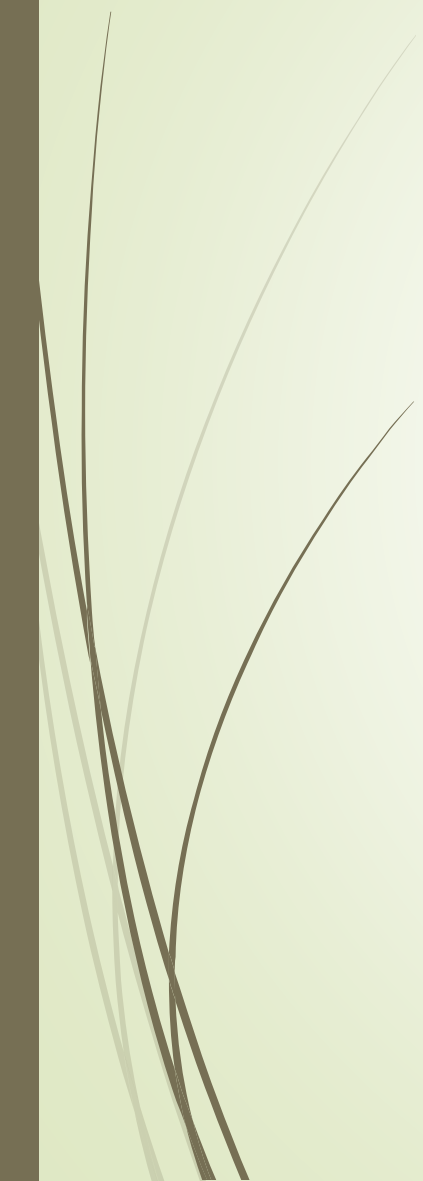




Statistical tests

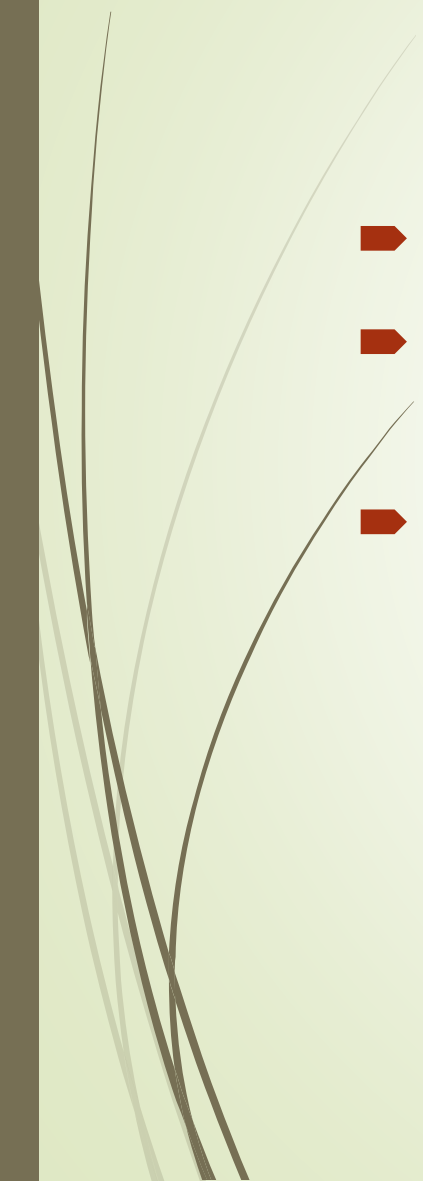


Dr. Tudor Călinici

2026



Objectives

- ▶ Test the normality of quantitative data
 - ▶ Test the difference of the means using parametrical tests
 - ▶ Test the association using Chi-squared test
- 



Statistical tests

- All the time when we analyze the population based on samples, we must apply statistical tests in order to sustain our findings
- Define hypotheses, apply the test, interpret p
 - To choose the appropriate test for a practical condition
 - The goal of the test should be appropriate
 - The specific conditions of the test must be met



Research

1. Formulate the hypotheses (fundamental research, documentation, etc.)

2. Experiment (sample)

3. Analyzing data

- ▶ The statistical test will be used to
 - ▶ Test the internal quality of the sample
 - ▶ Verify if the hypotheses is confirmed in the sample

Testing hypothesis

- Define hypothesis
- Choose the statistical test
 - It may involve other statistical tests
- Interpret the results

Quantitative variables

- We are interested to assess the performance of an automatic blood pressure monitor, so we'll compare the results given by this device with the results given by a manual blood pressure monitor.
- The automatic blood pressure monitor is **NOT CONFORM** if its results vary more than 10% of the results given by a manual blood pressure monitor.
- In order to do that we'll measure the blood pressure in a sample of 98 persons with both devices
- variable: BP – continuous quantitative variable



Quantitative variables

- H_0 : there is no difference between the results given by the automatic blood pressure monitor and the results given by a manual blood pressure monitor.
- H_1 : there is difference (more than 10%) between the results given by the automatic blood pressure monitor and the results given by a manual blood pressure monitor.
- Which statistical test should we use?



Testing the difference between quantitative data

Data follow normal distribution

- **Z**
- **T**
- **ANOVA**
- **...**

Parametrical tests

Data not follow normal distribution

- **Kruskal-Wallis**
- **Mann-Whitney U**
- **Mann-Whitney-Wilcoxon**
- **Wilcoxon signed-rank**
- **...**

Non - Parametrical tests

Statistical tests

Test the normality of the data



Quantitative variables

- We are interested to assess the performance of an automatic blood pressure monitor, so we'll compare the results of that with the results given by a manual blood pressure monitor.

- The automatic blood pressure monitor is NOT CONFORM if its results vary more than 10% of the results given by a manual blood pressure monitor.

- In order to do that we'll measure the blood pressure in a sample of 98 persons with both devices - variable: SBP/DBP – continuous quantitative variable

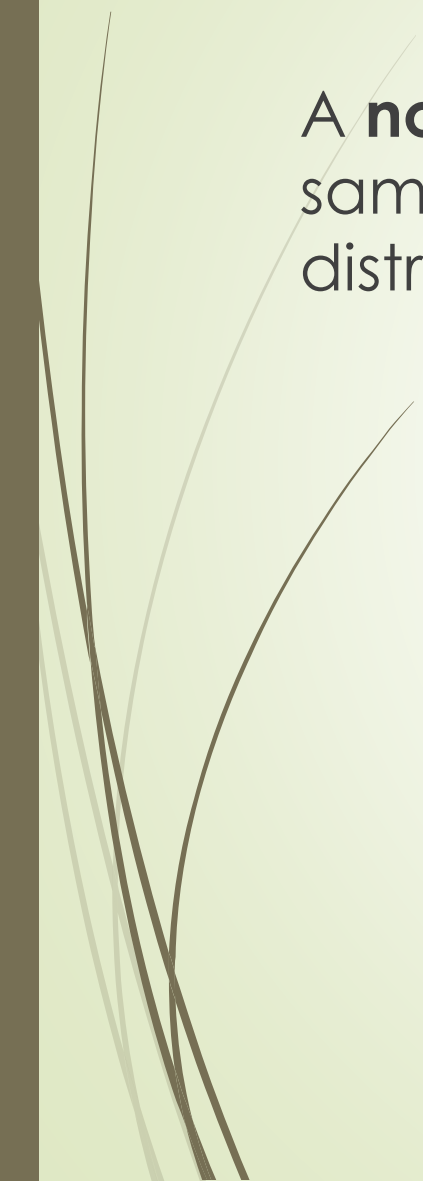
- Should we use a parametrical test, or should we use a non-parametrical test?





Quantitative variables

A **normality test** is used to determine whether sample data has been drawn from a normally distributed population (within some tolerance).



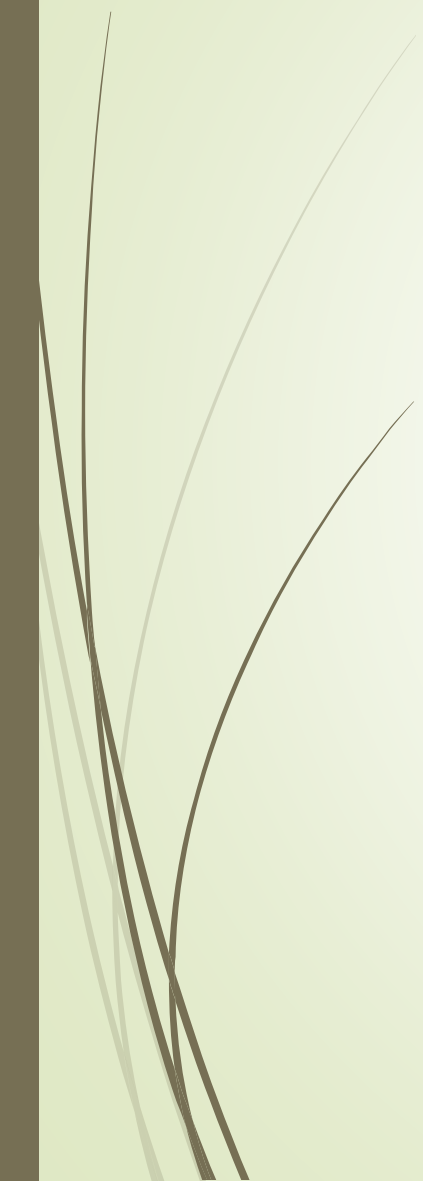


Normality tests

- Kolmogorov-Smirnov (only if sample size ≥ 50)
- Shapiro-Wilk
- Test the internal quality of the sample



Normality tests - Hypotheses

- **H0: There is no difference between the sample distribution and normal distribution**
 - **H1: There is difference between the sample distribution and normal distribution**
- 



Normality tests - decision

The result of the test: p

If **$p \leq 0,05$** H_0 can be rejected, H_1 could be accepted with 95% of confidence – **the sample distribution is different than normal distribution**

If **$p > 0,05$** H_0 cannot be rejected, we cannot demonstrate that the sample distribution is different than normal distribution – **the sample can be considered as normal distributed**

Example - decision



- H_0 : There is no difference between the distribution of the data in the sample and the normal distribution
- H_1 : There is difference between the distribution of the data in the sample and the normal distribution
- We apply Kolmogorov-Smirnov (sample size=98 ≥ 50)
- $p=0,03$
- Because $p < 0,05$ we can reject H_0 and accept H_1 with 95% of confidence, so with 95% of confidence, there is difference between the distribution of the data in the sample and the normal distribution, **so we must use a non-parametrical statistical test**





- We decide to repeat the experiment; this time we have a sample size 32
- We apply Shapiro-Wilk test (sample size=32<50)
- $p=0,09$
- Because $p>0,05$ we cannot reject H_0 , we cannot demonstrate that there is difference between the distribution of the data in the sample and the normal distribution, so we must use a parametrical statistical test



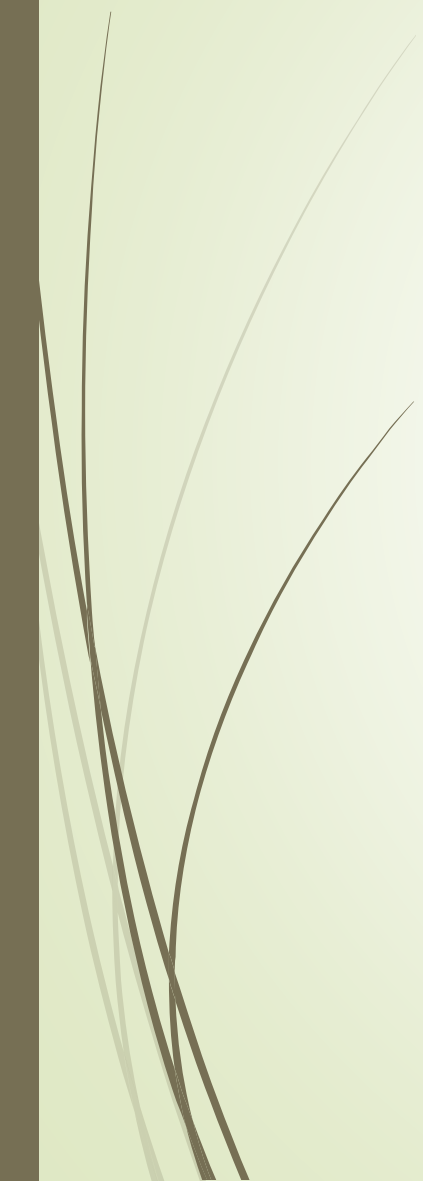


Conclusion

- When we have to analyze a sample according to a quantitative variable we must decide if we'll use parametrical or non-parametrical tests. For that we must check if the data distribution is similar with normal distribution using normality tests (Kolmogorov-Smirnov or Shapiro-Wilk test).
- If $p > 0,05$, parametrical tests should be used
- If $p \leq 0,05$ non-parametrical tests should be used

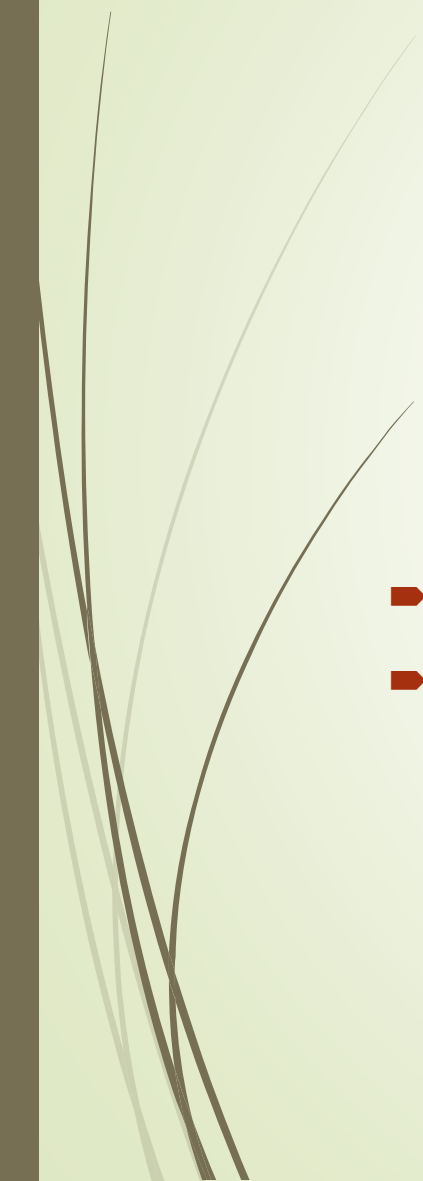


Remember!

- **The result of the normality test is only used to decide what kind of test you'll use to test the statistical hypothesis!**
 - **The normality test will be always followed by another statistical test – parametrical or non-parametrical**
- 



Parametrical tests

- Quantitative normal distributed variables
 - The estimator for the values of the variable of interest is the mean
- 



Z test , two samples, for means

- Conditions of application
 - Normal distributed data
 - The variance in population is known (or at least it can be estimated)
 - Big samples (sample size > 30)
- Null hypotheses: $m_1 = m_2$
- Alternative hypotheses: $m_1 \neq m_2$ (two tail test) or $m_1 > m_2$ (one tail test)

Problem

- A trainer is interested to check that a new training method leads to a better performance on long-jump challenge. For this, he compared the results of a group of 50 athletes trained with this new method - average of the jumps $m_1 = 730$ cm- with the results of a control group (size 83) – average of the jumps $m_2 = 710$ cm. The data are normal distributed. Usually, on long-jump, the expected variance is 30 cm. The confidence level is set to 95%.
- Conditions for Z test
 - Normal distributed data - YES
 - The variance in population is known (or at least it can be estimated) – YES – 30 cm
 - Big samples (sample size > 30) YES – sample size = $50 + 83 = 133$



Applying z test

- H_0 : there is no significant difference between the performance on long-jump challenge in the two groups (m_1 is not significant different than m_2)
- H_1 : the new method of training increase the performance on long-jump challenge ($m_1 > m_2$) – one tail test
- Apply the Z test $\Rightarrow p=0,005$
- $P < 0,05$ so with 95% of confidence we can say that the new method of training increase the performance on long-jump challenge



T test (student test) – two samples

- Conditions of application
 - Normal distributed data
- Null hypotheses: $m_1 = m_2$
- Alternative hypotheses: $m_1 \neq m_2$ (two tail test) or $m_1 > m_2$ (one tail test)

Are the samples independent?

Yes - T test for independent samples

- In order to test the efficacy of a drug in reducing the value of triglycerides we compare the level of the triglycerides of a group of people which had used that drug with the level of the triglycerides of a group of people which had used placebo

No - T test for paired samples

- In order to test the efficacy of a drug in reducing the value of triglycerides we measure the value of triglycerides before and after the treatment

T test for paired samples

- In order to test the efficacy of a drug in reducing the value of triglycerides we measure the value of triglycerides before and after the treatment (95% confidence)
- H_0 : There is no significant difference between the mean of the TG before and after the treatment
- H_1 : The mean of TG after the treatment is significantly lower than the mean of TG before treatment (one tail test)

Applying the test $\Rightarrow p=0,0000000005$

- $P < 0,05$, we reject H_0 , so, with 95% of confidence, the drug is efficient in reducing TG

TG before mg/dL	TG after mg/dL
304	154
302	152
251	221
426	296
387	157
296	126
293	223
488	438
383	313
432	282
341	211
214	164
201	100
484	254
342	232
409	219
353	183
223	73
224	71

T test for independent samples

In order to test the efficacy of a drug in reducing the value of triglycerides we compare the level of the triglycerides of a group of people which had used that drug with the level of the triglycerides of a group of people which had used placebo

There are two different T tests for independent samples (T test assuming equal variances (Student's) and T test assuming unequal variances (Welch's)) and we must choose which test is appropriate

The decision will be given by the result of another statistical test (F test for variances / Barlet test / Levine Test)



F test for variances / Barlet test / Levine Test

- H_0 : There is no significant difference between the variances of the two groups (the variances are equal)
- H_1 : There is significant difference between the variances of the two groups (the variances are unequal)
- If $p \leq 0,05$ we will apply test for unequal variances – (Welch's Test)
- If $p > 0,05$ we will apply test for equal variances (T Test)

Example - In order to test the efficacy of a drug in reducing the value of triglycerides we compare the level of the triglycerides of a group of people which had used that drug with the level of the triglycerides of a group of people which had used placebo – 95% confidence

- H_0 : There is no significant difference between the mean of the TG when we use drug or when we use placebo
- H_1 : The mean of TG when we use drug is significantly lower than the mean of TG when we use placebo (one tail test)
- First, we apply F test $\Rightarrow p=0,39 > 0,05$, so we must use T test for equal variances (Student's)
- We apply T test for equal variances $\Rightarrow p=0,36 > 0,05$ – we cannot demonstrate that this drug is more efficient than placebo in reducing the TG value

TG after drug mg/dL
154
152
221
296
157
126
223
438
313
282
211
164
100
183
73
71

TG after placebo mg/dL
169
152
226
301
152
141
238
453
308
287
216
164
95
254
242
224
198
88
66

	TG after drug mg/dL	TG after placebo mg/dL
Mean	197,75	209,1578947



Remember

- In order to apply T-test for independent samples you must first test the equality of the variances (F test / Levine test / Barlet test)
- The test for the equality of the variances (TEV) will be always followed by a T-test for independent samples as follows:
 - T-test for independent samples assuming equal variances (Student's) when the result of TEV $p > 0,05$
 - T-test for independent samples assuming un-equal variances (Welch's Test) when the result of TEV $p \leq 0,05$

Chi square test

- Is used to test the independence (association) between **qualitative** variables
- The null hypothesis: there is no association between variables
- The alternative hypothesis: there is association between variables

Chi square by example – dichotomical variables

- We are searching for hypothetical association between smoking and lung cancer. For that, from oncology department records, we select a group of 160 people who were diagnosed with lung cancer. We also select a control group of 240 persons which were never diagnosed with lung cancer. The study of this sample leads us to the following contingency table:

	Cancer present	Cancer absent	Total
Smoking yes	80	50	130
Smoking No	80	190	270
Total	160	240	400

The hypotheses, level of significance, critical region

- H_0 : there is no association between smoking and lung cancer
- H_1 : there is association between smoking and lung cancer
- We'll use a level of significance of 5%
- Critical region for 5% level of significance is $[3.84, +\infty)$

The parameter of the test

- The test consists by the comparison of the observed contingency table with the theoretical contingency table. The theoretical contingency table contains the theoretical distribution of the data when null hypothesis should be true

	Cancer present	Cancer absent	Total
Smoking yes	a^o	b^o	a^o+b^o
Smoking No	c^o	d^o	c^o+d^o
Total	a^o+c^o	b^o+d^o	$a^o+b^o+c^o+d^o$

Observed contingency table

	Cancer present	Cancer absent	Total
Smoking yes	a^t	b^t	a^t+b^t
Smoking No	c^t	d^t	c^t+d^t
Total	a^t+c^t	b^t+d^t	$a^t+b^t+c^t+d^t$

Theoretical contingency table

Computing the theoretical table

- The null hypotheses change the distribution of the data in the contingency table, but the totals will remain the same

	Cancer present	Cancer absent	Total
Smoking yes	$\frac{(a^0+c^0) * (a^0+b^0)}{(a^0+b^0+c^0+d^0)}$	$\frac{(b^0+d^0) * (a^0+b^0)}{(a^0+b^0+c^0+d^0)}$	a^0+b^0
Smoking No	$\frac{(a^0+c^0) * (c^0+d^0)}{(a^0+b^0+c^0+d^0)}$	$\frac{(b^0+d^0) * (c^0+d^0)}{(a^0+b^0+c^0+d^0)}$	c^0+d^0
Total	a^0+c^0	b^0+d^0	$a^0+b^0+c^0+d^0$

$$a^t = (a^0 + c^0) * (a^0 + b^0) / (a^0 + b^0 + c^0 + d^0)$$

$$b^t = (b^0 + d^0) * (a^0 + b^0) / (a^0 + b^0 + c^0 + d^0)$$

$$c^t = (a^0 + c^0) * (c^0 + d^0) / (a^0 + b^0 + c^0 + d^0)$$

$$d^t = (b^0 + d^0) * (c^0 + d^0) / (a^0 + b^0 + c^0 + d^0)$$

Compute the theoretical table

	Cancer present	Cancer absent	Total
Smoking yes	$\frac{160 * 130}{400}$	$\frac{240 * 130}{400}$	130
Smoking No	$\frac{160 * 270}{400}$	$\frac{270 * 240}{400}$	270
Total	160	240	400

	Cancer present	Cancer absent	Total
Smoking yes	52	78	130
Smoking No	108	162	270
Total	160	240	400

Chi square value

	Cancer present	Cancer absent	Total
Smoking yes	a^o	b^o	$a^o + b^o$
Smoking No	c^o	d^o	$c^o + d^o$
Total	$a^o + c^o$	$b^o + d^o$	$a^o + b^o + c^o + d^o$

	Cancer present	Cancer absent	Total
Smoking yes	80	50	130
Smoking No	80	190	270
Total	160	240	400

	Cancer present	Cancer absent	Total
Smoking yes	a^t	b^t	$a^t + b^t$
Smoking No	c^t	d^t	$c^t + d^t$
Total	$a^t + c^t$	$b^t + d^t$	$a^t + b^t + c^t + d^t$

	Cancer present	Cancer absent	Total
Smoking yes	52	78	130
Smoking No	108	162	270
Total	160	240	400

$$\text{Chi square} = \frac{(a^o - a^t)^2}{a^t} + \frac{(b^o - b^t)^2}{b^t} + \frac{(c^o - c^t)^2}{c^t} + \frac{(d^o - d^t)^2}{d^t}$$

$$\text{Chi square} = \frac{(80 - 52)^2}{52} + \frac{(50 - 78)^2}{78} + \frac{(80 - 108)^2}{108} + \frac{(190 - 162)^2}{162} = 37.2$$

Decision

- If Chi square is in the critical region, we can reject H_0 and accept H_1
- If Chi square is not in the critical region, we cannot reject H_0 and cannot demonstrate H_1
- Chi square = 37.2 in $[3.84, +\infty)$, we can reject H_0 and accept H_1
- With 95% of confidence there is association between smoking and lung cancer

Chi square using statistical software

- If you use a statistical software to apply Chi square test, you can make the decision by interpreting p value
- If $p \leq 0,05$ you can reject H_0 and accept H_1 with 95% of confidence
- If $p > 0,05$ you cannot reject H_0

- For our example $p = 0,00000000105$