Hypothesis testing

A hypothesis is an assertion about the unknown distribution of the population.

In the parametric case, we assume that the distribution of the sample is $f(Y|\theta)$, where θ is an unknown parameter. Let Θ be the possible values of the parameter θ.

Given a hypothesis, there exists a collection of value of θ satisfying that hypothesis. So, to each hypothesis, we associate a subset of Θ

Elements of a statistical test

We have two complementary statistical hypothesis:

H_0: null hypothesis : $\theta \in \Theta_0$

H_a: alternative or research hypothesis, $\theta \in \Theta - \Theta_0 = \Theta_1$

A test is a rule based on the data to get to one of the following decisions

(1) Reject H_0, The data supports the H_a. H_0 is false

(2) Do no reject H_0. The data does not give enough evidence to reject H_0. H_0 might be true

For some values $(y_1 -- y_n)$, we make the decision (1)

For some values $(y_1 -- y_n)$, we make the decision (2)

The rejection region is

$C = \{(y_1 -- y_n) :$ the test rejects H_0, if $(y_1 -- y_n)$ is observed$\}$

Given a hypotesting testing problem, there are many possible tests for that problem.

A type I error is made if H_0 is reject when H_0 is true
The probability of a type I error is denoted by α. The value of α is called the level of the test.

A type II is made if H_0 is accepted when H_a is true.
The probability of a type II error is denoted by β

true hypothesis \ decision	H_0	H_1
H_0	correct decision	type I error α
H_1	type II error β	correct decision

We want to make statistical decisions so that errors are minimized.
Usually, we do a test so that the significance less is $\leq \alpha$.
($\alpha = 0.05, 0.01, 0.10$)
We take the null hypothesis, unless the data strongly suggest that the null hypothesis is true.
Example

Attained significance level or p-value

The p-value or attained significance level is the smallest level of significance α for which the observed data indicate that H_0 should be rejected.

If the null hypothesis is rejected for extreme value of a statistic, then the p-value is the probability that we obtain a value of the statistic as extreme or more extreme than the value of the statistic from the data.

Common large sample tests

Let $\hat{\theta}$ be an estimator of θ.

Let $\sigma_{\hat{\theta}}$ be the standard deviation of $\hat{\theta}$

(1) To test $H_0: \theta = \theta_0$ (or $\theta \le \theta_0$) versus $H_a: \theta > \theta_0$

we reject H_0 if $\hat{\theta} \geq \theta_0 + z_\alpha \sigma_{\hat{\theta}}$

The p-value is $P\left(N(0,1) \geq \frac{\sqrt{n}}{\sigma_{\hat{\theta}}}(\underset{\uparrow \text{data}}{\hat{\theta}} - \theta_0)\right)$

(2) To test $H_0: \theta = \theta_0$ (or $\theta \geq \theta_0$) versus $H_a: \theta < \theta_0$

we reject H_0 if $\hat{\theta} \le \theta_0 + z_\alpha \sigma_{\hat{\theta}}$

The p-value is $P\left(N(0,1) \le \frac{\sqrt{n}}{\sigma_{\hat{\theta}}}(\hat{\theta} - \theta_0)\right)$

(3) To test $H_0: \theta = \theta_0$ versus $H_a: \theta \neq \theta_0$

we reject H_0 if $|\hat{\theta} - \theta_0| \geq z_\alpha \sigma_{\hat{\theta}}$

The p-value is $P\left(|N(0,1)| \geq \frac{\sqrt{n}}{\sigma_{\hat{\theta}}}|\hat{\theta} - \theta_0|\right)$