

Biostatistics I: Statistical tests for continuous data

Eleni-Rosalina Andrinopoulou

Department of Biostatistics, Erasmus Medical Center

✉ e.andrinopoulou@erasmusmc.nl

🐦 [@erandrinopoulou](https://twitter.com/erandrinopoulou)

One-sample test

Parametric

- ▶ One-sample t-test

Is the mean BMI of the students in my university different from the mean BMI of all students?

Non-parametric

- ▶ One-sample Wilcoxon signed rank test

Is the median score value of the students in my university different from the median score value of all students?

One-sample t-test: Theory

Scenario

Is the mean BMI of the students in my university different from the mean BMI of all students?

Connection with linear regression

$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where $x_i = 0$

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

One-sample t-test

Scenario

Is the mean BMI of the students in my university different from the mean BMI of all students?

Alternatively

$$H_0 : \mu = 0$$

$$H_1 : \mu \neq 0$$

More general

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

One-sample t-test

If **one-tailed**

Is the mean BMI of the students in my university larger than the mean BMI of all students?

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0$$

or

Is the mean BMI of the students in my university smaller than the mean BMI of all students?

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0$$

One-sample t-test

Test statistic

$$t = \frac{\bar{x} - \mu_0}{sd(x)/\sqrt{n}}$$

- ▶ Sample mean: \bar{x} (sample of students in my university)
- ▶ Population mean: μ_0 (all students)
- ▶ Standard deviation of the sample: $sd(x)$
- ▶ Number of subjects: n

Sampling distribution

- ▶ t -distribution with $df = n - 1$
- ▶ Critical values and p-value

One-sample t-test

Type I error

- ▶ Normally $\alpha = 0.05$

Draw conclusions

- ▶ Compare test statistic (t) with the critical values $_{\alpha/2}$ or the p-value with α

If **one-tailed**: Compare test statistic with the critical value $_{\alpha}$

Two-sample test (independent samples)

Parametric

- ▶ Two-sample t-test (independent samples)

Is the mean BMI of the students from group 1 different from the mean BMI of the students of group 2?

Non-parametric

- ▶ Two-sample Wilcoxon rank sum test

Is the distribution of the score values of the students in group 1 different from the distribution of the score values of the students in group 2?

Connection with linear regression!

Two-sample test (dependent samples)

Parametric

- ▶ Two-sample t-test (dependent samples)

Is the mean BMI of the students before the exams different from the mean BMI of the students after the exams?

Non-parametric

- ▶ Two-sample Wilcoxon signed rank test

Is the median score value of the students in my university different this year compared to next year?

Connection with linear regression!

M-sample test

Parametric

- ▶ Analysis of variance / F-test

Is the mean BMI of the students different in groups 1, 2 and 3?

Non-parametric

- ▶ M-sample Kruskal-Wallis test

Is the distribution of the score values of the students different in groups 1, 2 and 3?

Connection with linear regression!

Correlation test

Parametric

- ▶ Pearson correlation test

Is the height of the students in my university linearly associated with their weight?

Non-parametric

- ▶ Spearman correlation test

Is the height of the students in my university monotonically associated with their weight?

Connection with linear regression!