Biostatistics I: Hypothesis testing

Introduction

Eleni-Rosalina Andrinopoulou

Department of Biostatistics, Erasmus Medical Center

e.andrinopoulou@erasmusmc.nl

@erandrinopoulou

Erasmus MC
University Medical Center Rotterdam

In this Section

- Types of tests
- General procedure
- Example

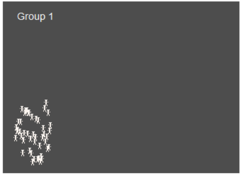
Types of tests

- **parametric** (assumptions about the distribution) / **non-parametric** (distribution-free)

Types of tests

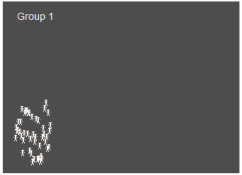
- **parametric** (assumptions about the distribution) / **non-parametric** (distribution-free)

- **one sample** / **two samples** / .. / *M* **samples**
  - compare one group with a value
  - compare two groups paired / unpaired
  - compare *M* groups



Types of tests

- **parametric** (assumptions about the distribution) / **non-parametric** (distribution-free)
- **one sample** / **two samples** / .. / M **samples**
  - compare one group with a value
  - compare two groups paired / unpaired
  - compare M groups



- **one-sided (one-tailed)** / **two-sided (two-tailed)**

$H_0 : \theta = \theta_0$

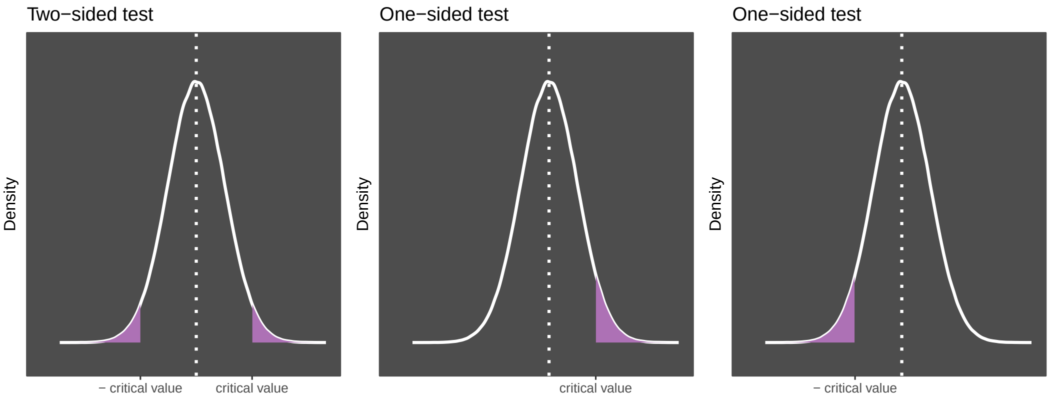
$H_1 : \theta \neq \theta_0$ (two-sided)
$H_1 : \theta > \theta_0$ (one-sided)
$H_1 : \theta < \theta_0$ (one-sided)

General procedure

- Choose a null hypothesis H_0 and an alternative hypothesis H_1
- Collect and visualize the data
- Choose and calculate the test statistic, which is a numerical summary of the data
- Determine the sampling distribution under the condition that the null-hypothesis holds
- Choose the type I error (significant level) α, usually α=0.05
- Determine the corresponding critical value(s)
- Compare the test statistic with critical value(s) and reject or not

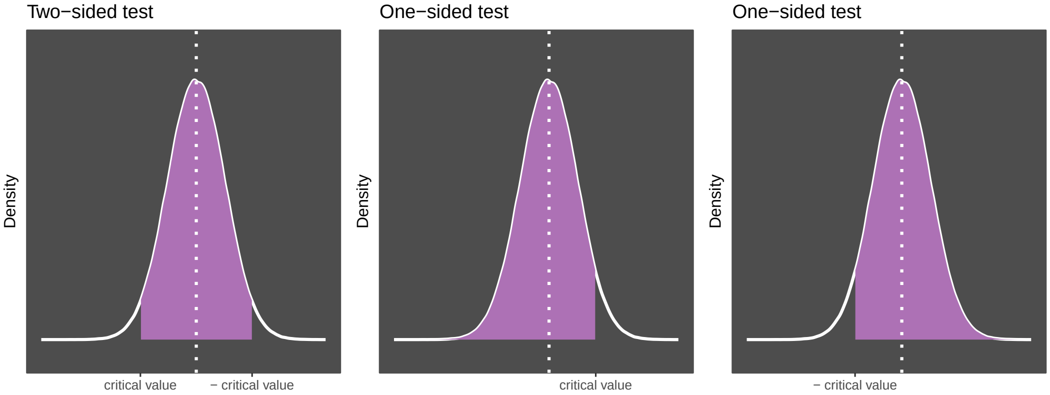
Draw conclusions

The purple area represents the rejection region of H_0



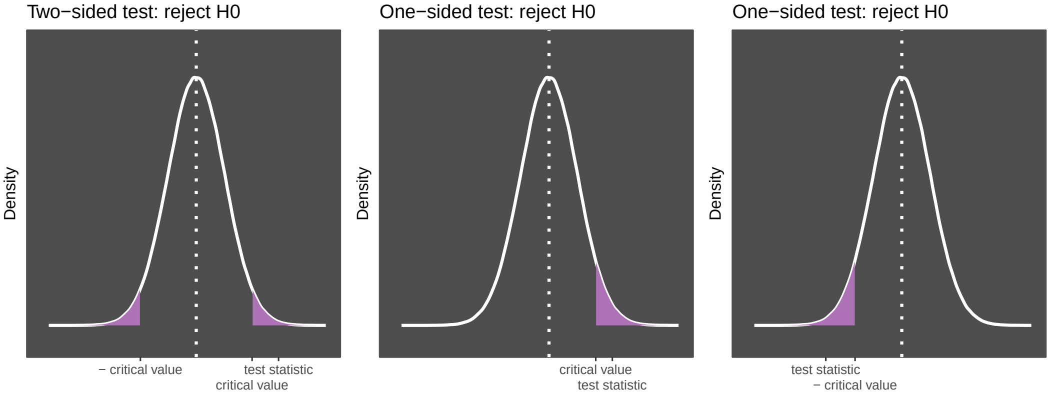
Draw conclusions

The purple area represents the non rejection region of H_0



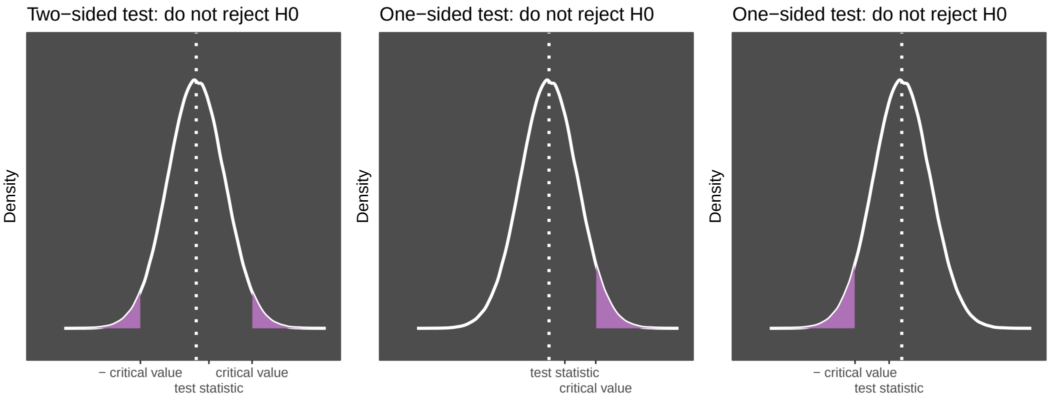
Draw conclusions

For example:



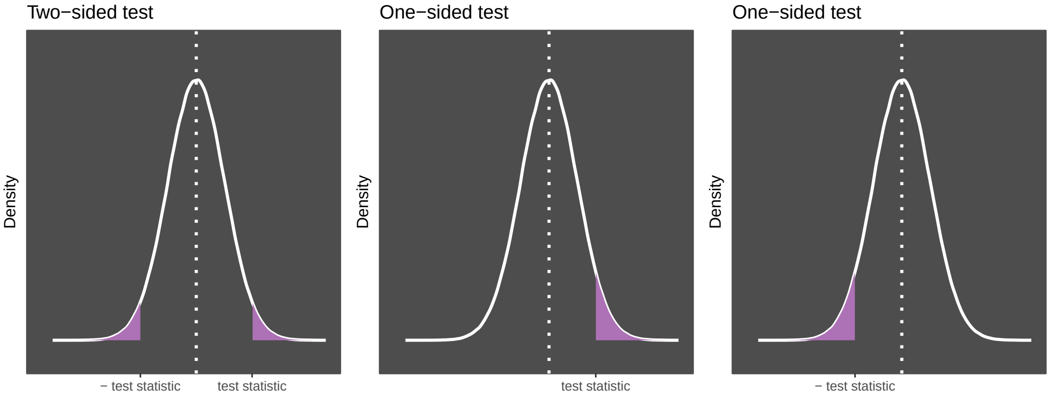
Draw conclusions

For example:



Draw conclusions

Translate that into probabilities: the p-value (purple area) is the probability of obtaining test results at least as extreme as the results observed in the original sample, under the assumption that the null hypothesis is correct



Confidence interval

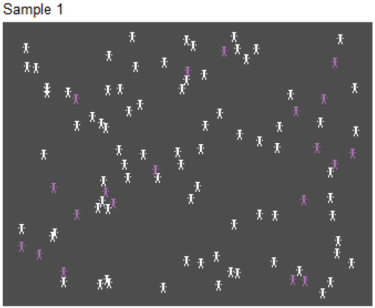
- Always present both the value and precision of the parameter of interest
- Statistically significant effect $\neq$ practically significant effect
- Confidence interval and p-value/critical value(s) always agree

Confidence level = 1 – significance level (α)

Example

We make the hypothesis that the mean height of Dutch women older than 18 is larger than 1.7 cm

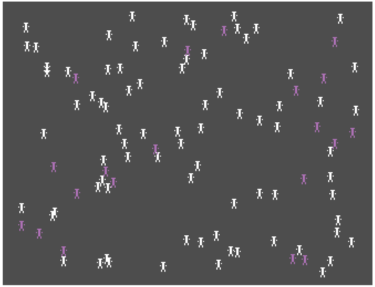
- Collect data! Select one representative sample



Example

The distribution of the statistic of those different samples is called sampling distribution

Sample 1



Example

This is a one sample, one-tailed test. We have:

- $H_0 : \mu = 1.7$
- $H_1 : \mu > 1.7$

where
μ is the mean height of all Dutch women older than 18

The test statistic is $z = \frac{\bar{x} - \mu_0}{sd(x)/\sqrt{n}}$,

μ_0 is the estimate under the null hypothesis

A test statistic is used to determine whether to reject (or not) the H_0

Example

Let's assume that:

- $n = 500$
- $\bar{x} = 1.79$
- $sd(x) = 1.1$

The test statistic will be $z = \frac{\bar{x} - \mu_0}{sd(x)/\sqrt{n}} = \frac{1.79 - 1.7}{1.1/\sqrt{500}} = 1.83$

Example

The sampling distribution depends on

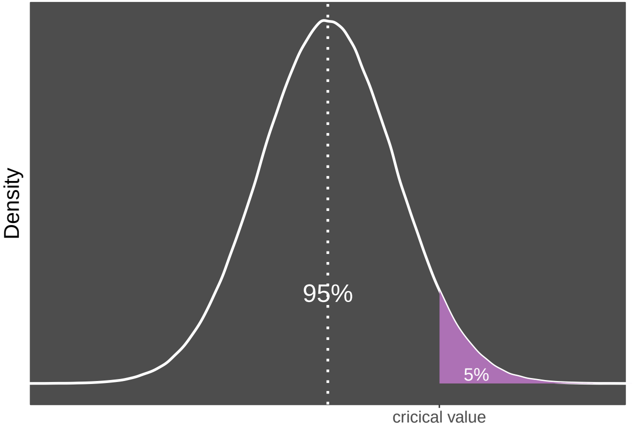
- the underlying distribution of the population
- the statistic
- the sampling procedure
- the sample size

The sampling distribution will be the normal distribution

Test statistic VS critical values from sampling distribution: how likely it is that we would get that statistic if we were sampling from a population that has the null hypothesis characteristics

Example

We choose the type I error (α) to be 0.05. The critical value$_\alpha$ is then obtained from the standard normal distribution and compared to the test statistic

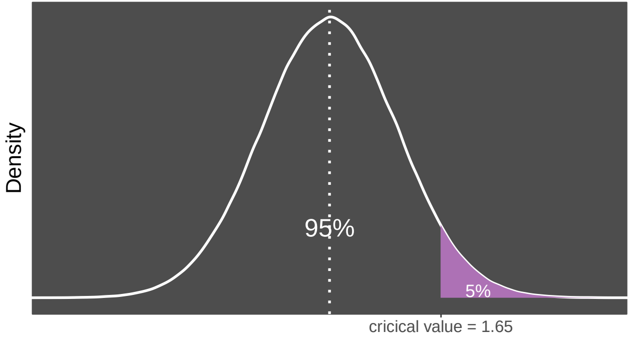


Example

Using R we get the critical values from standard normal distribution:
critical value$_\alpha$ = critical value$_{0.05}$

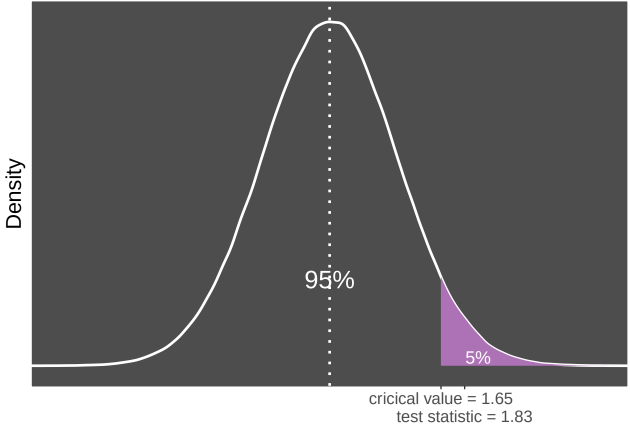
```r
qnorm(p = 0.05, lower.tail = FALSE)
```

```
[1] 1.644854
```



Example

test statistic > critical value$_{\alpha}$ $\Rightarrow$ reject the H_0

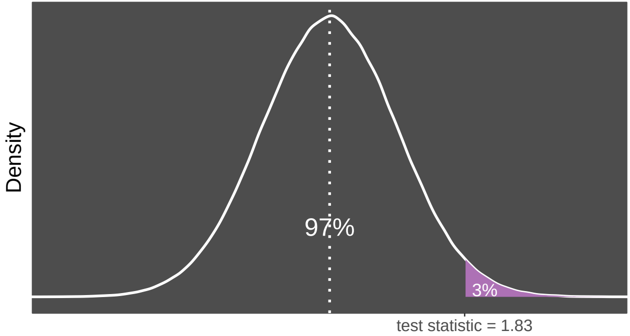


Example

The p-value is:

```
pnorm(q = 1.83, lower.tail = FALSE)
```

```
[1] 0.03362497
```



Example

What about the confidence interval?

Range of values for the unknown parameter

- Select the confidence level: represents the probability that the estimated interval will contain the true value of the parameter
- z follows the standard normal distribution

$$Pr(-1.96 \leq z \leq 1.96) = 0.95 \Rightarrow$$

$$Pr(-1.96 \leq \frac{\bar{x} - \mu_0}{sd(x)/\sqrt{n}} \leq 1.96) = 0.95 \Rightarrow$$

$$Pr(-1.96 * sd(x)/\sqrt{n} \leq \bar{x} - \mu_0 \leq 1.96 * sd(x)/\sqrt{n}) = 0.95 \Rightarrow$$

$$Pr(\bar{x} - 1.96 * sd(x)/\sqrt{n} \leq \mu_0 \leq \bar{x} + 1.96 * sd(x)/\sqrt{n}) = 0.95$$

Example

In our data:

95% CI: $[1.79 - 1.96 * 1.1/\sqrt{500}, 1.79 + 1.96 * 1.1/\sqrt{500}] \Rightarrow$
95% CI: $[1.69, 1.89]$

- We chose the confidence level to be 0.95
- We could, however, assume a 99% confidence interval