1 Hypothesis Testing

A *hypothesis* is a statement about the parameter space Θ. The *null hypothesis* Θ_0 is a subset of Θ of interest, typically suggested by some scientific theory. The *alternative hypothesis* $\Theta_1 = \Theta \backslash \Theta_0$ is the complement of Θ_0. *Hypothesis testing* is a decision whether to accept the null hypothesis or to reject it according to the observed evidence.

A *test function* is a mapping $\phi : \mathcal{X}^n \mapsto \{0,1\}$, where $\mathcal{X}$ is the sample space. We accept the null if $\phi(\mathbf{x}) = 0$, or reject it if $\phi(\mathbf{x}) = 1$. The *acceptance region* is defined as $A_\phi = \{\mathbf{x} \in \mathcal{X}^n : \phi(\mathbf{x}) = 0\}$, and the *rejection region* is $R_\phi = \{\mathbf{x} \in \mathcal{X}^n : \phi(\mathbf{x}) = 1\}$. The *power function* of the test ϕ is $\beta_\phi(\theta) = P_\theta(\phi(\mathbf{X}) = 1) = E_\theta(\phi(\mathbf{X}))$. The power function measures, at a given point, the probability that the test function rejects the null. The *size* of the test ϕ is a real number $\alpha = \sup_{\theta \in \Theta_0} \beta_\phi(\theta)$. The *level* of the test ϕ is a value $\alpha \in (0,1)$ such that $\alpha \geq \sup_{\theta \in \Theta_0} \beta_\phi(\theta)$, which is often used when it is difficult to get the exact supremum. The *probability of committing Type I error* is $\beta_\phi(\theta)$ for some $\theta \in \Theta_0$. The *probability of committing Type II error* is $1 - \beta_\phi(\theta)$ for $\theta \in \Theta_1$.

There has been a philosophical debate for decades about the hypothesis testing framework. At present the prevailing framework taught in statistics education is the frequentist perspective. A frequentist views the parameter as a fixed constant, and they are conservative about the Type I error. Only if overwhelming evidence is demonstrated should a researcher reject the null.

The definition of the test function is too general to be useful. We narrow it down to a set of meaningful test function. Under the notion of protecting the null hypothesis, a desirable test should have a small level. Conventionally we take $\alpha = 0.01$, 0.05 or 0.1.

We define $\Psi_\alpha = \left\{\phi : \sup_{\theta \in \Theta_0} \beta_\phi(\theta) \leq \alpha\right\}$ as the class of test functions of level smaller than α. A *uniformly most powerful test* $\phi^* \in \Psi_\alpha$ is a test function such that, for every $\phi \in \Psi_\alpha$

$$\beta_{\phi^*}(\theta) \geq \beta_\phi(\theta)$$

uniformly over $\theta \in \Theta_1$.

Example 1. Suppose a random sample of size 6 is generated from $(X_1, \ldots, X_6) \sim N(\theta, 1)$, where θ is unknown. We want to infer the population mean of the normal distribution. The null hypothesis is H_0: $\theta \leq 0$ and the alternative is H_1: $\theta > 0$. The test function $\phi(\mathbf{X}) = 1\left(\bar{X} \geq 1.64/\sqrt{6}\right)$ is the most powerful test among all tests of level 0.05. The power function of ϕ^* is $\beta_{\phi^*}(\theta) = \Phi\left(\sqrt{6}\theta - 1.64\right)$.

2 Confidence Interval

An *interval estimate* is a function $C : \mathcal{X}^n \mapsto \{\Theta' : \Theta' \subseteq \Theta\}$ that maps a point in the sample space to a subset of the parameter space. The *coverage probability* of an *interval estimator* $C(\mathbf{X})$ is defined as $P_\theta(\theta \in C(\mathbf{X}))$. The coverage probability is frequency that the interval estimator captures the true parameter that generates the sample.

Exercise 1. Suppose a random sample of size 6 is generated from $(X_1, \ldots, X_6) \sim N(\theta, 1)$. Find the coverage probability of the random interval $\left[\bar{X} - 1.96/\sqrt{6}, \bar{X} + 1.96/\sqrt{6}\right]$.

Hypothesis testing and confidence interval are closely related. Sometime it is difficult to directly construct the confidence interval, but easier to test a hypothesis. One way to construct confidence interval is by inverting a corresponding hypothesis testing problem.

Suppose $A_{\phi}(\theta)$ the acceptance region of a test ϕ whose size is α and the null is θ. If $C(\mathbf{x})$ is constructed as

$$C(\mathbf{x}) = \{\theta \in \Theta : \mathbf{x} \in A_{\phi}(\theta)\},$$

the coverage probability $P_{\theta}(\theta \in C(\mathbf{X})) = 1 - \alpha$.