Statistics

Lecture 9

Hypothesis testing: Parametric tests

SILESIAN UNIVERSITY
SCHOOL OF BUSINESS ADMINISTRATION IN KARVINA

David Bartl
Statistics
INM/BASTA

Outline of the lecture

- *z*-tests for the means
  - type I and type II error
  - significance level α
  - power of the test $1 - \beta$
- *t*-tests for the means
- χ^2-test for the variance
- *z*-test for the population proportion
- *p*-value of the test

Theorems & Confidence intervals

We presented the theorems last time:

- If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$
- If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$
- If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

We used the theorems to establish the confidence intervals.

Theorems & Confidence intervals

Theorem:

- If $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$

It also holds by the Lindenberg-Lévy Central Limit Theorem:

- If $X_1, X_2, \dots, X_n$ are independent and identically distributed and n is large, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$

Confidence interval:

$$\mu \in \left[\bar{x} - \frac{z(\alpha/2)\sigma}{\sqrt{n}}, \bar{x} + \frac{z(\alpha/2)\sigma}{\sqrt{n}}\right]$$

Theorems & Confidence intervals

Theorem:

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$

Confidence interval:

$$\mu \in \left[\bar{x} - \frac{t_{n-1}(\alpha/2)s}{\sqrt{n}}, \bar{x} + \frac{t_{n-1}(\alpha/2)s}{\sqrt{n}}\right]$$

where $t_{n-1}(\alpha/2) = F^{-1}\left(1 - \frac{\alpha}{2}\right)$

where F is the cumulative distribution function of

Theorems & Confidence intervals

Theorem:

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

Confidence interval:

$$\sigma^2 \in \left[\frac{(n-1)s^2}{\chi^2_{n-1}(\alpha/2)}, \frac{(n-1)s^2}{\chi^2_{n-1}(1-\alpha/2)}\right]$$

where $\chi^2_{n-1}(\alpha/2) = F^{-1}\left(1-\frac{\alpha}{2}\right)$ and $\chi^2_{n-1}(1-\alpha/2) = F^{-1}\left(\frac{\alpha}{2}\right)$

where F is the cumulative distribution function of

Theorems & Confidence intervals

Now, we use the same theorems

- If $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$
- If $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$
- If $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$
- …

to establish some parametric statistical tests about the parameters (μ, σ^2, …) of the respective probability distribution.

z-tests for the means

- One-sample *z*-test for the population mean
- Paired-sample *z*-test for the difference of the population means
- Two-sample *z*-test for the difference of the population means

One-sample z-test for the population mean

Recall the theorem:

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$

Example or motivation:

Assume that $X \sim \mathcal{N}(\mu, \sigma^2)$ is a random variable following the normal probability distribution. Assume that we do know the variance σ^2, but we do not know the true value of the population mean $\mu \in \mathbb{R}$.

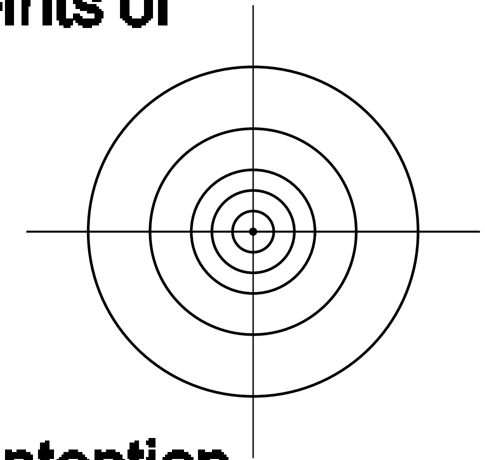
We conjecture / We assume / We … / that the population mean $\mu = \mu_0$, i.e. the (unknown) population mean μ is equal to some prescribed value $\mu_0 \in \mathbb{R}$.

One-sample z-test for the population mean

Example: An archer shoots an arrow against the plane.

The sample space $\Omega = \mathbb{R}^2 = \{[x, y] : x, y \in \mathbb{R}\}$ is the set of all the points of the plane. The random variable X is the x-coordinate of the hit, i.e.

$X(\omega) = X([x, y]) = x$.



We know the archer's variance σ^2, but we do not know the archer's intention,

i.e. we do not know the point which the archer intends to hit,

i.e. we do not know the archer's mean μ.

We conjecture that the archer's intention is to hit the origin, i.e. $\mu = 0$.

One-sample z-test for the population mean

Let $x_1 = X(\omega_1),\ x_2 = X(\omega_2),\ \ldots,\ x_n = X(\omega_n)$ be the numerical results of n trials of a random experiment, where $X \sim \mathcal{N}(\mu, \sigma^2)$, such as the x-coordinates of the archer's n hits.

We know the variance σ^2, but we do not know the mean μ.

We state the **null hypothesis** (about the mean):

$$H_0: \quad \mu = \mu_0$$

where $\mu_0 \in \mathbb{R}$ is some number such that we conjecture that the true mean could equal the μ_0.

One-sample *z*-test for the population mean

The **meaning of the null hypothesis** (such as $H_0\colon \mu = \mu_0$ in our example) is that

- the observed distinct values are caused by the randomness only

  (according to the assumed distribution, such as $X \sim \mathcal{N}(\mu, \sigma^2)$ in our example)
- there are no other factors causing the distinct values
- everything is all right, no need to reconfigure anything
- all factors under the consideration are equivalent (have the same effect)

One-sample *z*-test for the population mean

Having stated the **null hypothesis**

$$H_0: \quad \mu = \mu_0$$

we also state the **alternative hypothesis** (denoted by H_1 or H_A).

There are three options how to state the alternative hypothesis:

- two-sided: $H_1: \quad \mu \neq \mu_0$
- one-sided: $H_1: \quad \mu < \mu_0$
- one-sided: $H_1: \quad \mu > \mu_0$

One-sample *z*-test for the population mean

Which alternative hypothesis ($\mu \neq \mu_0$ or $\mu < \mu_0$ or $\mu > \mu_0$) do we choose?

→ That depends upon our knowledge of the situation.

In our example:

- If we suspect that the archer's intention is to hit a point different from the given point (such as the origin), we choose $H_1\colon \mu \neq \mu_0$.
- If we conjecture that the archer's intention is to hit a point to the left of the given point (such as the origin), we choose $H_1\colon \mu < \mu_0$.
- If we conjecture that the archer's intention is to hit a point to the right of the given point (such as the origin), we choose $H_1\colon \mu > \mu_0$.

One-sample *z*-test for the population mean

Under our assumptions ($x_1, \dots, x_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent and $\mu = \mu_0$), it follows by the Theorem that

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1) \qquad \textit{exactly}$$

Also, if $x_1, \dots, x_n$ are independent and identically distributed and n is large and $\mu = \mu_0$, then, by the Lindenberg-Lévy Central Limit Theorem,

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1) \qquad \textit{approximately}$$

One-sample *z*-test for the population mean

Thus, having the n measurements $x_1, x_2, \ldots, x_n$, we calculate the statistic

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

where

- $\bar{x} = (\sum_{i=1}^{n} x_i)/n$ is the sample mean
- n is the sample size
- μ_0 is the conjectured or estimated population mean
- $\sigma = \sqrt{\sigma^2}$ is the known standard deviation

We know (or assume) that $Z \sim \mathcal{N}(0, 1)$.

One-sample *z*-test for the population mean

We have $Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \sim \mathcal{N}(0, 1)$.

Then, if $-\infty \leq a < b \leq +\infty$, the probability that $a < Z < b$ is

$$P(a < Z < b) = \int_a^b \varphi(x)\,\mathrm{d}x$$

where

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}$$

is the density of the normalized normal distribution.

One-sample *z*-test for the population mean

Consider the first case $(H_1\colon \mu \neq \mu_0)$ first. We have:

$$H_0\colon \quad \mu = \mu_0$$

$$H_1\colon \quad \mu \neq \mu_0$$

Knowing that $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$, the probability

$$P(-c < Z < +c)$$

is quite high,

so having $-c < Z < +c$ accords with H_0, if $c > 0$ is large enough.

One-sample *z*-test for the population mean

On the other hand, if H_0 is true, then it is quite improbable that $Z \notin (-c, +c)$.

Therefore, **if we observe that**

$$Z \leq -c \qquad \text{or} \qquad +c \leq Z$$

then we may conclude that H_0 **is probably not true,**

i.e. **we reject the null hypothesis** H_0.

Therefore, the statistical test proceeds as follows:

(see below)

One-sample *z*-test for the population mean

Statistical one-sample z-test with two-sided alternative hypothesis ($\mu \neq \mu_0$):

- choose **the level of significance**, a small number $\alpha > 0$, a very popular value is $\alpha = 5\ \%$, other popular values are $10\ \%$ or $1\ \%$ or $0.1\ \%$ etc.
- find the **critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} \varphi(x)\,\mathrm{d}x = \frac{1}{2}\alpha \qquad \text{and} \qquad \int_{+c}^{+\infty} \varphi(x)\,\mathrm{d}x = \frac{1}{2}\alpha$$

for $\alpha = 5\ \%$, the (two-sided) critical value is $c = 1.959963\ldots$

- if $Z \in (-\infty, -c] \cup [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-c, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Type I and Type II error

There are exactly four possibilities when testing the null hypothesis H_0:

- the null hypothesis (H_0) is actually true & we do not reject it — OK
- the null hypothesis (H_0) is actually true & we reject it — type I error
- the null hypothesis (H_0) is actually not true & we do not reject it — type II error
- the null hypothesis (H_0) is actually not true & we reject it — OK

The purpose is that the probability of the type I error and that of the type II error is as small as possible.

Type I and Type II error

What is the probability of type I error

(the null hypothesis (H_0) is actually true & we reject it)?

The probability is equal to the significance level α, usually $\alpha = 5\,\%$.

Recall: The null hypothesis H_0 is rejected if and only if

$Z \in (-\infty, -c] \cup [+c, +\infty)$, i.e. if and only if $|Z| \geq c$.

The critical value c is such that – if H_0 holds true – then $P(|Z| \geq c) = \alpha$,

i.e. the probability of the type I error (rejecting H_0 when it is true) is α.

Type I and Type II error

What is the probability of type II error

(the null hypothesis (H_0) is actually false & we fail to reject it)?

The probability of type II error is denoted by β.

The power of the test is the probability $1-\beta$

It is much more difficult to calculate the probability β of type II error.

It must be calculated for each test separately.

Type I and Type II error

To calculate the probability β of type II error, consider that the null hypothesis H_0 is not true $(\mu \neq \mu_0)$ and we fail to reject it $\left(|Z| = \left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| < c\right)$.

By the Theorem then, we have $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{x-\mu_0+(\mu_0-\mu)}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$.

Then the probability of the type II error (H_0 not true & fail to reject it) is:

$$P\left(-c < \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < +c\right) = P\left(-c < \frac{\bar{x}-\mu+(\mu_0-\mu)}{\sigma/\sqrt{n}} < +c\right) =$$

$$= P\left(\frac{\mu-\mu_0}{\sigma/\sqrt{n}} - c < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \frac{\mu-\mu_0}{\sigma/\sqrt{n}} + c\right) =$$

$$= \beta = \int_{(\mu-\mu_0)/(\sigma/\sqrt{n})-c}^{(\mu-\mu_0)/(\sigma/\sqrt{n})+c} \varphi(x)\,\mathrm{d}x$$

Type I and Type II error

Notice that, if the true μ is close to the hypothesized μ_0 $(\mu \approx \mu_0)$, then $\frac{\mu-\mu_0}{\sigma/\sqrt{n}} \approx 0$, hence

$$\beta = \int_{(\mu-\mu_0)/(\sigma/\sqrt{n})-c}^{(\mu-\mu_0)/(\sigma/\sqrt{n})+c} \varphi(x)\,\mathrm{d}x \approx \int_{-c}^{+c} \varphi(x)\,\mathrm{d}x = 1-\alpha = 95\ \%$$

if $\alpha = 5\ \%$, say.

It is recommended that β should be ≤ 20 %.

Therefore, if we wish to have $\beta \approx 20\ \%$ or $\beta \leq 20\ \%$,

then we must not consider the true μ close to the hypothesized μ_0.

One-sample *z*-test for the population mean

Consider now the second case (H_1: $\mu < \mu_0$). We have:

$$H_0: \quad \mu = \mu_0$$

$$H_1: \quad \mu < \mu_0$$

Knowing that $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$ and assuming that $c > 0$ is large enough,

what is the probability that $-c < Z$?

If H_0 is true, then it is quite improbable that $Z \notin (-c, +\infty)$.

Therefore, if we observe that $Z \leq -c$, then we may conclude that

H_0 is probably not true, i.e. **we reject the null hypothesis H_0.**

One-sample z-test for the population mean

Statistical one-sample z-test with one-sided alternative hypothesis ($\mu < \mu_0$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$, other popular values are $\alpha = 10\ \%$ or $\alpha = 1\ \%$ or $\alpha = 0.1\ \%$ etc.
- find **the critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} \varphi(x)\,\mathrm{d}x = \alpha$$

  for $\alpha = 5\ \%$, the (one-sided) critical value is $-c = -1.64485\ldots$

- if $Z \in (-\infty, -c]$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-c, +\infty)$, then **do not reject** (or fail to reject) the null hypothesis

One-sample *z*-test for the population mean

Consider finally the third case $(H_1\!: \mu > \mu_0)$. We have:

$$H_0\!: \quad \mu = \mu_0$$

$$H_1\!: \quad \mu > \mu_0$$

Knowing that $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$ and assuming that $c > 0$ is large enough,

what is the probability that $Z < +c$?

If H_0 is true, then it is quite improbable that $Z \notin (-\infty, +c)$.

Therefore, if we observe that $+c \leq Z$, then we may conclude that

H_0 is <u>probably not true</u>, i.e. **we <u>reject the null hypothesis</u> H_0.**

One-sample *z*-test for the population mean

Statistical one-sample z-test with one-sided alternative hypothesis ($\mu > \mu_0$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$, other popular values are $\alpha = 10\ \%$ or $\alpha = 1\ \%$ or $\alpha = 0.1\ \%$ etc.
- find the **critical value** $c > 0$ so that

$$\int_{+c}^{+\infty} \varphi(x)\,\mathrm{d}x = \alpha$$

  for $\alpha = 5\ \%$, the (one-sided) critical value is $+c = +1.64485\ldots$

- if $Z \in [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-\infty, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Paired-sample z-test for the difference of the pop.means

Example or motivation:

Let us have a sample of n objects, e.g. n patients.

We do two measurements with each of the objects (patients)

— before some treatment

— after the treatment

The purpose it to learn whether the treatment has any effect.

(Hence the null hypothesis: "The treatment has no effect.")

Let $x_1, x_2, \dots, x_n$ be the values measured before the treatment, and

let $y_1, y_2, \dots, y_n$ be the values measured after the treatment.

Paired-sample *z*-test for the difference of the pop.means

That is, the measurement x_i and y_i is done with the i-th object (patient) before and after the treatment for $i = 1, 2, \dots, n$.

We assume that $X \sim \mathcal{N}(\mu^{\text{before}}, \sigma_X^2)$, i.e. the random variable of the measurement before the treatment follows the normal distribution, and that $Y \sim \mathcal{N}(\mu^{\text{after}}, \sigma_Y^2)$, i.e. the random variable of the measurement after the treatment also follows the normal distribution. We do not know the true values of the population means μ^{before} and μ^{after}, <u>but we do assume that we know the variances</u> σ_X^2 and σ_Y^2.

Paired-sample *z*-test for the difference of the pop.means

We formulate the **null hypothesis**:

the treatment has no effect, i.e. the population means are the same

$$H_0: \quad \mu^{\text{before}} = \mu^{\text{after}}$$

Recall that we do not know the true population means μ^{before} and μ^{after}. We only test the hypothesis by having done a sample of n pairs of measurements.

Formulate the alternative hypothesis:

- two-sided: $H_1: \quad \mu^{\text{before}} \neq \mu^{\text{afer}}$ (the treatment has some effect)
- one-sided: $H_1: \quad \mu^{\text{before}} < \mu^{\text{after}}$ (the treatment increases / …
- one-sided: $H_1: \quad \mu^{\text{before}} > \mu^{\text{after}}$ … / decreases the quantity)

Paired-sample z-test for the difference of the pop.means

Recall the theorem:

If $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$

Notice also:

If $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu^{\text{before}}, \sigma_X^2)$ and $Y_1, Y_2, \dots, Y_n \sim \mathcal{N}(\mu^{\text{after}}, \sigma_Y^2)$ are independent,

then the differences

$$X_1 - Y_1,\ X_2 - Y_2,\ \dots,\ X_n - Y_n \sim \mathcal{N}(\mu^{\text{before}} - \mu^{\text{after}},\ \sigma_X^2 + \sigma_Y^2)$$

Now, the hypothesis $\mu^{\text{before}} = \mu^{\text{after}}$ is equivalent to that

the mean of the difference $X - Y$ is $\mu = \mu_0 = 0.$

Paired-sample z-test for the difference of the pop.means

We have thus

reduced

the paired-sample z-test for the difference of the population means

to

the one-sample z-test for the population mean,

which we already know.

Paired-sample z-test for the difference of the pop.means

Having the n pairs of the measurements $x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n$, calculate the statistic

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2 + \sigma_Y^2}/\sqrt{n}} \qquad \text{or} \qquad Z = \frac{\bar{x} - \bar{y} - \mu_0}{\sqrt{\sigma_X^2 + \sigma_Y^2}/\sqrt{n}}$$

where

- $\bar{x} = (\sum_{i=1}^{n} x_i)/n$ is the sample mean of the measurements before the treatment
- $\bar{y} = (\sum_{i=1}^{n} y_i)/n$ is the sample mean of the measurements after the treatment
- n is the number of the pairs
- $\mu_0 = 0$ for no difference of the means $(\mu^{\text{before}} = \mu^{\text{after}})$
- $\mu_0 = \text{const.}$ for a general difference of the means $(\mu^{\text{before}} = \mu^{\text{after}} + \text{const.})$

Paired-sample *z*-test for the difference of the pop.means

In the first case $(H_1\colon\ \mu^{\text{before}} \neq \mu^{\text{after}})$, we have:

$$H_0\colon \quad \mu^{\text{before}} = \mu^{\text{after}}$$

$$H_1\colon \quad \mu^{\text{before}} \neq \mu^{\text{after}}$$

Knowing that

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2 + \sigma_Y^2}/\sqrt{n}} \sim \mathcal{N}(0,1)$$

we fail to reject H_0 iff

$$-c < Z < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(-c < Z < +c) = 1 - \alpha$ where the probability α of type I error is small.

Paired-sample *z*-test for the difference of the pop.means

Statistical paired-sample z-test for the difference of the population means with two-sided alternative hypothesis ($\mu^{\text{before}} \neq \mu^{\text{after}}$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} \varphi(x)\,\mathrm{d}x = \frac{1}{2}\alpha \quad \text{and} \quad \int_{+c}^{+\infty} \varphi(x)\,\mathrm{d}x = \frac{1}{2}\alpha$$

  for $\alpha = 5\,\%$, the (two-sided) critical value is $c = 1.959963\ldots$

- if $Z \in (-\infty, -c] \cup [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-c, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Paired-sample *z*-test for the difference of the pop.means

In the second case (H_1: $\mu^{\text{before}} < \mu^{\text{after}}$), we have:

$$H_0: \quad \mu^{\text{before}} = \mu^{\text{after}}$$

$$H_1: \quad \mu^{\text{before}} < \mu^{\text{after}}$$

Knowing that

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2 + \sigma_Y^2} / \sqrt{n}} \sim \mathcal{N}(0,1)$$

we <u>fail to reject</u> H_0 iff

$$-c < Z$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(-c < Z) = 1 - \alpha$ where the probability α of type I error is small.

Paired-sample *z*-test for the difference of the pop.means

Statistical paired-sample z-test for the difference of the population means with one-sided alternative hypothesis ($\mu^{\text{before}} < \mu^{\text{after}}$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find **the critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} \varphi(x)\,\mathrm{d}x = \alpha$$

  for $\alpha = 5\,\%$, the (one-sided) critical value is $-c = -1.64485\ldots$

- if $Z \in (-\infty, -c]$, **the critical region**, then **<u>reject</u>** the null hypothesis
- if $Z \in (-c, +\infty)$, then **<u>do not reject</u>** (or <u>fail to reject</u>) the null hypothesis

Paired-sample *z*-test for the difference of the pop.means

In the third case $(H_1\colon\ \mu^{\text{before}} > \mu^{\text{after}})$, we have:

$$H_0\colon \quad \mu^{\text{before}} = \mu^{\text{after}}$$

$$H_1\colon \quad \mu^{\text{before}} > \mu^{\text{after}}$$

Knowing that

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2 + \sigma_Y^2}/\sqrt{n}} \sim \mathcal{N}(0,1)$$

we <u>fail to reject</u> H_0 iff

$$Z < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(Z < +c) = 1 - \alpha$ where the probability α of type I error is small.

Paired-sample *z*-test for the difference of the pop.means

Statistical paired-sample *z*-test for the difference of the population means with one-sided alternative hypothesis ($\mu^{\text{before}} > \mu^{\text{after}}$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$
- find **the critical value** $c > 0$ so that

$$\int_{+c}^{+\infty} \varphi(x)\,\mathrm{d}x = \alpha$$

  for $\alpha = 5\ \%$, the (one-sided) critical value is $+c = +1.64485 \ldots$

- if $Z \in [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-\infty, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Two-sample *z*-test for the difference of the popul. means

Motivation:

We have two unknown random variables X and Y. We ask (test the hypothesis) whether the population means of both random variables are the same.

We assume that both random variables are normal, i.e. $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, and that we know their variances σ_X^2 and σ_Y^2.

We sample the variable X m-times, so we have the sample $x_1, x_2, \dots, x_m$.

We sample the variable Y n-times, so we have the sample $y_1, y_2, \dots, y_n$.

Two-sample z-test for the difference of the popul. means

Having the m observations $x_1, x_2, \dots, x_m$ of the random variable $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and having the n observations $y_1, y_2, \dots, y_n$ of the random variable $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, we formulate the **null hypothesis**:

both samples come from the same population:
the values of the population means are the same

$$H_0: \quad \mu_X = \mu_Y$$

Recall that we do not know the true population means μ_X and μ_Y. We only test the hypothesis by means of two samples of m and n measurements.

Two-sample z-test for the difference of the popul. means

Having the m observations $x_1, x_2, \ldots, x_m$ of the random variable $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$,

the n observations $y_1, y_2, \ldots, y_n$ of the random variable $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, and

$$H_0: \quad \mu_X = \mu_Y$$

formulate the alternative hypothesis:

- two-sided: $H_1: \quad \mu_X \neq \mu_Y$ (the means are different)
- one-sided: $H_1: \quad \mu_X < \mu_Y$ (the first mean < the second mean)
- one-sided: $H_1: \quad \mu_X > \mu_Y$ (the first mean > the second mean)

Two-sample z-test for the difference of the popul. means

Recall the theorem:

If $X_1, X_2, \dots, X_m \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y_1, Y_2, \dots, Y_n \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are independent, then

$$\frac{\bar{X} - \mu_X}{\sigma_X / \sqrt{m}} \sim \mathcal{N}(0,1) \quad \text{and} \quad \frac{\bar{Y} - \mu_Y}{\sigma_Y / \sqrt{n}} \sim \mathcal{N}(0,1)$$

equivalently

$$\bar{X} \sim \mathcal{N}\left(\mu_X, \frac{\sigma_X^2}{m}\right) \quad \text{and} \quad \bar{Y} \sim \mathcal{N}\left(\mu_Y, \frac{\sigma_Y^2}{n}\right)$$

Therefore:

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}\right)$$

Two-sample *z*-test for the difference of the popul. means

<u>**We have shown:**</u>

If $X_1, X_2, \dots, X_m \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y_1, Y_2, \dots, Y_n \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are independent, then

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}\right)$$

equivalently

$$\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}} \sim \mathcal{N}(0,1)$$

Two-sample z-test for the difference of the popul. means

Having the m measurements $x_1, x_2, \dots, x_m$ and n measurements $y_1, y_2, \dots, y_n$, calculate the statistic

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}} \quad \text{or} \quad Z = \frac{\bar{x} - \bar{y} - \mu_0}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}}$$

where

- $\bar{x} = \sum_{i=1}^{m} x_i/m$ is the sample mean of the <u>first</u> sample
- $\bar{y} = \sum_{j=1}^{n} y_j/n$ is the sample mean of the <u>second</u> sample
- m and n is the size of the first and second, respectively, sample
- $\mu_0 = 0$ for no difference of the means $(\mu_X = \mu_Y)$
- $\mu_0 = \text{const.}$ for a general difference of the means $(\mu_X = \mu_Y + \text{const.})$

Two-sample *z*-test for the difference of the popul. means

In the first case $(H_1\colon\ \mu_X \neq \mu_Y)$, we have:

$$H_0\colon \quad \mu_X = \mu_Y$$

$$H_1\colon \quad \mu_X \neq \mu_Y$$

Knowing that

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}} \sim \mathcal{N}(0,1)$$

we <u>fail to reject</u> H_0 iff

$$-c < Z < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that $P(-c < Z < +c) = 1 - \alpha$ where the probability α of type I error is small.

Two-sample z-test for the difference of the popul. means

Statistical two-sample z-test for the difference of the population means with two-sided alternative hypothesis ($\mu_X \neq \mu_Y$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} \varphi(x)\,\mathrm{d}x = \frac{1}{2}\alpha \qquad \text{and} \qquad \int_{+c}^{+\infty} \varphi(x)\,\mathrm{d}x = \frac{1}{2}\alpha$$

for $\alpha = 5\,\%$, the (two-sided) critical value is $c = 1.959963\ldots$

- if $Z \in (-\infty, -c] \cup [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-c, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Two-sample *z*-test for the difference of the popul. means

In the second case $(H_1\colon\ \mu_X < \mu_Y)$, we have:

$$H_0\colon \quad \mu_X = \mu_Y$$

$$H_1\colon \quad \mu_X < \mu_Y$$

Knowing that

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}} \sim \mathcal{N}(0,1)$$

we fail to reject H_0 iff

$$-c < Z$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(-c < Z) = 1 - \alpha$ where the probability α of type I error is small.

Two-sample z-test for the difference of the popul. means

Statistical two-sample *z*-test for the difference of the population means with one-sided alternative hypothesis ($\mu_X < \mu_Y$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$
- find **the critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} \varphi(x)\,\mathrm{d}x = \alpha$$

  for $\alpha = 5\ \%$, the (one-sided) critical value is $-c = -1.64485 \ldots$

- if $Z \in (-\infty, -c]$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-c, +\infty)$, then **do not reject** (or fail to reject) the null hypothesis

Two-sample *z*-test for the difference of the popul. means

In the third case $(H_1\colon\ \mu_X > \mu_Y)$, we have:

$$H_0\colon \quad \mu_X = \mu_Y$$

$$H_1\colon \quad \mu_X > \mu_Y$$

Knowing that

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}} \sim \mathcal{N}(0,1)$$

we fail to reject H_0 iff

$$Z < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(Z < +c) = 1 - \alpha$ where the probability α of type I error is small.

Two-sample *z*-test for the difference of the popul. means

Statistical two-sample *z*-test for the difference of the population means with one-sided alternative hypothesis ($\mu_X > \mu_Y$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical value** $c > 0$ so that

$$\int_{+c}^{+\infty} \varphi(x)\,\mathrm{d}x = \alpha$$

  for $\alpha = 5\,\%$, the (one-sided) critical value is $+c = +1.64485\ldots$
- if $Z \in [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $Z \in (-\infty, +c)$, then **do not reject** (or fail to reject) the null hypothesis

t-tests for the means

- One-sample *t*-test for the population mean
- Paired-sample *t*-test for the difference of the population means
- Two-sample *t*-test for the difference of the population means

One-sample *t*-test for the population mean

<u>Recall the theorem:</u>

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$

Recall that, above, we have been using this theorem for the *z*-tests about the mean.

Notice, however, that we hardly ever know the variance σ^2 of the random variable X in practice.

That is why we recall another theorem now.

One-sample *t*-test for the population mean

Recall another theorem:

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

where

- $s^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$ is the sample variance of the random variables
- χ^2_{n-1} is Pearson's χ^2-distribution with $n-1$ degrees of freedom

One-sample *t*-test for the population mean

<u>Recall the corollary of the two theorems:</u>

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$

where

- $\bar{X} = \sum_{i=1}^{n} X_i / n$ is the sample mean of the random variables
- $s = \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 / (n-1)}$ is the sample standard deviation of the random variables
- t_{n-1} is Student's *t*-distribution with $n-1$ degrees of freedom

One-sample *t*-test for the population mean

Recall the corollary:

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$

Example or motivation:

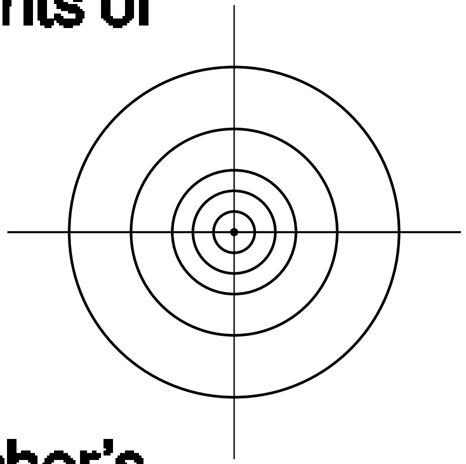
Assume that $X \sim \mathcal{N}(\mu, \sigma^2)$ is a random variable following the normal probability distribution with some mean $\mu \in \mathbb{R}$ and with some variance $\sigma^2 \in \mathbb{R}_0^+$.

Knowing neither the variance σ^2 nor the true value of the population mean $\mu \in \mathbb{R}$, we conjecture / we assume / we … / that the population mean $\mu = \mu_0$, i.e. the (unknown) population mean μ is equal to some prescribed value $\mu_0 \in \mathbb{R}$.

One-sample *t*-test for the population mean

Example: An archer shoots an arrow against the plane.

The sample space $\Omega = \mathbb{R}^2 = \{[x, y] : x, y \in \mathbb{R}\}$ is the set of all the points of the plane. The random variable X is the x-coordinate of the hit, i.e. $X(\omega) = X([x, y]) = x$.



We do not know the archer's variance σ^2 and we do not know the archer's intention, i.e. we do not know the point which the archer intends to hit, i.e. we do not know the archer's mean μ.

We conjecture that the archer's intention is to hit the origin, i.e. $\mu = 0$.

One-sample *t*-test for the population mean

Let $x_1 = X(\omega_1),\ x_2 = X(\omega_2),\ \ldots,\ x_n = X(\omega_n)$ be the numerical results of n trials of a random experiment, where $X \sim \mathcal{N}(\mu, \sigma^2)$,

such as the x-coordinates of the archer's n hits.

We do not know the variance σ^2 and we do not know the mean μ.

We state the **null hypothesis** (about the mean):

$$H_0: \quad \mu = \mu_0$$

where $\mu_0 \in \mathbb{R}$ is some number such that we conjecture that the true mean could equal the μ_0.

One-sample *t*-test for the population mean

Having stated the **null hypothesis**

$$H_0: \quad \mu = \mu_0$$

we also state the **alternative hypothesis:**

- two-sided: $H_1: \quad \mu \neq \mu_0$
- one-sided: $H_1: \quad \mu < \mu_0$
- one-sided: $H_1: \quad \mu > \mu_0$

One-sample *t*-test for the population mean

Which alternative hypothesis ($\mu \neq \mu_0$ or $\mu < \mu_0$ or $\mu > \mu_0$) do we choose?

→ That depends upon our knowledge of the situation.

In our example:

- If we suspect that the archer's intention is to hit a point different from the given point (such as the origin), we choose H_1: $\mu \neq \mu_0$.
- If we conjecture that the archer's intention is to hit a point to the left of the given point (such as the origin), we choose H_1: $\mu < \mu_0$.
- If we conjecture that the archer's intention is to hit a point to the right of the given point (such as the origin), we choose H_1: $\mu > \mu_0$.

One-sample *t*-test for the population mean

Under our assumptions ($x_1, \dots, x_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent and $\mu = \mu_0$), it follows by the Theorem that

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$$

Thus, having the n measurements $x_1, x_2, \dots, x_n$, we calculate the statistic

$$T = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

We know (or assume) that $T \sim t_{n-1}$.

One-sample *t*-test for the population mean

We have $T = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$.

Then, if $-\infty \leq a < b \leq +\infty$, the probability that $a < T < b$ is

$$P(a < T < b) = \int_a^b f(x)\,\mathrm{d}x$$

where

$$f(x) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\sqrt{\pi(n-1)}}\left(1 + \frac{x^2}{n-1}\right)^{-\frac{k}{2}}$$

is the density of Student's *t*-distribution with $n-1$ degrees of freedom.

The gamma function

$$\Gamma(z) = \int_0^{+\infty} x^z e^{-x} \, dx \qquad \text{for} \quad z \in \mathbb{C} \ \text{ such that } \ \mathrm{Re}(z) > 0$$

It is easy to calculate:

$$\Gamma(1) = 1$$

$$\Gamma(z+1) = z\Gamma(z)$$

<u>Therefore:</u>

$$\Gamma(n+1) = n! \qquad \text{for} \quad n = 0, 1, 2, 3, \ldots$$

The gamma function – another definition (due to Euler)

$$\Gamma(z) = \frac{1}{z}\prod_{n=1}^{\infty}\frac{\left(1+\frac{1}{n}\right)^{z}}{1+\frac{z}{n}} \qquad \text{for} \qquad z \in \mathbb{C} \setminus \{0, -1, -2, -3, \ldots\}$$

One-sample *t*-test for the population mean

Consider the first case $(H_1\colon \mu \neq \mu_0)$ first. We have:

$$H_0\colon \quad \mu = \mu_0$$

$$H_1\colon \quad \mu \neq \mu_0$$

Knowing that $T = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$, the probability

$$P(-c < T < +c)$$

is quite high,

so having $-c < T < +c$ accords with H_0, if $c > 0$ is large enough.

One-sample *t*-test for the population mean

On the other hand, if H_0 is true, then it is quite improbable that $T \notin (-c, +c)$.

Therefore, **if we observe that**

$$T \leq -c \qquad \text{or} \qquad +c \leq T$$

then we may conclude that H_0 **is <u>probably not true</u>,**

i.e. **we <u>reject the null hypothesis</u>** H_0.

Therefore, the statistical test proceeds as follows:

(see below)

One-sample *t*-test for the population mean

Statistical one-sample *t*-test with two-sided alternative hypothesis ($\mu \neq \mu_0$):

- choose **the level of significance**, a small number $\alpha > 0$, a very popular value is $\alpha = 5$ %, other popular values are 10 % or 1 % or 0.1 % etc.
- find the **critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha \qquad \text{and} \qquad \int_{+c}^{+\infty} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha$$

where f is the density of the *t*-distribution with $n-1$ degrees of freedom

- if $T \in (-\infty, -c] \cup [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-c, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Type I and Type II error

There are exactly four possibilities when testing the null hypothesis H_0:

- the null hypothesis (H_0) is actually true & we do not reject it — OK
- the null hypothesis (H_0) is actually true & we reject it — type I error
- the null hypothesis (H_0) is actually not true & we do not reject it — type II error
- the null hypothesis (H_0) is actually not true & we reject it — OK

The probability of the type I error is the **<u>significance level</u>** α

The probability of the type II error is β

The **<u>power of the test</u>** is the probability $1-\beta$

One-sample *t*-test for the population mean

Consider now the second case (H_1: $\mu < \mu_0$). We have:

$$H_0: \quad \mu = \mu_0$$

$$H_1: \quad \mu < \mu_0$$

Knowing that $T = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$ and assuming that $c > 0$ is large enough,

what is the probability that $-c < T$?

If H_0 is true, then it is quite improbable that $T \notin (-c, +\infty)$.

Therefore, if we observe that $T \leq -c$, then we may conclude that

H_0 is <u>probably not true</u>, i.e. we <u>reject the null hypothesis</u> H_0.

One-sample *t*-test for the population mean

Statistical one-sample *t*-test with one-sided alternative hypothesis ($\mu < \mu_0$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$, other popular values are $\alpha = 10\ \%$ or $\alpha = 1\ \%$ or $\alpha = 0.1\ \%$ etc.
- find **the critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} f(x)\,\mathrm{d}x = \alpha$$

  where f is the density of the *t*-distribution with $n-1$ degrees of freedom
- if $T \in (-\infty, -c]$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-c, +\infty)$, then **do not reject** (or fail to reject) the null hypothesis

One-sample *t*-test for the population mean

Consider finally the third case (H_1: $\mu > \mu_0$). We have:

$$H_0: \quad \mu = \mu_0$$

$$H_1: \quad \mu > \mu_0$$

Knowing that $T = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$ and assuming that $c > 0$ is large enough,

what is the probability that $T < +c$?

If H_0 is true, then it is quite improbable that $T \notin (-\infty, +c)$.

Therefore, if we observe that $+c \leq T$, then we may conclude that

H_0 **is <u>probably not true</u>,** i.e. **we <u>reject the null hypothesis</u>** H_0.

One-sample *t*-test for the population mean

Statistical one-sample *t*-test with one-sided alternative hypothesis ($\mu > \mu_0$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$,

  other popular values are $\alpha = 10\ \%$ or $\alpha = 1\ \%$ or $\alpha = 0.1\ \%$ etc.

- find the **critical value** $c > 0$ so that

$$\int_{+c}^{+\infty} f(x)\,\mathrm{d}x = \alpha$$

  where f is the density of the *t*-distribution with $n - 1$ degrees of freedom

- if $T \in [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-\infty, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Paired-sample *t*-test for the difference of the pop.means

<u>Example or motivation:</u>

Let us have a sample of n objects, e.g. n patients.

We do two measurements with each of the objects (patients)

— before some treatment

— after the treatment

The purpose it to learn whether the treatment has any effect.

(Hence the null hypothesis: "The treatment has no effect.")

Let $x_1, x_2, \ldots, x_n$ be the values measured before the treatment, and

let $y_1, y_2, \ldots, y_n$ be the values measured after the treatment.

Paired-sample *t*-test for the difference of the pop.means

That is, the measurement x_i and y_i is done with the i-th object (patient) before and after the treatment for $i = 1, 2, \dots, n$.

We assume that $X \sim \mathcal{N}(\mu^{\text{before}}, \sigma_X^2)$, i.e. the random variable of the measurement before the treatment follows the normal distribution, and that $Y \sim \mathcal{N}(\mu^{\text{after}}, \sigma_Y^2)$, i.e. the random variable of the measurement after the treatment also follows the normal distribution, for some $\mu^{\text{before}}, \mu^{\text{after}} \in \mathbb{R}$ and for some $\sigma_X^2, \sigma_Y^2 \in \mathbb{R}_0^+$.
We do not know the true values of the population means μ^{before} and μ^{after}, and we do not know the true values of the variances σ_X^2 and σ_Y^2.

Paired-sample *t*-test for the difference of the pop.means

We formulate the <u>**null hypothesis**</u>:

the treatment has no effect, i.e. the population means are the same

$$H_0: \quad \mu^{\text{before}} = \mu^{\text{after}}$$

Recall that we do not know the true population means μ^{before} and μ^{after}. <u>We only test the hypothesis</u> by having done a sample of n pairs of measurements.

Formulate the alternative hypothesis:

- two-sided: $H_1: \quad \mu^{\text{before}} \neq \mu^{\text{afer}}$ (the treatment has <u>some effect</u>)
- one-sided: $H_1: \quad \mu^{\text{before}} < \mu^{\text{after}}$ (the treatment <u>increases</u> / …
- one-sided: $H_1: \quad \mu^{\text{before}} > \mu^{\text{after}}$ … / <u>decreases</u> the quantity)

Paired-sample *t*-test for the difference of the pop.means

Recall the theorem:

If $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$

Notice also:

If $X_1, X_2, \dots, X_n \sim \mathcal{N}\left(\mu^{\text{before}}, \sigma_X^2\right)$ and $Y_1, Y_2, \dots, Y_n \sim \mathcal{N}\left(\mu^{\text{after}}, \sigma_Y^2\right)$ are independent,

then the differences

$$X_1 - Y_1,\ X_2 - Y_2,\ \dots,\ X_n - Y_n \sim \mathcal{N}\left(\mu^{\text{before}} - \mu^{\text{after}},\ \sigma_X^2 + \sigma_Y^2\right)$$

Now, the hypothesis $\mu^{\text{before}} = \mu^{\text{after}}$ is equivalent to that

the mean of the difference $X - Y$ is $\mu = \mu_0 = 0.$

Paired-sample *t*-test for the difference of the pop.means

We have thus

reduced

the paired-sample *t*-test for the difference of the population means

to

the one-sample *t*-test for the population mean,

which we already know.

Paired-sample *t*-test for the difference of the pop.means

Having the n pairs of the measurements $x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n$, calculate the statistic

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{s^2}/\sqrt{n}} \quad \text{or} \quad T = \frac{\bar{x} - \bar{y} - \mu_0}{\sqrt{s^2}/\sqrt{n}}$$

where

- $\bar{x} = \left(\sum_{i=1}^{n} x_i\right)/n$ is the sample mean of the measurements <u>before</u> the treatment
- $\bar{y} = \left(\sum_{i=1}^{n} y_i\right)/n$ is the sample mean of the measurements <u>after</u> the treatment
- $\mu_0 = 0$ for no difference of the means $(\mu^{\text{before}} = \mu^{\text{after}})$
- $\mu_0 = \text{const.}$ for a general difference of the means $(\mu^{\text{before}} = \mu^{\text{after}} + \text{const.})$
- $s^2 = \left(\sum_{i=1}^{n} (x_i - y_i - \bar{x} + \bar{y})^2\right)/(n-1)$ is the sample variance of the differences

Paired-sample *t*-test for the difference of the pop.means

In the first case $(H_1\colon\ \mu^{\text{before}} \neq \mu^{\text{after}})$, we have:

$$H_0\colon \quad \mu^{\text{before}} = \mu^{\text{after}}$$

$$H_1\colon \quad \mu^{\text{before}} \neq \mu^{\text{after}}$$

Knowing that

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{s^2}/\sqrt{n}} \sim t_{n-1}$$

we fail to reject H_0 iff

$$-c < T < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(-c < T < +c) = 1 - \alpha$ where the probability α of type I error is small.

Paired-sample *t*-test for the difference of the pop.means

Statistical paired-sample *t*-test for the difference of the population means with two-sided alternative hypothesis ($\mu^{\text{before}} \neq \mu^{\text{after}}$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find **the critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha \qquad \text{and} \qquad \int_{+c}^{+\infty} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha$$

  where f is the density of the *t*-distribution with $n-1$ degrees of freedom
- if $T \in (-\infty, -c] \cup [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-c, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Paired-sample *t*-test for the difference of the pop.means

In the second case $(H_1\colon\ \mu^{\text{before}} < \mu^{\text{after}})$, we have:

$$H_0\colon \quad \mu^{\text{before}} = \mu^{\text{after}}$$

$$H_1\colon \quad \mu^{\text{before}} < \mu^{\text{after}}$$

Knowing that

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{s^2}/\sqrt{n}} \sim t_{n-1}$$

we <u>fail to reject</u> H_0 iff

$$-c < T$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(-c < T) = 1 - \alpha$ where the probability α of type I error is small.

Paired-sample *t*-test for the difference of the pop.means

Statistical paired-sample *t*-test for the difference of the population means with one-sided alternative hypothesis ($\mu^{\text{before}} < \mu^{\text{after}}$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find **the critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} f(x)\,\mathrm{d}x = \alpha$$

  where f is the density of the *t*-distribution with $n-1$ degrees of freedom
- if $T \in (-\infty, -c]$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-c, +\infty)$, then **do not reject** (or fail to reject) the null hypothesis

Paired-sample *t*-test for the difference of the pop.means

In the third case $(H_1\colon\ \mu^{\text{before}} > \mu^{\text{after}})$, we have:

$$H_0\colon \quad \mu^{\text{before}} = \mu^{\text{after}}$$

$$H_1\colon \quad \mu^{\text{before}} > \mu^{\text{after}}$$

Knowing that

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{s^2}/\sqrt{n}} \sim t_{n-1}$$

we <u>fail to reject</u> H_0 iff

$$T < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(T < +c) = 1 - \alpha$ where the probability α of type I error is small.

Paired-sample *t*-test for the difference of the pop.means

Statistical paired-sample *t*-test for the difference of the population means with one-sided alternative hypothesis ($\mu^{\text{before}} > \mu^{\text{after}}$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find **the critical value** $c > 0$ so that

$$\int_{+c}^{+\infty} f(x)\,\mathrm{d}x = \alpha$$

  where f is the density of the *t*-distribution with $n-1$ degrees of freedom
- if $T \in [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-\infty, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

<u>Motivation:</u>

We have two unknown random variables X and Y. We ask (test the hypothesis) whether the population means of both random variables are the same.

We assume that both random variables are normal, i.e. $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, for some $\mu_X, \mu_Y \in \mathbb{R}$ and for some $\sigma_X^2, \sigma_Y^2 \in \mathbb{R}_0^+$.

Although we do not know the means μ_X, μ_Y nor the variances σ_X^2, σ_Y^2, we assume that

¡¡¡ ¡¡¡ ¡¡¡ $\sigma_X^2 = \sigma_Y^2$!!! !!! !!!

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Having the m observations $x_1, x_2, \dots, x_m$ of the random variable $X \sim \mathcal{N}(\mu_X, \sigma^2)$ and having the n observations $y_1, y_2, \dots, y_n$ of the random variable $Y \sim \mathcal{N}(\mu_Y, \sigma^2)$, we formulate the **null hypothesis**:

both samples come from the same population:
the values of the population means are the same

$$H_0: \quad \mu_X = \mu_Y$$

Recall that we do not know the true population means μ_X and μ_Y. **We only test the hypothesis** by means of two samples of m and n measurements **with the same variance.**

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Having the m observations $x_1, x_2, \dots, x_m$ of the random variable $X \sim \mathcal{N}(\mu_X, \sigma^2)$,

the n observations $y_1, y_2, \dots, y_n$ of the random variable $Y \sim \mathcal{N}(\mu_Y, \sigma^2)$, and

$$H_0: \quad \mu_X = \mu_Y$$

formulate the alternative hypothesis:

- two-sided: $H_1: \quad \mu_X \neq \mu_Y$ (the means are different)
- one-sided: $H_1: \quad \mu_X < \mu_Y$ (the first mean < the second mean)
- one-sided: $H_1: \quad \mu_X > \mu_Y$ (the first mean > the second mean)

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Recall the theorem:

If $X_1, X_2, \ldots, X_m \sim \mathcal{N}(\mu_X, \sigma^2)$ and $Y_1, Y_2, \ldots, Y_n \sim \mathcal{N}(\mu_Y, \sigma^2)$ are independent, then

$$\frac{\bar{X} - \mu_X}{\sigma/\sqrt{m}} \sim \mathcal{N}(0,1) \quad \text{and} \quad \frac{\bar{Y} - \mu_Y}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

equivalently

$$\bar{X} \sim \mathcal{N}\left(\mu_X, \frac{\sigma^2}{m}\right) \quad \text{and} \quad \bar{Y} \sim \mathcal{N}\left(\mu_Y, \frac{\sigma^2}{n}\right)$$

Therefore:

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_X - \mu_Y, \frac{\sigma^2}{m} + \frac{\sigma^2}{n}\right)$$

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

We have shown:

If $X_1, X_2, \ldots, X_m \sim \mathcal{N}(\mu_X, \sigma^2)$ and $Y_1, Y_2, \ldots, Y_n \sim \mathcal{N}(\mu_Y, \sigma^2)$ are independent, then

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_X - \mu_Y, \frac{\sigma^2}{m} + \frac{\sigma^2}{n}\right)$$

equivalently

$$\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim \mathcal{N}(0, 1)$$

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X=\sigma_Y$

Recall the other theorem:

If $X_1, X_2, \dots, X_m \sim \mathcal{N}(\mu_X, \sigma^2)$ and $Y_1, Y_2, \dots, Y_n \sim \mathcal{N}(\mu_Y, \sigma^2)$ are independent, then

$$\frac{(m-1)s_X^2}{\sigma^2} \sim \chi^2_{m-1} \quad \text{and} \quad \frac{(n-1)s_Y^2}{\sigma^2} \sim \chi^2_{n-1}$$

Therefore:

$$\frac{(m-1)s_X^2 + (n-1)s_Y^2}{\sigma^2} \sim \chi^2_{m+n-2}$$

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X=\sigma_Y$

Recall also that, if

$$Z = \frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim \mathcal{N}(0,1)$$

and

$$Y = \frac{(m-1)s_X^2+(n-1)s_Y^2}{\sigma^2} \sim \chi^2_{m+n-2}$$

then

$$T = \frac{Z}{\sqrt{\frac{Y}{m+n-2}}} \sim t_{m+n-2}$$

by the definition of Student's *t*-distribution.

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X=\sigma_Y$

Therefore:

$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{(m-1)s_X^2+(n-1)s_Y^2}{m+n-2}}\sqrt{\frac{1}{m}+\frac{1}{n}}} = \frac{\dfrac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{m}+\frac{1}{n}}}}{\sqrt{\dfrac{\frac{(m-1)s_X^2+(n-1)s_Y^2}{\sigma^2}}{m+n-2}}} \sim \quad t_{m+n-2}$$

where

$$s_X^2 = \frac{\sum_{i=1}^{m}(X_i-\bar{X})^2}{m-1} \qquad \text{and} \qquad s_Y^2 = \frac{\sum_{j=1}^{n}(Y_i-\bar{Y})^2}{n-1}$$

are the sample variances.

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Having the m measurements $x_1, x_2, \ldots, x_m$ and n measurements $y_1, y_2, \ldots, y_n$, recall that

- $\bar{x} = \sum_{i=1}^{m} x_i / m$ is the sample mean of the first sample
- $\bar{y} = \sum_{j=1}^{n} y_j / n$ is the sample mean of the second sample
- $s_x^2 = \sum_{i=1}^{m} (x_i - \bar{x})^2 / (m-1)$ is the sample variance of the first sample
- $s_y^2 = \sum_{j=1}^{n} (y_j - \bar{y})^2 / (n-1)$ is the sample variance of the second sample
- m is the size of the first sample
- n is the size of the second sample

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Having the m measurements $x_1, x_2, \dots, x_m$ and n measurements $y_1, y_2, \dots, y_n$, calculate the statistic

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(m-1)s_x^2 + (n-1)s_y^2}{m+n-2}} \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

for no difference of the means $(\mu_X = \mu_Y)$

We know (or assume) that $T \sim t_{m+n-2}$

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Or, having the m measurements $x_1, x_2, \dots, x_m$ and n measurements $y_1, y_2, \dots, y_n$, calculate the statistic

$$T = \frac{\bar{x} - \bar{y} - \mu_0}{\sqrt{\frac{(m-1)s_x^2 + (n-1)s_y^2}{m+n-2}}\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

for a general difference of the means $(\mu_X = \mu_Y + \mu_0)$

We know (or assume) that $T \sim t_{m+n-2}$

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

In the first case $(H_1\colon\ \mu_X \neq \mu_Y)$, we have:

$$H_0\colon \quad \mu_X = \mu_Y$$

$$H_1\colon \quad \mu_X \neq \mu_Y$$

Knowing that

$$T \sim t_{m+n-2}$$

we <u>fail to reject</u> H_0 iff

$$-c < T < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(-c < T < +c) = 1 - \alpha$ where the probability α of type I error is small.

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Statistical two-sample *t*-test for the difference of the population means with two-sided alternative hypothesis ($\mu_X \neq \mu_Y$) and with the same variances:

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha \qquad \text{and} \qquad \int_{+c}^{+\infty} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha$$

  where f is the density of the *t*-distribution with $m + n - 2$ degrees of freedom

- if $T \in (-\infty, -c] \cup [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-c, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X=\sigma_Y$

In the second case $(H_1\colon \mu_X < \mu_Y)$, we have:

$$H_0\colon \quad \mu_X = \mu_Y$$

$$H_1\colon \quad \mu_X < \mu_Y$$

Knowing that

$$T \sim t_{m+n-2}$$

we <u>fail to reject</u> H_0 iff

$$-c < T$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(-c < T) = 1 - \alpha$ where the probability α of type I error is small.

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Statistical two-sample *t*-test for the difference of the population means with one-sided alternative hypothesis ($\mu_X < \mu_Y$) and with the same variances :

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$
- find **the critical value** $c > 0$ so that

$$\int_{-\infty}^{-c} f(x)\,\mathrm{d}x = \alpha$$

  where f is the density of the *t*-distribution with $m + n - 2$ degrees of freedom
- if $T \in (-\infty, -c]$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-c, +\infty)$, then **do not reject** (or fail to reject) the null hypothesis

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X=\sigma_Y$

In the third case $(H_1\colon\ \mu_X > \mu_Y)$, we have:

$$H_0\colon\quad \mu_X = \mu_Y$$

$$H_1\colon\quad \mu_X > \mu_Y$$

Knowing that

$$T \sim t_{m+n-2}$$

we <u>fail to reject</u> H_0 iff

$$T < +c$$

where the critical value $c > 0$, under the assumption that H_0 is true, is such that

$P(T < +c) = 1 - \alpha$ where the probability α of type I error is small.

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X = \sigma_Y$

Statistical two-sample *t*-test for the difference of the population means with one-sided alternative hypothesis ($\mu_X > \mu_Y$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical value** $c > 0$ so that

$$\int_{+c}^{+\infty} f(x)\,\mathrm{d}x = \alpha$$

where f is the density of the *t*-distribution with $m + n - 2$ degrees of freedom

- if $T \in [+c, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $T \in (-\infty, +c)$, then **do not reject** (or fail to reject) the null hypothesis

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X \neq \sigma_Y$

Consider two normal random variables $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, for some $\mu_X, \mu_Y \in \mathbb{R}$ and for some $\sigma_X^2, \sigma_Y^2 \in \mathbb{R}_0^+$.

We ask (test the hypothesis) whether the population means of both random variables are the same.

Once $\sigma_X^2 = \sigma_Y^2$ is not assumed, the things get complicated.

We have an approximate result only.

Theorem (Satterthwaite's approximation):

If the random variables $X_1, X_2, \dots, X_m \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y_1, Y_2, \dots, Y_n \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are independent, then the statistic

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{s_X^2}{m} + \frac{s_Y^2}{n}}} \sim t_\nu \qquad \textit{approximately}$$

where

$$\nu = \frac{\left(\frac{s_X^2}{m} + \frac{s_Y^2}{n}\right)^2}{\frac{s_X^4}{m^2(m-1)} + \frac{s_Y^4}{n^2(n-1)}}$$

Two-sample *t*-test for the diff. of the pop. means // $\sigma_X \neq \sigma_Y$

Exercise:

Use the last Theorem (Satterthwaite's approximation) to formulate a statistical two-sample *t*-test for the difference of the population means with two-sided / one-sided alternative hypothesis (not assuming the same variance).

χ^2-test for the variance

- χ^2-test for the variance

χ^2-test for the population variance

<u>Recall the theorem:</u>

If $X_1, X_2, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent, then $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

<u>Example or motivation:</u>

Assume that $X \sim \mathcal{N}(\mu, \sigma^2)$ is a random variable following the normal probability distribution. We do not know the population mean $\mu \in \mathbb{R}$ and we do not know the true value of the population variance $\sigma^2 \in \mathbb{R}^+$ either.

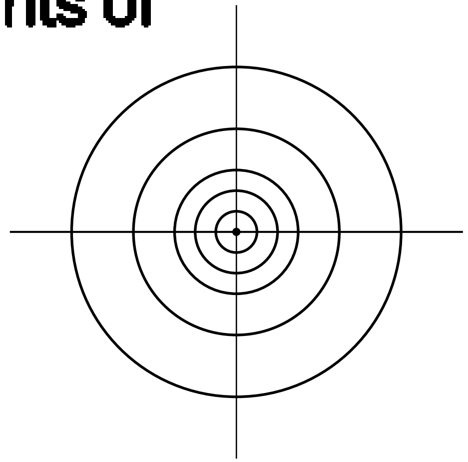
We conjecture / We assume / We … / that the population variance $\sigma^2 = \sigma_0^2$, i.e. the (unknown) population variance σ^2 is equal to some prescribed value $\sigma_0^2 \in \mathbb{R}^+$.

χ^2-test for the population variance

Example: An archer shoots an arrow against the plane.

The sample space $\Omega = \mathbb{R}^2 = \{[x,y] : x, y \in \mathbb{R}\}$ is the set of all the points of the plane. The random variable X is the x-coordinate of the hit, i.e.

$X(\omega) = X([x,y]) = x.$



We do not know the archer's mean μ and

we do not know the archer's variance σ^2.

We conjecture that the archer's variance could be $\sigma^2 = \sigma_0^2$

where $\sigma_0^2 \in \mathbb{R}^+$ is some prescribed value.

χ^2-test for the population variance

Let $x_1 = X(\omega_1),\ x_2 = X(\omega_2),\ \ldots,\ x_n = X(\omega_n)$ be the numerical results of n trials of a random experiment, where $X \sim \mathcal{N}(\mu, \sigma^2)$,

such as the x-coordinates of the archer's n hits.

We do not know the mean μ and we do not know the variance σ^2.

We state the **null hypothesis** (about the variance):

$$H_0: \quad \sigma^2 = \sigma_0^2$$

where $\sigma_0^2 \in \mathbb{R}^+$ is some number such that we conjecture that the true variance could equal the σ_0^2.

χ^2-test for the population variance

Having stated the **null hypothesis**

$$H_0: \quad \sigma^2 = \sigma_0^2$$

we also state the **alternative hypothesis:**

- two-sided: $H_1: \quad \sigma^2 \neq \sigma_0^2$
- one-sided: $H_1: \quad \sigma^2 < \sigma_0^2$
- one-sided: $H_1: \quad \sigma^2 > \sigma_0^2$

χ^2-test for the population variance

Under our assumptions ($x_1, \dots, x_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent and $\sigma^2 = \sigma_0^2$),

it follows by the Theorem that

$$\frac{(n-1)s^2}{\sigma^2} = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

Thus, having the n measurements $x_1, x_2, \dots, x_n$, we calculate the statistic

$$X^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

We know (or assume) that $X^2 \sim \chi^2_{n-1}$.

χ^2-test for the population variance

Consider the first case $(H_1\colon \sigma^2 \neq \sigma_0^2)$ first. We have:

$$H_0\colon \quad \sigma^2 = \sigma_0^2$$

$$H_1\colon \quad \sigma^2 \neq \sigma_0^2$$

Knowing that $X^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}$ and assuming that $0 < c < d$ is small enough and large enough, respectively, what is the probability that $c < X^2 < d$?

If H_0 is true, then it is quite improbable that $X^2 \notin (c, d)$.

Therefore, if we observe that $X^2 \leq c$ or $d \leq X^2$, then we may conclude that H_0 is <u>probably not true</u>, i.e. we <u>reject the null hypothesis</u> H_0.

χ^2-test for the population variance

Statistical χ^2-test with two-sided alternative hypothesis ($\sigma^2 \neq \sigma_0^2$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$, other popular values are $\alpha = 10\,\%$ or $\alpha = 1\,\%$ or $\alpha = 0.1\,\%$ etc.
- find the **critical values** $0 < c < d$ so that

$$\int_0^c f(x)\,\mathrm{d}x = \frac{\alpha}{2} \quad \text{and} \quad \int_d^{+\infty} f(x)\,\mathrm{d}x = \frac{\alpha}{2}$$

  where f is the density of the χ^2-distribution with $n-1$ degrees of freedom
- if $X^2 \in [0, c] \cup [d, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $X^2 \in (c, d)$, then **do not reject** (or fail to reject) the null hypothesis

χ^2-test for the population variance

Consider now the second case $(H_1\!: \sigma^2 < \sigma_0^2)$. We have:

$$H_0\!: \quad \sigma^2 = \sigma_0^2$$

$$H_1\!: \quad \sigma^2 < \sigma_0^2$$

Knowing that $X^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}$ and assuming that $0 < c$ is small enough,

what is the probability that $c < X^2$?

If H_0 is true, then it is quite improbable that $X^2 \notin (c, +\infty)$.

Therefore, **if we observe that** $X^2 \leq c$, **then we may conclude** that

H_0 **is probably not true**, i.e. **we reject the null hypothesis** H_0.

χ^2-test for the population variance

Statistical χ^2-test with one-sided alternative hypothesis ($\sigma^2 < \sigma_0^2$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$, other popular values are $\alpha = 10\ \%$ or $\alpha = 1\ \%$ or $\alpha = 0.1\ \%$ etc.
- find the **critical value** $c > 0$ so that

$$\int_0^c f(x)\,\mathrm{d}x = \alpha$$

  where f is the density of the χ^2-distribution with $n-1$ degrees of freedom
- if $X^2 \in [0, c]$, **the critical region**, then **reject** the null hypothesis
- if $X^2 \in (c, +\infty)$, then **do not reject** (or fail to reject) the null hypothesis

χ^2-test for the population variance

Consider finally the third case $(H_1\colon \sigma^2 > \sigma_0^2)$. We have:

$$H_0\colon \quad \sigma^2 = \sigma_0^2$$

$$H_1\colon \quad \sigma^2 > \sigma_0^2$$

Knowing that $X^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}$ and assuming that $d > 0$ is large enough,

what is the probability that $X^2 < d$?

If H_0 is true, then it is quite improbable that $X^2 \notin [0, d)$.

Therefore, if we observe that $d \leq X^2$, then we may conclude that

H_0 is <u>probably not true</u>, i.e. we <u>reject the null hypothesis</u> H_0.

χ^2-test for the population variance

Statistical χ^2-test with one-sided alternative hypothesis ($\sigma^2 > \sigma_0^2$):

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$, other popular values are $\alpha = 10\ \%$ or $\alpha = 1\ \%$ or $\alpha = 0.1\ \%$ etc.
- find the **critical value** $d > 0$ so that

$$\int_d^{+\infty} f(x)\,\mathrm{d}x = \alpha$$

  where f is the density of the χ^2-distribution with $n-1$ degrees of freedom
- if $X^2 \in [d, +\infty)$, **the critical region**, then **reject** the null hypothesis
- if $X^2 \in [0, d)$, then **do not reject** (or fail to reject) the null hypothesis

z-test for the population proportion

- One-sample binomial test for the population proportion
- One-sample *z*-test for the population proportion

One-sample z-test for the population proportion

Motivation: We are tossing a coin repeatedly, and we ask:

¿ Is the coin fair ?

More generally:

Consider a Bernoulli trial, with the probability of the success being $p \in (0, 1)$,

and with the probability of the failure being $q = 1 - p$.

We do not know the true probability p.

We conjecture / We assume / We … / that the probability $p = p_0$, i.e.

the (unknown) probability p is equal to some prescribed value $p_0 \in (0, 1)$,

e.g., in the case of the coin, conjecture that $p_0 = 50\ \%$ (meaning the coin is fair).

One-sample *z*-test for the population proportion

Let $X_1, X_2, X_3, \dots$ be a sequence of random variables, each of which represents the result of a Bernoulli Trial (independent of the other trials), attaining the value $X_i = 1$ (success) with a fixed probability $p \in (0, 1)$ and attaining the value $X_i = 0$ (failure) with the fixed probability $q = 1 - p$. (So that $p + q = 1$ and $p, q > 0$.) <u>We do not know the probability</u> p.

We conjecture / We assume / We … / that the probability $p = p_0$, i.e. the (unknown) probability p is equal to some prescribed value $p_0 \in (0, 1)$.

One-sample *z*-test for the population proportion

Having stated the **null hypothesis**

$$H_0: \quad p = p_0$$

we also state the **alternative hypothesis**:

- two-sided: $H_1: \quad p \neq p_0$
- one-sided: $H_1: \quad p < p_0$
- one-sided: $H_1: \quad p > p_0$

One-sample binomial test for the population proportion

Let $x_1 = X(\omega_1),\ x_2 = X(\omega_2),\ \ldots,\ x_n = X(\omega_n)$ be the outcomes of the n Bernoulli trials where the probability of success ("1") is p and the probability of failure ("0") is $q = 1 - p$.

For every $k \in \{0, 1, \ldots, n\}$, the probability that $\sum_{i=1}^{n} x_i = k$ is $\binom{n}{k} p^k q^{n-k}$.

Let $K, L \in \{0, 1, \ldots, n\}$ be such that $K < L$.

The probability that $\sum_{i=1}^{n} x_i \in \{0, 1, \ldots, K\}$ is $\sum_{k=0}^{K} \binom{n}{k} p^k q^{n-k}$.

The probability that $\sum_{i=1}^{n} x_i \in \{L, L+1, \ldots, n\}$ is $\sum_{k=L}^{n} \binom{n}{k} p^k q^{n-k}$.

One-sample binomial test for the population proportion

Consider the first case $(H_1\colon p \neq p_0)$ first. We have:

$$H_0\colon \quad p = p_0$$
$$H_1\colon \quad p \neq p_0$$

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical values** $K, L \in \{0, 1, \dots, n\}$ so that

  K is the largest number and L is the least number such that

$$\sum_{k=0}^{K} \binom{n}{k} p_0^k q_0^{n-k} \leq \frac{\alpha}{2} \qquad \text{and} \qquad \sum_{k=L}^{n} \binom{n}{k} p_0^k q_0^{n-k} \leq \frac{\alpha}{2}$$

- if $\sum_{i=1}^{n} x_i \in \{0, \dots, K\} \cup \{L, \dots, n\}$, **the critical region**, then **reject** the null hyp.
- if $\sum_{i=1}^{n} x_i \in \{K+1, \dots, L-1\}$, then **do not reject** (or fail to reject) the null hyp.

One-sample binomial test for the population proportion

Consider now the second case ($H_1\colon p < p_0$). We have: $H_0\colon \quad p = p_0$

$H_1\colon \quad p < p_0$

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical value** $K \in \{0, 1, \dots, n\}$ so that K is the largest number such that

$$\sum_{k=0}^{K} \binom{n}{k} p_0^k q_0^{n-k} \le \alpha$$

- if $\sum_{i=1}^{n} x_i \in \{0, \dots, K\}$, **the critical region**, then **reject** the null hypothesis
- if $\sum_{i=1}^{n} x_i \in \{K + 1, \dots, n\}$, then **do not reject** (or fail to reject) the null hypothesis

One-sample binomial test for the population proportion

Consider finally the third case (H_1: $p > p_0$). We have:

$$H_0: \quad p = p_0$$
$$H_1: \quad p > p_0$$

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find the **critical value** $L \in \{0, 1, \dots, n\}$ so that L is the least number such that

$$\sum_{k=L}^{n} \binom{n}{k} p_0^k q_0^{n-k} \leq \alpha$$

- if $\sum_{i=1}^{n} x_i \in \{L, \dots, n\}$, **the critical region**, then <u>reject</u> the null hypothesis
- if $\sum_{i=1}^{n} x_i \in \{0, \dots, L-1\}$, then <u>**do not reject**</u> (or <u>fail to reject</u>) the null hypothesis

One-sample z-test for the population proportion

It is inconvenient to calculate the sums $\sum_{k=0}^{K}\binom{n}{k}p_0^k q_0^{n-k}$ and $\sum_{k=L}^{n}\binom{n}{k}p_0^k q_0^{n-k}$ if n is large. It is more convenient then to approximate the sums by using <u>the de Moivre-Laplace Central Limit Theorem:</u>

Let $p \in (0,1)$ be the probability of success in a Bernoulli Trial, and let $q = 1 - p$ be the probability of failure in the trial. Whenever $-\infty \leq a < b \leq +\infty$, it then holds

$$\frac{\sum_{k=A_n}^{B_n}\binom{n}{k}p^k q^{n-k} - np}{\sqrt{npq}} \longrightarrow \underbrace{\int_a^b \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t}_{\Phi(b)-\Phi(a)} \qquad \text{as} \quad n \to \infty$$

where $A_n = [np + a\sqrt{npq}] \geq 0$ and $B_n = [np + b\sqrt{npq}] \leq n$ if $n \geq \max\left(\frac{q}{p}a^2, \frac{p}{q}b^2\right)$.

One-sample *z*-test for the population proportion

De Moivre-Laplace Central Limit Theorem (reformulated):

Let $X \sim \mathrm{Bi}(n, p)$ with $p \in (0,1)$, and put $q = 1 - p$ for short.

Whenever $-\infty \leq a < b \leq +\infty$, it then holds

$$P\left(a < \frac{X - np}{\sqrt{npq}} < b\right) \to \underbrace{\int_a^b \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t}_{\Phi(b)-\Phi(a)} \quad \text{as} \quad n \to \infty$$

and the convergence is uniform with respect to a and b.

One-sample *z*-test for the population proportion

Consider the first case $(H_1\colon p = p_0)$ first. We have:

$$H_0\colon \quad p = p_0$$
$$H_1\colon \quad p \neq p_0$$

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find $c > 0$ so that

$$\int_{-\infty}^{-c} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \,\mathrm{d}t = \frac{\alpha}{2} \qquad \text{and} \qquad \int_{+c}^{+\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \,\mathrm{d}t = \frac{\alpha}{2}$$

- if $\sum_{i=1}^{n} x_i \leq np_0 - c\sqrt{np_0q_0}$ or $np_0 + c\sqrt{np_0q_0} \leq \sum_{i=1}^{n} x_i$, **the critical region**, then **reject** the null hypothesis
- if $np_0 - c\sqrt{np_0q_0} < \sum_{i=1}^{n} x_i < np_0 + c\sqrt{np_0q_0}$, then **do not reject** (or fail to reject) the null hypothesis

One-sample *z*-test for the population proportion

Consider now the second case (H_1: $p < p_0$). We have: H_0: $p = p_0$
H_1: $p < p_0$

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\ \%$
- find $c > 0$ so that

$$\int_{-\infty}^{-c} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t = \alpha$$

- if $\sum_{i=1}^{n} x_i \leq np_0 - c\sqrt{np_0q_0}$, **the critical region**, then **reject** the null hypothesis
- if $np_0 - c\sqrt{np_0q_0} < \sum_{i=1}^{n} x_i$, then **do not reject** (or fail to reject) the null hyp.

One-sample *z*-test for the population proportion

Consider finally the third case $(H_1\colon p > p_0)$. We have: $H_0\colon \quad p = p_0$

$H_1\colon \quad p > p_0$

- choose **the level of significance**, a small number $\alpha > 0$, such as $\alpha = 5\,\%$
- find $c > 0$ so that

$$\int_{+c}^{+\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t = \alpha$$

- if $np_0 + c\sqrt{np_0q_0} \le \sum_{i=1}^{n} x_i$, **the critical region**, then **<u>reject</u>** the null hypothesis
- if $\sum_{i=1}^{n} x_i < np_0 + c\sqrt{np_0q_0}$, then **<u>do not reject</u>** (or <u>fail to reject</u>) the null hyp.

p-value of the test

- The general outline

  of a statistical hypothesis test
- The *p*-value of a test

The general outline of a statistical hypothesis test

- A statistical test consists in the study of the outcomes of a random experiment.
- We put down a hypothesis about the probability distribution of the outcomes of the random experiment.
- We also make up a statistic S – a formula, i.e. a mathematical expression – and we **prove** (!) as a mathematical Theorem that, under our hypotheses, the statistic S follows a certain probability distribution.
- We carry out the random experiment several (or many) times.
- We put down the results of the experiment, i.e., count positive results, count the negative results, and so on.

The general outline of a statistical hypothesis test

- We substitute the results (the counts, and so on) into the mathematical expression-statistic S, which is a random variable thus (its value depends on the results of the random experiment).
- We then choose the **significance level** α – a small probability – such as **α = 5 % = 0.05** (i.e. "one error per twenty trials").
  (Other popular choices include α = 10 % = 0.1 or α = 1 % = 0.01.)
- By using the mathematical Theorem, which we proved (see above), we find the **critical region** $C \subseteq \mathbb{R}$ so that – if our hypotheses are true – then **the probability of the event that** $S \in C$ is **$\leq \alpha$**.

The general outline of a statistical hypothesis test

- The critical region C is usually a closed interval or the union of two closed intervals.
- Finally, make a statistical conclusion:

- If $S \in C$, then **reject** the hypothesis.
- If the hypotheses are true, then it is quite improbable that $S \in C$;
  the probability is $\leq \alpha$. So we are making a mistake – type I error,
  i.e. rejecting a hypothesis which is true – about once per twenty trials,
  if $\alpha = 5$ %.

The general outline of a statistical hypothesis test

- If $S \notin C$, then **do not reject** (or **fail to reject**) the hypothesis.
- The fact that we fail to reject the hypothesis is <u>not</u> a confirmation that the hypothesis is true!
- Since the statistic S is a random variable, it may happen by chance that $S \notin C$ even if the hypothesis is false.
- This situation – failing to reject a false hypothesis – is a type II error.
  The probability of type II error is β, and this probability is difficult to calculate…
  If $\alpha = 5\ \%$, then the probability β should be $\leq 20\ \%$. (Is it $\leq 20\ \%$?)
  The probability $1 - \beta$ is the **power of the test.**

The *p*-value of the test

The above outline of the test is as follows:

- Choose the significance level α (such as $\alpha = 5$ %).
- Depending upon the α, find the critical region $C_\alpha \subseteq \mathbb{R}$ so that – if the hypothesis is true – then the probability that $(S \in C_\alpha)$ is $\leq \alpha$.
- Carry out the experiment, enumerate the expression S, and see if $S \in C_\alpha$.

<u>Another procedure:</u>

- Carry out the experiment and enumerate the expression S.
- Find the least number $p \in (0, 1)$ such that $S \in C_p$.
- This value p is the ***p*-value of the test.**

hypothesis.)