Hypothesis testing

Lecturer:
Francesco Dolce

Department of Applied Mathematics
Faculty of Information Technology
Czech Technical University in Prague

Probability and Statistics

BIE-PST, WS 2024/25, Lecture 11

Content

- **Probability theory:**
  - Events, probability, conditional probability, Bayes' Theorem, independence of events.
  - Random variables, distribution function, functions of random variables, characteristics of random variables: expected value, variance, moments, generating function, quantiles, critical values, important discrete and continuous distributions.
  - Random vectors, joint and marginal distributions, functions of random vectors, independence of random variables, conditional distribution, conditional expected value, covariance and correlation.
  - Markov's and Chebyshev's inequality, weak law of large numbers, strong law of large numbers, Central limit theorem.
- **Mathematical statistics:**
  - Point estimators, sample mean, sample variance, properties of point estimators, Maximum likelihood method.
  - Interval estimators, hypothesis testing, one-sided vs. two-sided alternatives, linear regression, estimators of regression parameters, testing of linear model.

Recap

Suppose we observe a **random sample** $X_1, \ldots, X_n$ (independent and identically distributed random variables) from an **unknown distribution**. We aim to estimate:

- The **shape** of the distribution – its type and parametric family.
- The **parameters** of the distribution.

The expectation $\mathrm{E}\, X_i = \mu$ and the variance $\mathrm{var}\, X_i = \sigma^2$ are most often estimated by the sample mean and the sample variance:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad s_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2.$$

Confidence intervals or **interval estimates** for a parameter θ of a distribution are such bounds $L = L(\mathbf{X})$, $U = U(\mathbf{X})$, for which

$$\mathrm{P}(L < \theta < U) = 1 - \alpha.$$

α is chosen as small, typically 5% or 1%. Then we speak of $(1-\alpha)\%$-confidence intervals. The **two-sided** confidence intervals for the expectation μ of a random sample from the **normal distribution** with **known variance** can be found as

$$\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\, , \, \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right),$$

where z denotes the corresponding critical value of the standard normal distribution.

Hypothesis testing – motivation

Often we need to verify a claim about some property of a studied distribution, with only a random sample at our disposal.

Hypothesis testing – motivation

Often we need to verify a claim about some property of a studied distribution, with only a random sample at our disposal.

Delivery of goods

We accept a delivery of goods if the percentage of faulty items is less then some given boundary, e.g., 5%.

Hypothesis testing – motivation

Often we need to verify a claim about some property of a studied distribution, with only a random sample at our disposal.

Delivery of goods

We accept a delivery of goods if the percentage of faulty items is less then some given boundary, e.g., 5%.

Often it is not possible to check the quality of all items. Therefore we select a random sample and test just the quality of sampled items. On the basis of the result we need to decide whether to accept the delivery or not.

Hypothesis testing – motivation

Often we need to verify a claim about some property of a studied distribution, with only a random sample at our disposal.

Delivery of goods

We accept a delivery of goods if the percentage of faulty items is less then some given boundary, e.g., 5%.

Often it is not possible to check the quality of all items. Therefore we select a random sample and test just the quality of sampled items. On the basis of the result we need to decide whether to accept the delivery or not.

Improving a technological procedure

A new technological procedure has been invented. Before using it in normal operation we need to decide whether the new procedure is actually better than the old one or not.

Hypothesis testing – motivation

Often we need to verify a claim about some property of a studied distribution, with only a random sample at our disposal.

Delivery of goods

We accept a delivery of goods if the percentage of faulty items is less then some given boundary, e.g., 5%.

Often it is not possible to check the quality of all items. Therefore we select a random sample and test just the quality of sampled items. On the basis of the result we need to decide whether to accept the delivery or not.

Improving a technological procedure

A new technological procedure has been invented. Before using it in normal operation we need to decide whether the new procedure is actually better than the old one or not.

We measure samples of n_1 and n_2 products manufactured using the old and the new procedure. On the basis of these samples we need to decide whether there is a difference between the old and the new method or not.

Hypothesis testing – hypotheses

Let us observe a **random sample** from some distribution.

Statements about this distribution, which cannot be surely confirmed, are called **hypotheses**.

The mechanism of verifying the validity of hypotheses based on observed data is called **hypothesis testing**.

Hypothesis testing – hypotheses

Let us observe a **random sample** from some distribution.
Statements about this distribution, which cannot be surely confirmed, are called **hypotheses**.

The mechanism of verifying the validity of hypotheses based on observed data is called **hypothesis testing**.

We use two basic notions:

- **Null hypothesis** H_0 denoting the statement which we want to confirm or reject.

Hypothesis testing – hypotheses

Let us observe a **random sample** from some distribution.

Statements about this distribution, which cannot be surely confirmed, are called **hypotheses**.

The mechanism of verifying the validity of hypotheses based on observed data is called **hypothesis testing**.

We use two basic notions:

- **Null hypothesis** H_0 denoting the statement which we want to confirm or reject.
- **Alternative hypothesis** H_A is a converse statement against which we reject H_0.

Hypothesis testing – hypotheses

Let us observe a **random sample** from some distribution.
Statements about this distribution, which cannot be surely confirmed, are called **hypotheses**.

The mechanism of verifying the validity of hypotheses based on observed data is called **hypothesis testing**.

We use two basic notions:

- **Null hypothesis** H_0 denoting the statement which we want to confirm or reject.
- **Alternative hypothesis** H_A is a converse statement against which we reject H_0.

Hypotheses types:

- **Non-parametric** – random sample from a general distribution. The statements deal with various properties of the distribution (e.g., median), or the shape of the distribution (goodness-of-fit tests).

Hypothesis testing – hypotheses

Let us observe a **random sample** from some distribution.
Statements about this distribution, which cannot be surely confirmed, are called **hypotheses**.

The mechanism of verifying the validity of hypotheses based on observed data is called **hypothesis testing**.

We use two basic notions:

- **Null hypothesis** H_0 denoting the statement which we want to confirm or reject.
- **Alternative hypothesis** H_A is a converse statement against which we reject H_0.

Hypotheses types:

- **Non-parametric** – random sample from a general distribution. The statements deal with various properties of the distribution (e.g., median), or the shape of the distribution (goodness-of-fit tests).
- **Parametric** – random sample from a distribution given by parameters $\theta \in \mathbb{R}^d$. We test statements regarding the value of θ.

Hypothesis testing – types of errors

The **test** of a given null hypothesis H_0 against an alternative hypothesis H_A is a **decision procedure** with two possible results: either **rejecting** or **not rejecting** the null hypothesis H_0.

Hypothesis testing – types of errors

The **test** of a given null hypothesis H_0 against an alternative hypothesis H_A is a **decision procedure** with two possible results: either **rejecting** or **not rejecting** the null hypothesis H_0.

While deciding we can make one of two errors:

- Rejecting H_0 even if it is valid – **type I error**.

Hypothesis testing – types of errors

The **test** of a given null hypothesis H_0 against an alternative hypothesis H_A is a **decision procedure** with two possible results: either **rejecting** or **not rejecting** the null hypothesis H_0.

While deciding we can make one of two errors:

- Rejecting H_0 even if it is valid – **type I error**.
- Not rejecting H_0 even if it is not valid (if H_A is valid) – **type II error**.

Hypothesis testing – types of errors

The **test** of a given null hypothesis H_0 against an alternative hypothesis H_A is a **decision procedure** with two possible results: either **rejecting** or **not rejecting** the null hypothesis H_0.

While deciding we can make one of two errors:

- Rejecting H_0 even if it is valid – **type I error**.
- Not rejecting H_0 even if it is not valid (if H_A is valid) – **type II error**.

Often it is possible to have only one of these errors under control.

- We proceed so that the probability of making the type I error is less than or equal to some small given α, which is called the **level of significance**.
- Usually we take $\alpha = 5\%$ or $\alpha = 1\%$.

Hypothesis testing – types of errors

The **test** of a given null hypothesis H_0 against an alternative hypothesis H_A is a **decision procedure** with two possible results: either **rejecting** or **not rejecting** the null hypothesis H_0.

While deciding we can make one of two errors:

- Rejecting H_0 even if it is valid – **type I error**.
- Not rejecting H_0 even if it is not valid (if H_A is valid) – **type II error**.

Often it is possible to have only one of these errors under control.

- We proceed so that the probability of making the type I error is less than or equal to some small given α, which is called the **level of significance**.
- Usually we take $\alpha = 5\%$ or $\alpha = 1\%$.
- The type II error can be either small or large depending on the sample size.
- The probability of **not making** type II error is called the **power** of the test.

Hypothesis testing – possible results

Possible results of testing:

- We reject the hypothesis H_0 in favor of H_A, with a small probability α of making an error.

Hypothesis testing – possible results

Possible results of testing:

- We reject the hypothesis H_0 in favor of H_A, with a small probability α of making an error.
- We do not reject H_0.

Hypothesis testing – possible results

Possible results of testing:

- We reject the hypothesis H_0 in favor of H_A, with a small probability α of making an error.
- We do not reject H_0.

The position of both hypotheses is not symmetric.

As a null hypothesis we choose that one, where the wrongful rejection, i.e., making type I error, would be more serious.

The hypothesis which we need to confirm is chosen as the alternative hypothesis H_A.

Hypothesis testing – possible results

Possible results of testing:

- We reject the hypothesis H_0 in favor of H_A, with a small probability α of making an error.
- We do not reject H_0.

The position of both hypotheses is not symmetric.

As a null hypothesis we choose that one, where the wrongful rejection, i.e., making type I error, would be more serious.

The hypothesis which we need to confirm is chosen as the alternative hypothesis H_A.

Rejection of H_0 in favor of H_A is a strong result.

Hypothesis testing – possible results

Generally we say that we “test a hypothesis H_0 against an alternative H_A at the level of significance α”.

Hypothesis testing – possible results

Generally we say that we “test a hypothesis H_0 against an alternative H_A at the level of significance α”.

- If we reject a hypothesis H_0 and thus can very reliable state that the alternative H_A holds, we say that the statement of H_A is “statistically significant”.

Hypothesis testing – possible results

Generally we say that we "test a hypothesis H_0 against an alternative H_A at the level of significance α".

- If we reject a hypothesis H_0 and thus can very reliable state that the alternative H_A holds, we say that the statement of H_A is "statistically significant".
- If we do not reject H_0, the statement of H_A is called "statistically insignificant".

Parametric tests and confidence intervals – two-sided

Let $X_1, \ldots, X_n$ be a random sample from a distribution with a parameter θ.

We want to test the hypothesis (**two-sided alternative**):

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta \neq \theta_0,$$

for some fixed value θ_0.

Parametric tests and confidence intervals – two-sided

Let $X_1, \ldots, X_n$ be a random sample from a distribution with a parameter θ.

We want to test the hypothesis (**two-sided alternative**):

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta \neq \theta_0,$$

for some fixed value θ_0.

Let further $\big(L(\mathbf{X}), U(\mathbf{X})\big)$ be the two-sided $100(1-\alpha)\%$ confidence interval for θ based on a random sample. Thus it holds that

$$\mathrm{P}\left(\theta \in (L, U)\right) = 1 - \alpha.$$

Parametric tests and confidence intervals – two-sided

Let $X_1, \ldots, X_n$ be a random sample from a distribution with a parameter θ.

We want to test the hypothesis (**two-sided alternative**):

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta \neq \theta_0,$$

for some fixed value θ_0.

Let further $\big(L(\mathbf{X}), U(\mathbf{X})\big)$ be the two-sided $100(1-\alpha)\%$ confidence interval for θ based on a random sample. Thus it holds that

$$\mathrm{P}\left(\theta \in (L, U)\right) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, U)$.

Parametric tests and confidence intervals – two-sided

Let $X_1, \ldots, X_n$ be a random sample from a distribution with a parameter θ.

We want to test the hypothesis (**two-sided alternative**):

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta \neq \theta_0,$$

for some fixed value θ_0.

Let further $\big(L(\mathbf{X}), U(\mathbf{X})\big)$ be the two-sided $100(1-\alpha)\%$ confidence interval for θ based on a random sample. Thus it holds that

$$\mathrm{P}\left(\theta \in (L, U)\right) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, U)$.
- We do not reject H_0 if $\theta_0 \in (L, U)$.

Parametric tests and confidence intervals – two-sided

We verify that this way we can test the hypothesis at given **significance level** α.

Parametric tests and confidence intervals – two-sided

We verify that this way we can test the hypothesis at given **significance level** α.

If the null hypothesis H_0 holds, i.e., $\theta = \theta_0$, for the **type I error** we have

$$\begin{aligned}\mathrm{P}(\text{reject } H_0 | H_0 \text{ holds}) &= \mathrm{P}\left(\theta_0 \notin (L,U) | \theta = \theta_0\right) = \mathrm{P}\left(\theta \notin (L,U)\right) \\ &= 1 - \mathrm{P}\left(\theta \in (L,U)\right) = 1 - (1-\alpha) = \alpha,\end{aligned}$$

because (L, U) is the $100(1-\alpha)\%$ confidence interval for θ.

The **level of significance** of our test is indeed α.

Parametric tests and confidence intervals – two-sided

We verify that this way we can test the hypothesis at given **significance level** α.

If the null hypothesis H_0 holds, i.e., $\theta = \theta_0$, for the **type I error** we have

$$\begin{aligned}\mathrm{P}(\text{reject } H_0 | H_0 \text{ holds}) &= \mathrm{P}\left(\theta_0 \notin (L,U) | \theta = \theta_0\right) = \mathrm{P}\left(\theta \notin (L,U)\right) \\ &= 1 - \mathrm{P}\left(\theta \in (L,U)\right) = 1 - (1-\alpha) = \alpha,\end{aligned}$$

because (L,U) is the $100(1-\alpha)\%$ confidence interval for θ.

The **level of significance** of our test is indeed α.

There are more possible decision rules. However, it can be shown (see literature) that for a general class of distributions, using the $(1-\alpha)$ confidence intervals, the probability of making type II error is lowest for any test with level of significance α. Therefore we can obtain the **most powerful test** against the given alternative.

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

For testing we use the one-sided interval of the type corresponding to the alternative hypothesis, i.e., upper $100(1-\alpha)\%$ confidence interval $(L, +\infty)$ for θ. It holds that

$$\mathrm{P}\left(\theta \in (L, +\infty)\right) = \mathrm{P}(\theta > L) = 1 - \alpha.$$

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

For testing we use the one-sided interval of the type corresponding to the alternative hypothesis, i.e., upper $100(1-\alpha)\%$ confidence interval $(L, +\infty)$ for θ. It holds that

$$\mathrm{P}\left(\theta \in (L, +\infty)\right) = \mathrm{P}(\theta > L) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, +\infty)$, i.e., $\theta_0 < L$.

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

For testing we use the one-sided interval of the type corresponding to the alternative hypothesis, i.e., upper $100(1-\alpha)\%$ confidence interval $(L, +\infty)$ for θ. It holds that

$$\mathrm{P}\left(\theta \in (L, +\infty)\right) = \mathrm{P}(\theta > L) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, +\infty)$, i.e., $\theta_0 < L$.
- We do not reject the hypothesis H_0 if $\theta_0 \in (L, +\infty)$, i.e., $\theta_0 \geq L$.

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

For testing we use the one-sided interval of the type corresponding to the alternative hypothesis, i.e., upper $100(1-\alpha)\%$ confidence interval $(L, +\infty)$ for θ. It holds that

$$\mathrm{P}\left(\theta \in (L, +\infty)\right) = \mathrm{P}(\theta > L) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, +\infty)$, i.e., $\theta_0 < L$.
- We do not reject the hypothesis H_0 if $\theta_0 \in (L, +\infty)$, i.e., $\theta_0 \geq L$.

Remarks:

- If θ_0 is outside the interval, the alternative H_A holds with a large probability.

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

For testing we use the one-sided interval of the type corresponding to the alternative hypothesis, i.e., upper $100(1-\alpha)\%$ confidence interval $(L, +\infty)$ for θ. It holds that

$$\mathrm{P}\left(\theta \in (L, +\infty)\right) = \mathrm{P}(\theta > L) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, +\infty)$, i.e., $\theta_0 < L$.
- We do not reject the hypothesis H_0 if $\theta_0 \in (L, +\infty)$, i.e., $\theta_0 \geq L$.

Remarks:

- If θ_0 is outside the interval, the alternative H_A holds with a large probability.
- The **level of significance** is again α.

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

For testing we use the one-sided interval of the type corresponding to the alternative hypothesis, i.e., upper $100(1-\alpha)\%$ confidence interval $(L, +\infty)$ for θ. It holds that

$$\mathrm{P}\left(\theta \in (L, +\infty)\right) = \mathrm{P}(\theta > L) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, +\infty)$, i.e., $\theta_0 < L$.
- We do not reject the hypothesis H_0 if $\theta_0 \in (L, +\infty)$, i.e., $\theta_0 \geq L$.

Remarks:

- If θ_0 is outside the interval, the alternative H_A holds with a large probability.
- The **level of significance** is again α.
- We proceed analogously for $H_A : \theta < \theta_0$.

Parametric tests and confidence intervals – one-sided

Now we want to test the hypothesis against a **one-sided alternative**

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

For testing we use the one-sided interval of the type corresponding to the alternative hypothesis, i.e., upper $100(1-\alpha)\%$ confidence interval $(L, +\infty)$ for θ. It holds that

$$\mathrm{P}\left(\theta \in (L, +\infty)\right) = \mathrm{P}(\theta > L) = 1 - \alpha.$$

We decide as follows:

- We reject the hypothesis H_0 if $\theta_0 \notin (L, +\infty)$, i.e., $\theta_0 < L$.
- We do not reject the hypothesis H_0 if $\theta_0 \in (L, +\infty)$, i.e., $\theta_0 \geq L$.

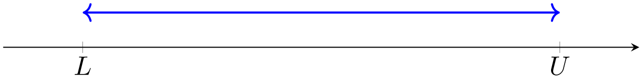
Remarks:

- If θ_0 is outside the interval, the alternative H_A holds with a large probability.
- The **level of significance** is again α.
- We proceed analogously for $H_A : \theta < \theta_0$.
- The null hypothesis can also be formulated in a compound form as:

$$H_0 : \theta \leq \theta_0 \qquad \text{against} \qquad H_A : \theta > \theta_0.$$

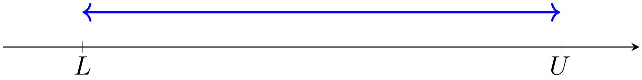
Parametric tests – illustration

Reject $H_0 : \theta = \theta_0$ in favor of the two-sided alternative $H_A : \theta \neq \theta_0$, if θ_0 does not lie in the **two-sided** confidence interval.

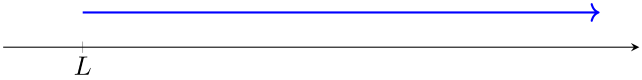


Parametric tests – illustration

Reject $H_0 : \theta = \theta_0$ in favor of the two-sided alternative $H_A : \theta \neq \theta_0$, if θ_0 does not lie in the **two-sided** confidence interval.

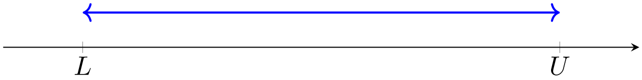


Reject $H_0 : \theta = \theta_0$ in favor of the one-sided alternative $H_A : \theta > \theta_0$, if θ_0 does not lie in the **upper one-sided** confidence interval.

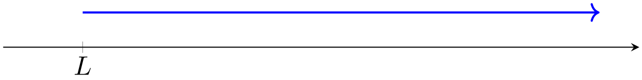


Parametric tests – illustration

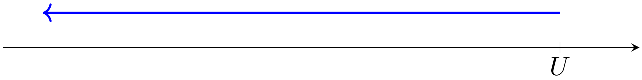
Reject $H_0 : \theta = \theta_0$ in favor of the two-sided alternative $H_A : \theta \neq \theta_0$, if θ_0 does not lie in the **two-sided** confidence interval.



Reject $H_0 : \theta = \theta_0$ in favor of the one-sided alternative $H_A : \theta > \theta_0$, if θ_0 does not lie in the **upper one-sided** confidence interval.



Reject $H_0 : \theta = \theta_0$ in favor of the one-sided alternative $H_A : \theta < \theta_0$, if θ_0 does not lie in the **lower one-sided** confidence interval.



Parametric tests – testing procedure

Testing procedure:

- Choose the level of significance α.

Parametric tests – testing procedure

Testing procedure:

- Choose the level of significance α.
- Measure (observe) the random sample.

Parametric tests – testing procedure

Testing procedure:

- Choose the level of significance α.
- Measure (observe) the random sample.
- Construct a $(1-\alpha)$ confidence interval corresponding to the alternative hypothesis H_A.

Parametric tests – testing procedure

Testing procedure:

- Choose the level of significance α.
- Measure (observe) the random sample.
- Construct a $(1-\alpha)$ confidence interval corresponding to the alternative hypothesis H_A.
- Reject H_0 if θ_0 is outside of the confidence interval.

Tests for the parameters of the normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Tests for the parameters of the normal distribution

Let $X_1, \ldots, X_n$ be a random sample from N(μ, σ^2).

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the level of significance α:

Tests for the parameters of the normal distribution

Let $X_1, \ldots, X_n$ be a random sample from N(μ, σ^2).

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the level of significance α:

- For known variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \, , \; \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right).$$

Tests for the parameters of the normal distribution

Let $X_1, \dots, X_n$ be a random sample from N(μ, σ^2).

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the level of significance α:

- For known variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \, , \, \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right).$$

- For unknown variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - t_{\alpha/2,n-1} \frac{s_n}{\sqrt{n}} \, , \, \bar{X}_n + t_{\alpha/2,n-1} \frac{s_n}{\sqrt{n}} \right).$$

Tests for the parameters of the normal distribution

Let $X_1, \ldots, X_n$ be a random sample from N(μ, σ^2).

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the level of significance α:

- For known variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\,,\ \bar{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

- For unknown variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - t_{\alpha/2,n-1}\frac{s_n}{\sqrt{n}}\,,\ \bar{X}_n + t_{\alpha/2,n-1}\frac{s_n}{\sqrt{n}}\right).$$

$H_0 : \sigma^2 = \sigma_0^2$ against the alternative $H_A : \sigma^2 \neq \sigma_0^2$ at the level of significance α:

Tests for the parameters of the normal distribution

Let $X_1, \dots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the level of significance α:

- For known variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \, , \, \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right).$$

- For unknown variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - t_{\alpha/2,n-1} \frac{s_n}{\sqrt{n}} \, , \, \bar{X}_n + t_{\alpha/2,n-1} \frac{s_n}{\sqrt{n}} \right).$$

$H_0 : \sigma^2 = \sigma_0^2$ against the alternative $H_A : \sigma^2 \neq \sigma_0^2$ at the level of significance α:

- We reject hypothesis H_0 if σ_0^2 is not in the interval

$$\left(\frac{(n-1)s_n^2}{\chi^2_{\alpha/2,n-1}} \, , \, \frac{(n-1)s_n^2}{\chi^2_{1-\alpha/2,n-1}} \right).$$

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from N(μ, σ^2).

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu > \mu_0$ at the level of significance α:

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from N(μ, σ^2).

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu > \mu_0$ at the level of significance α:

- For known variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - z_\alpha \frac{\sigma}{\sqrt{n}} \, , \, +\infty \right).$$

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu > \mu_0$ at the level of significance α:

- For known variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - z_\alpha \frac{\sigma}{\sqrt{n}} \, , \, +\infty \right).$$

- For unknown variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - t_{\alpha,n-1} \frac{s_n}{\sqrt{n}} \, , \, +\infty \right).$$

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

$H_0 : \mu = \mu_0$ against the alternative $H_A : \mu > \mu_0$ at the level of significance α:

- For known variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - z_\alpha \frac{\sigma}{\sqrt{n}} \, , \, +\infty \right).$$

- For unknown variance σ^2 we reject hypothesis H_0 if μ_0 is not in the interval

$$\left(\bar{X}_n - t_{\alpha,n-1} \frac{s_n}{\sqrt{n}} \, , \, +\infty \right).$$

$H_0 : \sigma^2 = \sigma_0^2$ against the alternative $H_A : \sigma^2 > \sigma_0^2$ at the level of significance α:

- We reject the hypothesis H_0 if σ_0^2 is not in the interval

$$\left(\frac{(n-1)s_n^2}{\chi^2_{\alpha,n-1}} \, , \, +\infty \right).$$

Tests for parameter of normal distribution – example

Example

We have $n = 35$ observations of random variable with distribution $\mu = \mathrm{E}\,X$:

$$\text{90\% interval } A: \quad (0.4055, 5.3945)$$

$$\text{95\% interval } B: \quad (-0.0724, 5.8724)$$

Test hypothesis

$$H_0 : \mu = 0 \qquad \textit{against} \qquad H_A : \mu > 0$$

at the significance level 5% ($\alpha = 0.05$) and 2.5% ($\alpha = 0.025$).

Tests for parameter of normal distribution – example

Example

We have $n = 35$ observations of random variable with distribution $\mu = \mathrm{E}\,X$:

$$90\%\text{ interval } A : \quad (0.4055, 5.3945)$$

$$95\%\text{ interval } B : \quad (-0.0724, 5.8724)$$

Test hypothesis

$$H_0 : \mu = 0 \qquad \textit{against} \qquad H_A : \mu > 0$$

at the significance level 5% ($\alpha = 0.05$) and 2.5% ($\alpha = 0.025$).

The needed one-sided confidence interval is

- 5% – $(0.4055, +\infty)$ and because $0 \notin (0.4055, +\infty)$ we reject the null hypothesis at significance level $\alpha = 5\%$

Tests for parameter of normal distribution – example

Example

We have $n = 35$ observations of random variable with distribution $\mu = \mathrm{E}\, X$:

$$\text{90\% interval } A: \quad (0.4055, 5.3945)$$

$$\text{95\% interval } B: \quad (-0.0724, 5.8724)$$

Test hypothesis

$$H_0 : \mu = 0 \qquad \textit{against} \qquad H_A : \mu > 0$$

at the significance level 5% ($\alpha = 0.05$) and 2.5% ($\alpha = 0.025$).

The needed one-sided confidence interval is

- 5% – $(0.4055, +\infty)$ and because $0 \notin (0.4055, +\infty)$ we reject the null hypothesis at significance level $\alpha = 5\%$
- 2.5% – $(-0.0724, +\infty)$ and because $0 \in (-0.0724, +\infty)$ we cannot reject the null hypothesis at significance level $\alpha = 2.5\%$.

Hypothesis testing – p-value

In bibliography you can encounter the **p-value** approach.

Given observed data, the null hypothesis can not be rejected on every significance level α.

Hypothesis testing – p-value

In bibliography you can encounter the **p-value** approach.

Given observed data, the null hypothesis can not be rejected on every significance level α.

The minimal significance level at which we can reject hypothesis H_0 given the data at hand is called the **p-value**.

Hypothesis testing – p-value

In bibliography you can encounter the **p-value** approach.

Given observed data, the null hypothesis can not be rejected on every significance level α.

The minimal significance level at which we can reject hypothesis H_0 given the data at hand is called the **p-value**. The p-value depends on the random sample realization.

Hypothesis testing – p-value

In bibliography you can encounter the **p-value** approach.

Given observed data, the null hypothesis can not be rejected on every significance level α.

The minimal significance level at which we can reject hypothesis H_0 given the data at hand is called the **p-value**. The p-value depends on the random sample realization.

Meaning of the p-value

- Many statistical softwares give only the p-value as the output of a hypothesis test.
- If the p-value is smaller than our required significance level α we reject H_0.
- The size of the p-value informs us how strong is the rejection of H_0 is, or how weak the non-rejection.
- The smaller the p-value is, the more significant is the rejection of H_0.

Hypothesis testing – test statistics approach

We can use an other approach for hypotheses testing, based on comparing the tested value with its point estimate.

Hypothesis testing – test statistics approach

We can use an other approach for hypotheses testing, based on comparing the tested value with its point estimate.

For example, when observing a random sample from N(μ, σ^2), we could reject the hypothesis $H_0 : \mu = \mu_0$ if the sample mean $\bar{X}_n$ and the tested value μ_0 are far away from each other.

Hypothesis testing – test statistics approach

We can use an other approach for hypotheses testing, based on comparing the tested value with its point estimate.

For example, when observing a random sample from N(μ, σ^2), we could reject the hypothesis $H_0 : \mu = \mu_0$ if the sample mean $\bar{X}_n$ and the tested value μ_0 are far away from each other.

How far?

Hypothesis testing – test statistics approach

We can use an other approach for hypotheses testing, based on comparing the tested value with its point estimate.

For example, when observing a random sample from $\mathsf{N}(\mu, \sigma^2)$, we could reject the hypothesis $H_0 : \mu = \mu_0$ if the sample mean $\bar{X}_n$ and the tested value μ_0 are far away from each other.

How far?

We know that if H_0 holds and σ^2 is known, it holds that

$$\frac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n} \sim \mathsf{N}(0,1).$$

Therefore with a large probability $1 - \alpha$, the standardised distance should be within bounds given by the critical values of $\mathsf{N}(0,1)$.

Hypothesis testing – test statistics approach

We can use an other approach for hypotheses testing, based on comparing the tested value with its point estimate.

For example, when observing a random sample from $\mathsf{N}(\mu, \sigma^2)$, we could reject the hypothesis $H_0 : \mu = \mu_0$ if the sample mean $\bar{X}_n$ and the tested value μ_0 are far away from each other.

How far?

We know that if H_0 holds and σ^2 is known, it holds that

$$\frac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n} \sim \mathsf{N}(0, 1).$$

Therefore with a large probability $1 - \alpha$, the standardised distance should be within bounds given by the critical values of $\mathsf{N}(0, 1)$. We can thus reject H_0 in favor of $H_A : \mu \neq \mu_0$ if

$$\left| \frac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n} \right| > z_{\alpha/2}.$$

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Test $H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the significance level α:

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Test $H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the significance level α:

If the variance σ^2 is known, we can construct a **test statistic** $T = T(X_1, \ldots, X_n)$ as

$$T = \frac{\bar{X}_n - \mu_0}{\sigma} \sqrt{n}.$$

Parametric tests – normal distribution

Let $X_1, \dots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Test $H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the significance level α:
If the variance σ^2 is known, we can construct a **test statistic** $T = T(X_1, \dots, X_n)$ as

$$T = \frac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n}.$$

We reject H_0 in favor of H_A if the test statistic lies in the **critical region**:

$$W_\alpha = \{T : |T| > z_{\alpha/2}\}.$$

Parametric tests – normal distribution

Let $X_1, \dots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Test $H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the significance level α:
If the variance σ^2 is known, we can construct a **test statistic** $T = T(X_1, \dots, X_n)$ as

$$T = \frac{\bar{X}_n - \mu_0}{\sigma} \sqrt{n}.$$

We reject H_0 in favor of H_A if the test statistic lies in the **critical region**:

$$W_\alpha = \{T : |T| > z_{\alpha/2}\}.$$

We reject the hypothesis H_0 if $T \in W_\alpha$, meaning that $|T|$ is large enough and thus $\bar{X}_n$ is too far from μ_0. Alternatively, we do not reject if it holds that

$$-z_{\alpha/2} < \frac{\bar{X}_n - \mu_0}{\sigma} \sqrt{n} < z_{\alpha/2}.$$

Parametric tests – normal distribution

Let $X_1, \dots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Test $H_0 : \mu = \mu_0$ against the alternative $H_A : \mu \neq \mu_0$ at the significance level α:
If the variance σ^2 is known, we can construct a **test statistic** $T = T(X_1, \dots, X_n)$ as

$$T = \frac{\bar{X}_n - \mu_0}{\sigma} \sqrt{n}.$$

We reject H_0 in favor of H_A if the test statistic lies in the **critical region**:

$$W_\alpha = \{T : |T| > z_{\alpha/2}\}.$$

We reject the hypothesis H_0 if $T \in W_\alpha$, meaning that $|T|$ is large enough and thus $\bar{X}_n$ is too far from μ_0. Alternatively, we do not reject if it holds that

$$-z_{\alpha/2} < \frac{\bar{X}_n - \mu_0}{\sigma} \sqrt{n} < z_{\alpha/2}.$$

After separating μ_0 we obtain the same interval as the corresponding $(1 - \alpha)$ confidence interval.

Hypothesis testing – test statistics approach

The approach based on the construction of the test statistic T and the corresponding critical region of the test W_α can be summarized as follows:

Hypothesis testing – test statistics approach

The approach based on the construction of the test statistic T and the corresponding critical region of the test W_α can be summarized as follows:

Testing procedure

- Choose the level of significance α.

Hypothesis testing – test statistics approach

The approach based on the construction of the test statistic T and the corresponding critical region of the test W_α can be summarized as follows:

Testing procedure

- Choose the level of significance α.
- Measure (observe) a random sample.

Hypothesis testing – test statistics approach

The approach based on the construction of the test statistic T and the corresponding critical region of the test W_α can be summarized as follows:

Testing procedure

- Choose the level of significance α.
- Measure (observe) a random sample.
- Compute the test statistic T.

Hypothesis testing – test statistics approach

The approach based on the construction of the test statistic T and the corresponding critical region of the test W_α can be summarized as follows:

Testing procedure

- Choose the level of significance α.
- Measure (observe) a random sample.
- Compute the test statistic T.
- Find the corresponding critical region based on the outlying parts of the distribution of T.

Hypothesis testing – test statistics approach

The approach based on the construction of the test statistic T and the corresponding critical region of the test W_α can be summarized as follows:

Testing procedure

- Choose the level of significance α.
- Measure (observe) a random sample.
- Compute the test statistic T.
- Find the corresponding critical region based on the outlying parts of the distribution of T.
- Reject H_0 if $T \in W_\alpha$.

Hypothesis testing – test statistics approach

The approach based on the construction of the test statistic T and the corresponding critical region of the test W_α can be summarized as follows:

Testing procedure

- Choose the level of significance α.
- Measure (observe) a random sample.
- Compute the test statistic T.
- Find the corresponding critical region based on the outlying parts of the distribution of T.
- Reject H_0 if $T \in W_\alpha$.

✓ The critical region can be often converted to the corresponding confidence interval.

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Tests for the expectation at significance level α:

Parametric tests – normal distribution

Let $X_1, \ldots, X_n$ be a random sample from N(μ, σ^2).

Tests for the expectation at significance level α:

- For a known variance σ^2 the test statistic and critical regions are following:

H_0	H_A	test statistic T	critical region W_α
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \frac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n}$	$\lvert T \rvert > z_{\alpha/2}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T > z_\alpha$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -z_\alpha$

Parametric tests – normal distribution

Let $X_1, \dots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Tests for the expectation at significance level α:

- For a known variance σ^2 the test statistic and critical regions are following:

H_0	H_A	test statistic T	critical region W_α
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n}$	$\lvert T\rvert > z_{\alpha/2}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T > z_\alpha$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -z_\alpha$

- For unknown variance σ^2 the test statistic and critical regions are following:

H_0	H_A	test statistic T	critical region W_α
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\bar{X}_n - \mu_0}{s_n}\sqrt{n}$	$\lvert T\rvert > t_{\alpha/2,n-1}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T > t_{\alpha,n-1}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -t_{\alpha,n-1}$

Parametric tests – normal distribution

Let $X_1, \dots, X_n$ be a random sample from N(μ, σ^2).

Tests for the expectation at significance level α:

- For a known variance σ^2 the test statistic and critical regions are following:

H_0	H_A	test statistic T	critical region W_α
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n}$	$\lvert T\rvert > z_{\alpha/2}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T > z_\alpha$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -z_\alpha$

- For unknown variance σ^2 the test statistic and critical regions are following:

H_0	H_A	test statistic T	critical region W_α
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\bar{X}_n - \mu_0}{s_n}\sqrt{n}$	$\lvert T\rvert > t_{\alpha/2,n-1}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T > t_{\alpha,n-1}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -t_{\alpha,n-1}$

Tests for the variance at significance level α:

Parametric tests – normal distribution

Let $X_1, \dots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$.

Tests for the expectation at significance level α:

- For a known variance σ^2 the test statistic and critical regions are following:

H_0	H_A	test statistic T	critical region W_α
$\mu = \mu_0$	$\mu \neq \mu_0$		$\lvert T \rvert > z_{\alpha/2}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$T = \dfrac{\bar{X}_n - \mu_0}{\sigma}\sqrt{n}$	$T > z_\alpha$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -z_\alpha$

- For unknown variance σ^2 the test statistic and critical regions are following:

H_0	H_A	test statistic T	critical region W_α
$\mu = \mu_0$	$\mu \neq \mu_0$		$\lvert T \rvert > t_{\alpha/2,n-1}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$T = \dfrac{\bar{X}_n - \mu_0}{s_n}\sqrt{n}$	$T > t_{\alpha,n-1}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -t_{\alpha,n-1}$

Tests for the variance at significance level α:

H_0	H_A	test statistic T	critical region W_α
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$		$T > \chi^2_{\alpha/2,n-1} \vee T < \chi_{1-\alpha/2,n-1}$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$T = \dfrac{(n-1)s_n^2}{\sigma_0^2}$	$T > \chi^2_{\alpha,n-1}$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$		$T < \chi^2_{1-\alpha,n-1}$

Paired t-test

Suppose we observe a random sample of pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$. The variables within pairs can be dependent, but the pairs are independent between each other.

Paired t-test

Suppose we observe a random sample of pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$. The variables within pairs can be dependent, but the pairs are independent between each other.

Such a situation can describe the value of a certain marker measured on patients before and after a clinical procedure. We want to determine, whether the marker stayed the same, or if it has significantly increased or decreased after the procedure.

Paired t-test

Suppose we observe a random sample of pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$. The variables within pairs can be dependent, but the pairs are independent between each other.

Such a situation can describe the value of a certain marker measured on patients before and after a clinical procedure. We want to determine, whether the marker stayed the same, or if it has significantly increased or decreased after the procedure.

Suppose that all variables are normally distributed with $X_i \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathrm{N}(\mu_2, \sigma_2^2)$. We want to test $H_0 : \mu_1 = \mu_2$.

Paired t-test

Suppose we observe a random sample of pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$. The variables within pairs can be dependent, but the pairs are independent between each other.

Such a situation can describe the value of a certain marker measured on patients before and after a clinical procedure. We want to determine, whether the marker stayed the same, or if it has significantly increased or decreased after the procedure.

Suppose that all variables are normally distributed with $X_i \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathrm{N}(\mu_2, \sigma_2^2)$. We want to test $H_0 : \mu_1 = \mu_2$.

When we take $Z_i = Y_i - X_i$, the resulting variables will have the normal distribution with expectation of $\mu_{\text{diff}} = \mu_2 - \mu_1$.

Paired t-test

Suppose we observe a random sample of pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$. The variables within pairs can be dependent, but the pairs are independent between each other.

Such a situation can describe the value of a certain marker measured on patients before and after a clinical procedure. We want to determine, whether the marker stayed the same, or if it has significantly increased or decreased after the procedure.

Suppose that all variables are normally distributed with $X_i \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathrm{N}(\mu_2, \sigma_2^2)$. We want to test $H_0 : \mu_1 = \mu_2$.

When we take $Z_i = Y_i - X_i$, the resulting variables will have the normal distribution with expectation of $\mu_{\text{diff}} = \mu_2 - \mu_1$.

The test can then be performed in the same way as for a single sample from a normal distribution, testing $H_0 : \mu_{\text{diff}} = 0$ against $H_A : \mu_{\text{diff}} \neq 0$.

Paired t-test

Suppose we observe a random sample of pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$. The variables within pairs can be dependent, but the pairs are independent between each other.

Such a situation can describe the value of a certain marker measured on patients before and after a clinical procedure. We want to determine, whether the marker stayed the same, or if it has significantly increased or decreased after the procedure.

Suppose that all variables are normally distributed with $X_i \sim \mathsf{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathsf{N}(\mu_2, \sigma_2^2)$. We want to test $H_0 : \mu_1 = \mu_2$.

When we take $Z_i = Y_i - X_i$, the resulting variables will have the normal distribution with expectation of $\mu_{\text{diff}} = \mu_2 - \mu_1$.

The test can then be performed in the same way as for a single sample from a normal distribution, testing $H_0 : \mu_{\text{diff}} = 0$ against $H_A : \mu_{\text{diff}} \neq 0$.

Similarly for one-sided alternatives.

Paired t-test – example

Example – comparing fathers' and sons' heights

Suppose we want to determine whether the men's height increases between generations. We have observed five pairs of fathers and their sons, now adults. Their height was measured as follows (in centimeters):

height of father	X_i	172	176	180	184	186
height of son	Y_i	178	188	177	192	193
difference	$Z_i = Y_i - X_i$	6	12	-3	8	7

We test whether the expected sons' height is equal to the expected fathers' height, against the alternative that sons are significantly taller, using $\alpha = 5\%$.

Paired t-test – example

Example – comparing fathers' and sons' heights

Suppose we want to determine whether the men's height increases between generations. We have observed five pairs of fathers and their sons, now adults. Their height was measured as follows (in centimeters):

height of father	X_i	172	176	180	184	186
height of son	Y_i	178	188	177	192	193
difference	$Z_i = Y_i - X_i$	6	12	-3	8	7

We test whether the expected sons' height is equal to the expected fathers' height, against the alternative that sons are significantly taller, using $\alpha = 5\%$.

The upper one-sided 95% confidence interval for the expectation μ_{diff} of Z_i is

$$\left(\bar{Z}_n - t_{\alpha,n-1}\frac{s_Z}{\sqrt{n}}\,,\ +\infty\right) = \left(6 - 2.132\cdot\frac{5.52}{\sqrt{5}}\,,\ +\infty\right) = (0.735, +\infty).$$

The tested value $\mu_{\text{diff}} = 0$ does not lie in the interval, so we can reject the hypothesis in favor of the alternative that the sons are significantly taller than their fathers.
The test statistic and the p-value can be obtained in R using:

```
t.test(height_son,height_father,paired=T,alternative="greater")
```

Two sample t-test

Suppose we observe a random sample of $X_1, \ldots, X_{n_1}$ and an independent sample of $Y_1, \ldots, Y_{n_2}$.

Two sample t-test

Suppose we observe a random sample of $X_1, \ldots, X_{n_1}$ and an independent sample of $Y_1, \ldots, Y_{n_2}$.

Such a situation can describe the value of a certain marker measured on two independent groups of patients, each undergoing a different treatment. We want to determine whether the marker is equal for both treatment groups, or whether it differs significantly.

Two sample t-test

Suppose we observe a random sample of $X_1, \ldots, X_{n_1}$ and an independent sample of $Y_1, \ldots, Y_{n_2}$.

Such a situation can describe the value of a certain marker measured on two independent groups of patients, each undergoing a different treatment. We want to determine whether the marker is equal for both treatment groups, or whether it differs significantly.

Suppose that all variables are normally distributed with $X_i \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathrm{N}(\mu_2, \sigma_2^2)$. We want to test $H_0 : \mu_1 = \mu_2$.

Two sample t-test

Suppose we observe a random sample of $X_1, \dots, X_{n_1}$ and an independent sample of $Y_1, \dots, Y_{n_2}$.

Such a situation can describe the value of a certain marker measured on two independent groups of patients, each undergoing a different treatment. We want to determine whether the marker is equal for both treatment groups, or whether it differs significantly.

Suppose that all variables are normally distributed with $X_i \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathrm{N}(\mu_2, \sigma_2^2)$. We want to test $H_0 : \mu_1 = \mu_2$.

It can be shown that if H_0 holds, the statistic

$$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_\bullet}$$

follows the Student's t-distribution. The sample standard deviation $s_\bullet$ and the number of degrees of freedom depend on whether the samples have equal variances ($\sigma_1^2 = \sigma_2^2$) or not.

Two sample t-test

Suppose we observe a random sample of $X_1, \dots, X_{n_1}$ and an independent sample of $Y_1, \dots, Y_{n_2}$.

Such a situation can describe the value of a certain marker measured on two independent groups of patients, each undergoing a different treatment. We want to determine whether the marker is equal for both treatment groups, or whether it differs significantly.

Suppose that all variables are normally distributed with $X_i \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathrm{N}(\mu_2, \sigma_2^2)$. We want to test $H_0 : \mu_1 = \mu_2$.

It can be shown that if H_0 holds, the statistic

$$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_\bullet}$$

follows the Student's t-distribution. The sample standard deviation $s_\bullet$ and the number of degrees of freedom depend on whether the samples have equal variances ($\sigma_1^2 = \sigma_2^2$) or not.

The test can then be performed by comparing the test statistic T with the corresponding critical values of the t-distribution.

Two-sample tests – normal distribution

Let $X_1, \ldots, X_{n_1}$ be a random sample from N(μ_1, σ_1^2) and $Y_1, \ldots, Y_{n_2}$ be a random sample from N(μ_2, σ_2^2).

Two-sample tests – normal distribution

Let $X_1, \ldots, X_{n_1}$ be a random sample from N(μ_1, σ_1^2) and $Y_1, \ldots, Y_{n_2}$ be a random sample from N(μ_2, σ_2^2).

Tests for the equality of expectations under $\sigma_1^2 = \sigma_2^2$:

H_0	H_A	test statistic T	critical region W_α
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$T = \dfrac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_{12}} \sqrt{\dfrac{n_1 n_2}{n_1 + n_2}}$	$\lvert T \rvert > t_{\alpha/2, n_1+n_2-2}$
$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$		$T > t_{\alpha, n_1+n_2-2}$
$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$		$T < -t_{\alpha, n_1+n_2-2}$

- Where $s_{12} = \sqrt{\dfrac{(n_1-1)s_X^2 + (n_2-1)s_Y^2}{n_1+n_2-2}}$,
- t_{α, n_1+n_2-2} is the critical value of Student's t-distribution with $n_1 + n_2 - 2$ degrees of freedom.

Two-sample tests – normal distribution

Let $X_1, \ldots, X_{n_1}$ be a random sample from N(μ_1, σ_1^2) and $Y_1, \ldots, Y_{n_2}$ be a random sample from N(μ_2, σ_2^2).

Two-sample tests – normal distribution

Let $X_1, \ldots, X_{n_1}$ be a random sample from N(μ_1, σ_1^2) and $Y_1, \ldots, Y_{n_2}$ be a random sample from N(μ_2, σ_2^2).

Tests for the equality of expectations under $\sigma_1^2 \neq \sigma_2^2$:

H_0	H_A	test statistic T	critical region W_α
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_d}$	$\lvert T\rvert > t_{\alpha/2, n_d}$
$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$		$T > t_{\alpha, n_d}$
$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$		$T < -t_{\alpha, n_d}$

- Where $s_d = \sqrt{\frac{s_X^2}{n_1} + \frac{s_Y^2}{n_2}}$,
- $n_d = \dfrac{s_d^4}{\frac{1}{n_1 - 1}\left(\frac{s_X^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_Y^2}{n_2}\right)^2}$

Two sample t-test – example

Example – comparing men's heights from different countries

Suppose we want to determine whether the average men's height is the same in the Czech Republic and in Norway. We have observed five men from CZE and six men from NOR. Their heights were measured as follows (in centimeters):

height CZE	X_i	169	178	179	186	191	
height NOR	Y_i	175	182	183	189	191	192

We test whether the expected heights are equal, against the alternative that they are not, on $\alpha = 5\%$. We take the variances as equal.

Two sample t-test – example

Example – comparing men's heights from different countries

Suppose we want to determine whether the average men's height is the same in the Czech Republic and in Norway. We have observed five men from CZE and six men from NOR. Their heights were measured as follows (in centimeters):

height CZE	X_i	169	178	179	186	191	
height NOR	Y_i	175	182	183	189	191	192

We test whether the expected heights are equal, against the alternative that they are not, on $\alpha = 5\%$. We take the variances as equal. The test statistic using equal variances is

$$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_{12}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = -1.0545.$$

Two sample t-test – example

Example – comparing men's heights from different countries

Suppose we want to determine whether the average men's height is the same in the Czech Republic and in Norway. We have observed five men from CZE and six men from NOR. Their heights were measured as follows (in centimeters):

height CZE	X_i	169	178	179	186	191	
height NOR	Y_i	175	182	183	189	191	192

We test whether the expected heights are equal, against the alternative that they are not, on $\alpha = 5\%$. We take the variances as equal. The test statistic using equal variances is

$$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_{12}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = -1.0545.$$

The $\alpha/2$ critical value of the Student's t-distribution with $n_1 + n_2 - 2$ degrees of freedom is $t_{\alpha/2, n_1+n_2-2} = t_{0.025,9} = 2.262.$

Two sample t-test – example

Example – comparing men's heights from different countries

Suppose we want to determine whether the average men's height is the same in the Czech Republic and in Norway. We have observed five men from CZE and six men from NOR. Their heights were measured as follows (in centimeters):

height CZE	X_i	169	178	179	186	191	
height NOR	Y_i	175	182	183	189	191	192

We test whether the expected heights are equal, against the alternative that they are not, on $\alpha = 5\%$. We take the variances as equal. The test statistic using equal variances is

$$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_{12}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = -1.0545.$$

The $\alpha/2$ critical value of the Student's t-distribution with $n_1 + n_2 - 2$ degrees of freedom is $t_{\alpha/2, n_1+n_2-2} = t_{0.025,9} = 2.262.$

Since

$$1.0545 = |T| < t_{\alpha/2, n_1+n_2-2} = 2.262,$$

we do not reject the null hypothesis of equality. Based on our data we could not find a significant difference between the expected heights of men among the two countries.

Two sample t-test – example

Example – comparing men's heights from different countries

Suppose we want to determine whether the average men's height is the same in the Czech Republic and in Norway. We have observed five men from CZE and six men from NOR. Their heights were measured as follows (in centimeters):

height CZE	X_i	169	178	179	186	191	
height NOR	Y_i	175	182	183	189	191	192

We test whether the expected heights are equal, against the alternative that they are not, on $\alpha = 5\%$. We take the variances as equal. The test statistic using equal variances is

$$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s_{12}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = -1.0545.$$

The $\alpha/2$ critical value of the Student's t-distribution with $n_1 + n_2 - 2$ degrees of freedom is $t_{\alpha/2, n_1+n_2-2} = t_{0.025,9} = 2.262$.

Since

$$1.0545 = |T| < t_{\alpha/2, n_1+n_2-2} = 2.262,$$

we do not reject the null hypothesis of equality. Based on our data we could not find a significant difference between the expeced heights of men among the two countries.

The test statistic and the p-value can be obtained in R using:

```
t.test(height_cze,height_nor,paired=F,alternative="two.sided")
```

Two-sample tests – normal distribution

Let $X_1, \dots, X_{n_1}$ be a random sample from $\mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_1, \dots, Y_{n_2}$ be a random sample from $\mathrm{N}(\mu_2, \sigma_2^2)$.

Two-sample tests – normal distribution

Let $X_1, \ldots, X_{n_1}$ be a random sample from $\mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_1, \ldots, Y_{n_2}$ be a random sample from $\mathrm{N}(\mu_2, \sigma_2^2)$.

Tests for the equality of variances – F-test:

H_0	H_A	test statistic T	critical region W_α
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$		$T < F_{1-\alpha/2,n_1-1,n_2-1} \vee T > F_{\alpha/2,n_1-1,n_2-1}$
$\sigma_1^2 \leq \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$T = \frac{s_X^2}{s_Y^2}$	$T > F_{\alpha,n_1-1,n_2-1}$
$\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$		$T < F_{1-\alpha,n_1-1,n_2-1}$

- s_X^2 is the sample variance of the first random sample and s_Y^2 is the sample variance of the second sample.
- F_{α,n_1-1,n_2-1} is the critical value of the *Fisher-Snedecor F-distribution* with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.
- Important note: The F-test is particularly sensitive to the normality of X and Y. If we are not sure whether the data is normally distributed, it is better to use a different test or assume non-equal variances for the t-test.
- The test can be called in R using `var.test(height_cze,height_nor)`.

Recap

We want to test a **hypothesis** concerning a parameter of a distribution.

We test the **null hypothesis** H_0 against an **alternative hypothesis** H_A.

We set up a decision rule, which, based on the observed data, either **rejects** or **does not reject** H_0 in favor of H_A.

We don't want the probability of falsely rejecting H_0 to exceed a chosen **level of significance** α.

We can establish the decision rule based on the confidence intervals:

- Reject $H_0 : \theta = \theta_0$ in favor of $H_A : \theta \neq \theta_0$, if θ_0 does not lie in the $1 - \alpha$ two-sided confidence interval.
- Reject $H_0 : \theta = \theta_0$ in favor of $H_A : \theta > \theta_0$, if θ_0 does not lie in the $1 - \alpha$ one-sided upper confidence interval.
- Reject $H_0 : \theta = \theta_0$ in favor of $H_A : \theta < \theta_0$, if θ_0 does not lie in the $1 - \alpha$ one-sided lower confidence interval.

Based on the **test statistics** approach we can also solve paired and two-sample problems.