Hypothesis Testing

A **Hypothesis Test** is a statistical inference method used to test the significance of a proposed (hypothesized) relation between population statistics (parameters) and their corresponding sample estimators. In other words, hypothesis tests are used to determine if there is enough evidence in a sample to prove a hypothesis true for the entire population.

The test considers two hypotheses: the **Null Hypothesis**, which is a statement meant that's being tested, usually something like "there is no affect" with the intention of proving this false; and the **Alternate Hypothesis**, which is the statement meant to stand after the test is performed. The two hypotheses must be mutually exclusive, moreover, in most applications the two are complementary (one, the negation of the other). The test works by comparing the p-value to the level of significance (a chosen target). If the p-value is less than or equal to the level of significance, then the null hypothesis is rejected.

When analyzing data, only samples of a certain size might be manageable as efficient computations. In some situations the error terms follow a continuous or infinite distribution, hence the use of samples to suggest accuracy of the chosen test statistics. The method of hypothesis testing gives an advantage over guessing what distribution, or which parameters the data follows.

Contents

Definitions and Methodology

In statistical inference, properties (parameters) of a population are analyzed by sampling data sets. Given assumptions on the distribution, i.e. a statistical model of the data, certain hypotheses can be deduced from the known behavior of the model. These hypotheses must be tested against sampled data from the population.

DEFINITION

The **null hypothesis** (denoted H_0), is a statement that is assumed to be true. If the null hypotheses is rejected, then there is enough evidence (statistical significance) to accept the **alternate hypothesis** (denoted H_1). Before doing any test for significance, both hypothesis must be clearly stated and non-conflictive, i.e. mutually exclusive statements.

Rejecting the null hypothesis, given that it is true, is called a **Type I error** and it is denoted α, which is also its probability of occurrence. Failing to reject the null hypothesis, given that it is false, is called a **Type II error** and it is denoted β, which is also its probability of occurrence. Also α is known as the **significance level**, and $1-\beta$ is known as the **power** of the test.

	H_0 **is true**	H_0 **is false**
Reject H_0	Type I error	Correct Decision
Reject H_1	Correct Decision	Type II error

The **test statistic** is the standardized value following the sampled data under the assumption that the null hypothesis is true, and a chosen particular test. These tests depend on the statistic to be studied and the assumed distribution it follows, e.g. the population mean following a normal distribution. The **P-value** is the probability of observing an extreme test statistic in the direction of the alternate hypothesis, given that the null hypothesis is true. The **critical value** is the value of the assumed distribution of the test statistic such that the probability of making a Type I error is small.

Methodologies

Given an estimator $\hat{\theta}$ of a population statistic θ, following a probability distribution $P(T)$, computed from a sample $\mathcal{S}$. Given a significance level α and test statistic t^*.

- Define H_0 and H_1.
- Compute the test statistic t^*.

P-value Approach (most prevalent)

- Find the P-value using t^* (right-tailed).
- If the P-value is at most α reject H_0. Otherwise, reject H_1.

Critical Value Approach

- Find the critical value solving the equation $P(T \geq t_\alpha) = \alpha$ (right-tailed).
- If $t^* > t_\alpha$, reject H_0. Otherwise, reject H_1.

Note: Failing to reject H_0 only means inability to accept H_1, does not mean to accept H_0.

Examples

EXAMPLE

Assume a normally distributed population has recorded cholesterol levels with various statistics computed. From a sample of 100 subjects in the population the sample mean was 214.12 mg/dL (milligrams per deciliter), with a sample standard deviation of 45.71 mg/dL.

Perform a hypothesis test, with significance level 0.05, to test if there is enough evidence to conclude that the population mean is larger than 200 mg/dL.

Hypothesis Test

We will perform a hypothesis test using the P-value approach with a significance level $\alpha = 0.05$.

- Define H_0: $\mu = 200$.
- Define H_1: $\mu > 200$.

- Since our values are normally distributed, the test statistic is $z^* = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{214.12-200}{\frac{45.71}{\sqrt{100}}} \approx 3.09$.
- Using a standard normal distribution we find that our P-value is approximately 0.001.
- Since the P-value is at most $\alpha = 0.05$. We reject H_0.

Therefore, we can conclude that the test shows sufficient evidence to support the claim that μ is larger than 200 mg/dL.

If the sample size was smaller the normal and t-distributions behave differently. Also the question itself must be managed by a double tail test instead.

EXAMPLE

Assume a population's cholesterol levels are recorded and various statistics are computed. From a sample of 25 subjects the sample mean was 214.12 mg/dL (milligrams per deciliter), with a sample standard deviation of 45.71 mg/dL.

Perform a hypothesis test, with significance level 0.05, to test if there is enough evidence to conclude that the population mean is not equal to 200 mg/dL.

Hypothesis Test

We will perform a hypothesis test using the P-value approach with a significance level $\alpha = 0.05$ and the t-distribution with 24 degrees of freedom.

- Define H_0: $\mu = 200$.
- Define H_1: $\mu \neq 200$.
- Using the t-distribution, the test statistic is $t^* = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{214.12-200}{\frac{45.71}{\sqrt{25}}} \approx 1.54$.
- Using a t-distribution with 24 degrees of freedom, we find that our P-value is approximately $2(0.068) = 0.136$. We have multiplied by two since this is a two tailed argument, i.e. the mean can be smaller than, or larger than.
- Since the P-value is larger than $\alpha = 0.05$. We fail to reject H_0.

Therefore, the test does not show sufficient evidence to support the claim that μ is not equal to 200 mg/dL.

Hypothesis Test and Confidence Intervals

The complement of the rejection on a two tailed hypotheses test (with significance level α) for a population parameter θ, is equivalent to finding a confidence interval (with confidence level $1-\alpha$) for the population parameter θ. If the assumption on the parameter θ, falls inside the confidence interval then the test has failed to reject the null-hypothesis (with P-value more than α). Otherwise, if θ does not fall in the confidence interval then the null-hypothesis is rejected in favor of the alternate (with P-value at most α).

See Also

- statistics (estimation)
- normal distribution
- correlation
- confidence intervals

Cite as: Hypothesis Testing. *Brilliant.org*. Retrieved 04:17, July 26, 2018, from https://brilliant.org/wiki/hypothesis-testing/