

Module 3: Statistical inference (II)

Benjamin Smith

07/13/2026

Outline

This module we will review

- Basics of hypothesis testing
- Central Limit Theorem
- The Likelihood Ratio test

[Suppliment 3](#) discusses the Wald test, the Score test and their equivalence to the LR test.

Hypothesis testing

Definition (Hypothesis testing)

Suppose that we partition the parameter space Θ into two disjoint sets Θ_0 and Θ_1 and that we wish to test

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_1$$

We call H_0 the null hypothesis and H_1 the alternative hypothesis.

Rejection region

Let X be a random variable and let $\mathcal{X}$ be the range of X . Rejection region is a subset of outcomes $R \in \mathcal{X}$

$$X \in R \implies \text{reject } H_0$$

$$X \notin R \implies \text{retain (do not reject) } H_0$$

Usually, the rejection region is

$$R = \{x : T(x) > c\}$$

where T is a test statistic and c is a critical value.

Type I error and type II error

- Type I error, also known as a “false positive”: the error of rejecting a null hypothesis when it is actually true. $P(X \in R|H_0)$.
- Type II error, also known as a “false negative”: the error of not rejecting a null hypothesis when the alternative hypothesis is the true state of nature. $P(X \notin R|H_1)$.

Power function

Definition (Power function)

In a test of hypothesis about a parameter θ , let the null hypothesis be $H_0 : \theta = \theta_0$. The power function $\beta(\theta)$ is a function that gives, for any θ , the probability of rejecting the null hypothesis when the true parameter is equal to θ .

$$P(X \in R | \theta \text{ is the true parameter})$$

Note that the power function depends on the null hypothesis: if we change θ_0 , also the power function changes.

Size of a test

Definition (The size of a test)

The size of a test is defined to be

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

A test is said to have level α if its size is less than or equal to α .

Intuitively, we consider all the cases in which the null is true ($\theta \in \Theta_0$). For each case, we compute the probability of (incorrect) rejection. The size is equal to the largest value we find (worst-case scenario).

Exercise

Let $X_1, \dots, X_n \sim N(\mu, \sigma)$ where σ is known. We want to test $H_0 : \mu \leq 0$ versus $H_1 : \mu > 0$. Hence, $\Theta_0 = (-\infty, 0]$ and $\Theta_1 = (0, \infty)$.

Consider the test:

$$\text{reject } H_0 \text{ if } T > c$$

where $T = \bar{X}$. The rejection region is

$$R = \{(x_1, \dots, x_n) : T(x_1, \dots, x_n) > c\}$$

What is the power function? What is the size of the test?

Exercise (cont'd)

Let Z denote a standard Normal random variable. The power function is

$$\begin{aligned}\beta(\mu) &= \mathbb{P}_{\mu}(\bar{X} > c) \\ &= \mathbb{P}_{\mu}\left(\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} > \frac{\sqrt{n}(c - \mu)}{\sigma}\right) \\ &= \mathbb{P}\left(Z > \frac{\sqrt{n}(c - \mu)}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{\sqrt{n}(c - \mu)}{\sigma}\right)\end{aligned}$$

Exercise (cont'd)

$$\text{size} = \sup_{\mu \leq 0} \beta(\mu) = \beta(0) = 1 - \Phi\left(\frac{\sqrt{n}c}{\sigma}\right)$$

For a size α test, we set this equal to α and solve for c to get

$$c = \frac{\sigma\Phi^{-1}(1 - \alpha)}{\sqrt{n}}$$

We reject when $\bar{X} > \sigma\Phi^{-1}(1 - \alpha)/\sqrt{n}$.

Asymptotically normality

Definition

We say that an estimator $\hat{\theta}$ is asymptotically normal if:

$$\frac{\hat{\theta} - \theta}{\sqrt{\text{Var}(\hat{\theta})}} \xrightarrow{d} N(0, 1)$$

Theorem

If an estimator is asymptotically normal and the scaled squared standard error $\sqrt{n\widehat{\text{Var}}(\hat{\theta})} \xrightarrow{P} \sqrt{n\text{Var}(\hat{\theta})}$ then

$$\frac{\hat{\theta} - \theta}{\sqrt{\widehat{\text{Var}}(\hat{\theta})}} \xrightarrow{d} N(0, 1)$$

Central Limit Theorem

Let $X_1, \dots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables drawn from a distribution with mean μ and variance σ^2 . Then

$$\frac{\bar{X}_n - \mu}{\sqrt{\text{Var}(\bar{X}_n)}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0, 1) \text{ as } n \rightarrow \infty$$

Proof: Omitted. By characteristic functions.

Example

When $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$, then $\bar{X}_n$ satisfies that $\frac{\bar{X}_n - \mu}{\sqrt{\text{Var}(\bar{X}_n)}} \xrightarrow{d} N(0, 1)$ and $\sqrt{n \text{var}(\bar{X}_n)} = \sqrt{n \frac{s^2}{n}} = s \xrightarrow{P} \sigma = \sqrt{n \text{Var}(\bar{X}_n)}$. Then we can use the theorem above to conclude that

$$\frac{\bar{X}_n - \mu}{\sqrt{\widehat{\text{Var}}(\bar{X}_n)}} \xrightarrow{d} N(0, 1).$$

The likelihood ratio test

Let $\hat{\theta}_n$ be the MLE of θ .

$$\Delta_n = \ell(\hat{\theta}_n) - \ell(\theta_0) = \log \left(\frac{L(\hat{\theta}_n | \mathbf{x})}{L(\theta_0 | \mathbf{x})} \right) = \log \left(\frac{\sup_{\theta \in \Theta} L(\theta | \mathbf{x})}{L(\theta_0 | \mathbf{x})} \right) \geq 0$$

Under H_0 ,

$$2\Delta_n \xrightarrow{D} \chi_1^2$$

Uniformly most powerful test

Definition:

In statistical hypothesis testing, a uniformly most powerful (UMP) test is a hypothesis test which has the greatest power among all possible tests of a given size α .

Theorem (Neyman-Pearson lemma):

Let H_0 and H_1 be simple hypotheses. For a constant $c > 0$, suppose that the likelihood ratio test which rejects H_0 when $L(\mathbf{x}) < c$ has significance level α . Then for any other test of H_0 with significance level at most α , its power against H_1 is at most the power of this likelihood ratio test.

Resources

This tutorial is based on

- “All of statistics” Chapter 10 by Larry A. Wasserman.
- Arnaud Doucet’s STA 461 Lecture notes [[links](#)].