where $k \in [0, \infty)$ and we choose the values of k and γ from the size condition of the test i.e. $E_{\theta_0}(\phi(X)) = \alpha$.

In cases, where the X has a continuous distribution, the ratio $\frac{f_1((\underline{x})}{f_0((\underline{x})}$, may also have a continuous distribution (or in other words, it would be a continuous random variable). And in that case. $\underline{x} : f_1(\underline{x}) = kf_0(\underline{x})$ would have zero probability (with respect to the associated measure). As a result, the the above MP test would be of the form:

$$\phi(X) = \begin{cases} 1 & \text{if } f_1(\underline{x}) > kf_0(\underline{x}) \\ 0 & \text{if } f_1(\underline{x}) < kf_0(\underline{x}) \end{cases}$$

Follow the dedicated separate lecture note on this topic.

The Wald Test

Let

$$W = \frac{\hat{\theta}_n - \theta_0}{se}$$

where, se is the standard deviation of $\hat{\theta}_n$ calculated at θ_0. Under the usual conditions we have that under H_0, $W \sim N(0,1)$. Hence, an asymptotic level α test is to reject when $|W| > z_{\alpha/2}$. For example, with Bernoulli data, to test $H_0 : p = p_0$,

$$W = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

You can also use

$$W = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

In other words, to compute the standard error, you can replace θ with an estimate $\hat{\theta}$ or by null value θ_0.

The Likelihood Ratio Test (LRT)

This test is simple: reject H_0 if $\lambda(x^n) \leq c$ where

$$\lambda(x^n) = \frac{\sup_{\theta \in \Theta_0} L(\theta)}{\sup_{\theta \in \Theta_0 \cup \Theta_1} L(\theta)} = \frac{L(\hat{\theta}_0)}{L(\hat{\theta})}$$

where $\hat{\theta}_0$ maximises $L(\theta)$ subject to $\theta \in \Theta_0$.

Example

Let $X_1, \cdots, X_n \sim N(\theta, 1)$. Suppose

$$H_0 : \theta = \theta_0, \quad H_0 : \theta \neq \theta_0$$

after some algebra,

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

So,

$$R = \{x : \lambda \leq c\} = \{x : |\overline{X} - \theta_0| \geq c'\}$$

where, $c' = \sqrt{-2logc/n}$. Choosing c' to make this level α gives: reject if $|W| > z_{\alpha/2}$ where $W = \sqrt{n}(\overline{X} - \theta_0)$ which is the test we construct before.

Score Test

The score statistic is

$$S(\theta) = \frac{\partial}{\partial\theta} logf(X_1, \cdots, X_n; \theta) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} logf(X_i\theta).$$

Recall that $E_\theta S(\theta) = 0$ and $V_\theta S(\theta) = I_n(\theta)$. By the CLT,

$$Z = \frac{S(\theta_0)}{\sqrt{I_n(\theta_0)}} \sim N(0, 1)$$

under H_0. So, we reject if $|Z| > z_{\alpha/2}$. The advantage of the score test is that it does not require maximising the likelihood function.

Example

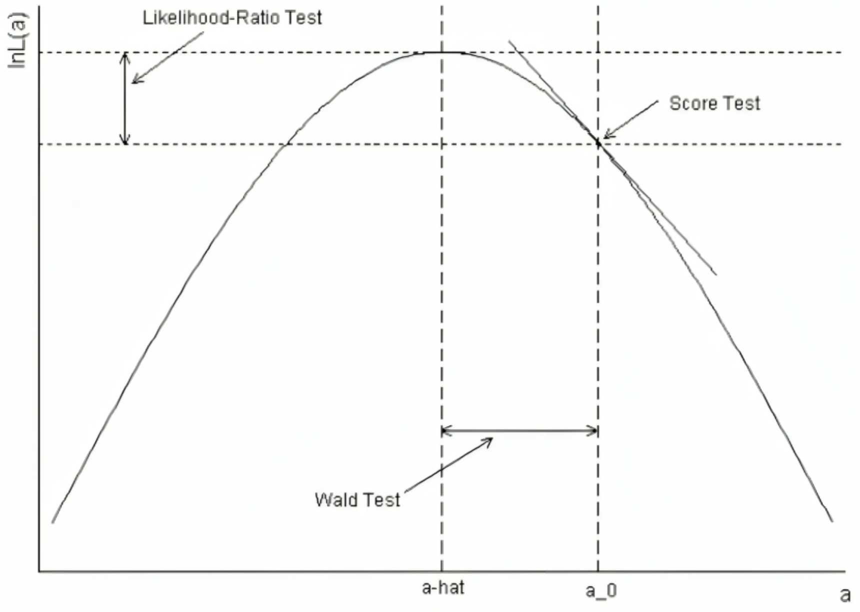
For the Binomial,

$$S(p) = \frac{n(\hat{p}_n - p}{p(1-p)}, \quad I_n(p) = \frac{n}{p(1-p)}$$

and so

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}.$$

This is the same as Wald test in this case.


lnL(a)
Likelihood-Ratio Test
Score Test
Wald Test
a-hat
a_0
a

p-value

When we test at a given level α we will reject or not reject. It is useful to summarise what levels we would reject at and what levels we would not reject at. **The p-values is the smallest α at which we would reject H_0.**
In other words, we reject at all $\alpha \geq p$. So, if the p-values is 0.03, then we would reject at $\alpha = 0.05$ but not at $\alpha = 0.01$. Hence, to test at level α when $p < \alpha$.

Theorem

Suppose we have a test of the form: Reject when $W(X^n) > c$. Then the p-values when $X^n = x^n$ is

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