9.10 TESTING OF STATISTICAL HYPOTHESES. STATISTICAL CONCLUSION ERRORS

Definition: Null and alternative hypothesis

- The *null* hypothesis is a statement to be tested in an experiment. It is labelled H_0.
- The *alternative hypothesis* is a statement that is contradictory to H_0. It is labelled H_a.

There are two option for a decision: reject H_0 or do not reject H_0.

Definition: Type I and Type II errors

- A ***Type I*** error occurs when we reject H_0 but in fact H_0 is true (incorrect decision).
- A ***Type II*** error occurs when we do not reject H_0 but in fact H_0 is false.

The probability of **Type I** error is denoted by α**.** The probability of **Type II** error is denoted by β**.**

Remark:
Type I error and **Type II** are of course **incorrect** decisions.

The remaining two decision are **correct**:

- Not to reject H_0 when, in fact, H_0 is true (correct decision).
- To reject H_0 when, in fact, H_0 is false (correct decision).

We will introduce two tests in which we want to verify the hypothesis that a population mean is $\boldsymbol{\mu_0}$. In first (Model I) we assume that standard deviation of the population $\boldsymbol{\sigma}$ is known.

In the second test (Model II) we do not know the value of $\boldsymbol{\sigma}$. In both cases we assume that a population follows a normal distribution $\boldsymbol{N(\mu,\sigma)}$.

Model I of the test is the following.

Suppose that a population follows a normal distribution $\boldsymbol{N\,(\mu,\sigma)}$ and that the standard deviation of the population $\boldsymbol{\sigma}$ is known. Null hypothesis is $\boldsymbol{H_0\colon \mu = \mu_0}$, where $\boldsymbol{\mu_0}$ is a predicted mean value for entire population. From a random sample $\boldsymbol{x}$ we want to check hypothesis $\boldsymbol{H_0\colon \mu = \mu_0}$ against the alternative hypothesis $\boldsymbol{H_1\colon \mu \neq \mu_0}$. We assume some preconceived or pre-set a "significance level" $\boldsymbol{\alpha}$ of the test. A pre-set $\boldsymbol{\alpha}$ is the probability that we reject $\boldsymbol{H_0}$ but in fact $\boldsymbol{H_0}$ is true i.e., the probability of a Type I error. If $\boldsymbol{\alpha}$ is not given, the accepted standard is to set $\boldsymbol{\alpha = 0.05}$.

We verify the null hypothesis H_0 as follows:

First, we compute a mean $\bar{x}$ of the sample x. Next, we compute the value of random variable u with standard normal distribution $N(0,1)$ using the formula:

$$u = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}.$$

From the standard normal distribution $N(0,1)$ table or using inverse normal distribution calculator (see for instance:

http://onlinestatbook.com/2/calculators/inverse_normal_dist.html),

or using some popular spreadsheet like Excel, OpenOffice Calc,
we find such critical value u_α that

$$P(|u| \geq u_\alpha) = \alpha.$$

The set of values of u defined by inequality $|u| \geq u_\alpha$ is a critical area of this test i.e if we get such value u that $|u| \geq u_\alpha$ the hypothesis H_0 must be rejected. If $|u| < u_\alpha$, we cannot reject hypothesis H_0.

The test described above is sometimes called a two-sided test since we have a two-sided critical area i.e.

$|u| \geq u_\alpha$ means the same that $u \in (-\infty, -u_\alpha] \cup [u_\alpha, \infty)$.

This occurs whenever we have the alternative hypothesis of the form $H_a: \mu \neq \mu_0$.

In some cases, however, we must perform one-sided tests.

When the alternative hypothesis is of the form $H_a: \mu < \mu_0$ we have a left-sided critical area i.e. $u \leq u_\alpha$. In this case we derive u_α such that

$$P(u \leq u_\alpha) = \alpha.$$

When the alternative hypothesis is of the form $H_a: \mu > \mu_0$ we have a right-sided critical area i.e. $u \geq u_\alpha$. In this case we derive u_α such that

$$P(u \geq u_\alpha) = \alpha.$$

Remark

We also perform Model I if a population follows a normal distribution $N(\mu, \sigma)$ and standard deviation is unknown, but the sample is large i.e., it contains several dozen elements. In this case we set $s = \sigma$.

Example 9.34

In 81 randomly picked manufacturing companies, material cost used in the production of one particular product was measured. The mean of these values is $\bar{x} = 540$ € and $s = 150$ €. On a significance level $\alpha = 0.05$, verify the hypothesis, that the average material cost of producing that product is 600 €.

Co-funded by the Erasmus+ Programme of the European Union

Solution:

The null hypothesis H_0 is that the average material cost of producing the product is 600 € i.e.,

$$H_0: \mu_0 = 600.$$

Since the sample is large, we can assume that $s = \sigma = 150.$ We apply Model I.

Compute

$$u = \frac{540 - 600}{150}\sqrt{81} = -3.6.$$

We have

$$P(|u| \geq u_\alpha) = 0.05$$

for $u_\alpha = 1.96$ (see table, spreadsheet or some inverse normal distribution calculator i.e., http://onlinestatbook.com/2/calculators/inverse_normal_dist.html (option: outside since the test is two-tailed).

We see that

$$|-3.6| = 3.6 \geq 1.96 = u_\alpha.$$

Hence on the significance level $\alpha = 0.05$ the hypothesis H_0 must be rejected in favour of the alternative hypothesis that the average material cost of producing considered product is not equal to 600 €.

Model II of the test is the following.

Suppose that a population follows a normal distribution $\boldsymbol{N(\mu, \sigma)}$ and that the mean μ and the standard deviation of the population σ is unknown. A small sample was selected at random from the population. Null hypothesis is $\boldsymbol{H_0: \mu = \mu_0}$. From a random sample $\boldsymbol{x}$ we want to check $\boldsymbol{H_0}$ against the alternative hypothesis $\boldsymbol{H_1: \mu \neq \mu_0}$. We assume some a "significance level" $\boldsymbol{\alpha}$ of the test.

We verify the null hypothesis H_0 as follows:

First, we compute a sample mean $\bar{x}$ and a sample standard deviation s

or estimator of the standard deviation $\hat{s}$. Next, we compute the value of a random variable t using the formula:

$$t = \frac{\bar{x} - \mu_0}{s}\sqrt{n-1} = \frac{\bar{x} - \mu_0}{\hat{s}}\sqrt{n}.$$

Co-funded by the Erasmus+ Programme of the European Union

Under condition that H_0 holds random variable t follows student's $t-$ distribution with $n-1$ degrees of freedom.

From the $t-$ distribution table or using inverse normal distribution calculator (see for instance: https://www.statology.org/inverse-t-distribution-calculator) or using some popular spreadsheet like Excel, OpenOffice Calc,

we find such critical value t_α that

$$P(|t| \geq t_\alpha) = \alpha.$$

The set of values of t defined by inequality $|t| \geq t_\alpha$ is a critical area of this test i.e., if we get from the formula such value t that $|t| \geq t_\alpha$ the hypothesis H_0 must be rejected in favour of H_1. If $|t| < t_\alpha$ we cannot reject the hypothesis H_0.

The test described above is of course two-sided (see Model I) since we have a two-sided critical area i.e.

$$|t| \geq t_\alpha$$

This occurs whenever we have the alternative hypothesis of the form $H_1: \mu \neq \mu_0$.

In some cases, however, we have to perform one-sided tests.

When the alternative hypothesis is of the form H_1: $\mu < \mu_0$ we have a left-sided critical area i.e. $t \leq t_\alpha$. In this case we derive t_α such that

$$P(t \leq t_\alpha) = \alpha.$$

When the alternative hypothesis is of the form $H_1: \mu > \mu_0$ we have a right-sided critical area i.e., $t \geq t_\alpha$. In this case we derive t_α such that

$$P(t \geq t_\alpha) = \alpha.$$

Example 9.35

A machine produces metal plates of certain dimensions, with their nominal thickness being $0.04\ mm$. After randomly selecting and measuring 25 plates, their average thickness was

$\bar{x} = 0.037\ mm$ and $\hat{s} = 0.005\ mm$. Is it possible to conclude that produced plates are thinner than $0.04\ mm$? Assume the significance level $\alpha = 0.01$.

Solution:

The null hypothesis H_0 is that, the average thickness of metal plates is not less than 0.04 mm,

$$H_0: \mu_0 \geq 0.04 \text{ and } H_1: \mu_0 < 0.04.$$

Co-funded by the Erasmus+ Programme of the European Union

Compute

$$t = \frac{0.037 - 0.04}{0.005}\sqrt{25} = -3$$

We use $t -$ distribution with 24 degree of freedom and test is left-sided.

From the $t -$ distribution table or using inverse normal distribution calculator (see for instance: https://www.statology.org/inverse-t-distribution-calculator) or using some popular spreadsheet like Excel, OpenOffice Calc,

we get $t_\alpha = -2.492.$

We have $t = -3 < -2.492 = t_\alpha.$

Hence t belong to a critical area. We must reject H_0 in favour of H_1. Produced plates are thinner than $0.04\ mm.$

Co-funded by the
Erasmus+ Programme
of the European Union