Probability & Statistics

November 21, 2020

Testing Hypotheses

Core concepts covered in this chapter:

- Null and alternative hypotheses
- Critical region
- Test statistics
- Rejection region
- Power function
- Type I and II errors
- Significance level

Problems of Testing Hypotheses

Hypothesis testing is about deciding whether the parameter θ lies in a particular subset Ω_0 (null hypothesis $H_0 : \theta \in \Omega_0$) of the parameter space Ω or in it's complement Ω_1 (alternative hypothesis $H_1 : \theta \in \Omega_1$).

The Null and Alternative Hypotheses

We are interested in knowing whether the parameter θ lies in a subset Ω_0 of Ω or in its complement Ω_1. We consider the null hypothesis $H_0 : \theta \in \Omega_0$ and the alternative hypothesis $H_1 : \theta \in \Omega_1$. Testing is a procedure for deciding which hypothesis to choose, sometimes based on the observed data. When performing a test, H_0 is said to be rejected or not rejected. The parameter space depends on the statistical modeling assuptions we make about the observed data.

Simple And Composite Hypotheses

Simple hypothesis Ω_i contains a single value of θ, and therefore the null hypothesis has the form $H_0 : \theta = \theta_0$.

Composite hypothesis Ω_i contains more than one value of θ (one-sided $H_0 : \theta \leq \theta_0$ or $H_0 : \theta \geq \theta_0$, two-sided $H_0 : \theta \neq \theta_0$).

The Critical Region and Test Statistics

Overview We want to infer some unknown parameter θ from a population. We consider a sample $X = (X_1, \ldots, X_n)$ obtained from a distribution that involves θ. We test the hypotheses $H_0 : \theta \in \Omega_0$ and $H_1 : \theta \in \Omega_1$. We denote S the set of all the possible values of the random sample X. S is partitioned in S_0 containing the values of X for which we do not reject the null hypothesis, and its complement called *critical region* $S_1 = S_0^c$ containing the values of X for which we reject it. We - almost - always express the *critical region* in terms of a *test statistic* $T = f(X)$: we denote R a subset of the real line such that we reject H_0 if $T \in R \iff X \in S_1$. We call R the *rejection region* of the test.

An example We are interested in determining the mean of a population. We consider a sample $X = (X_1, \dots, X_n)$ and its empirical mean $\overline{X_n}$. The hypotheses are $H_0 : \mu = \mu_0$ and $H_1 : \mu \neq \mu_0$. Intuitively, we reject the null hypothesis if $\overline{X_n}$ is far from μ_0. We actually choose a number c such that we reject H_0 if the distance from $\overline{X}$ to μ_0 is greater than c. Here, $S_0 = \{x = (x_1, \dots, x_n), |\overline{X} - \mu_0| \leq c\}$ and the *critical region* of the test is $S_1 = S_0^C = \{x, |\overline{X}_n - \mu_0| \geq c\}$. Put differently, we pick a *test statistics* $T = |\overline{X} - \mu_0|$ and reject H_0 if $T \geq c$. The interval $[c; +\infty[$ is then the *rejection rejection* of the test.

The Power Function and Types of Error

The whole point of this is to define the notion of *significance level* of a test.