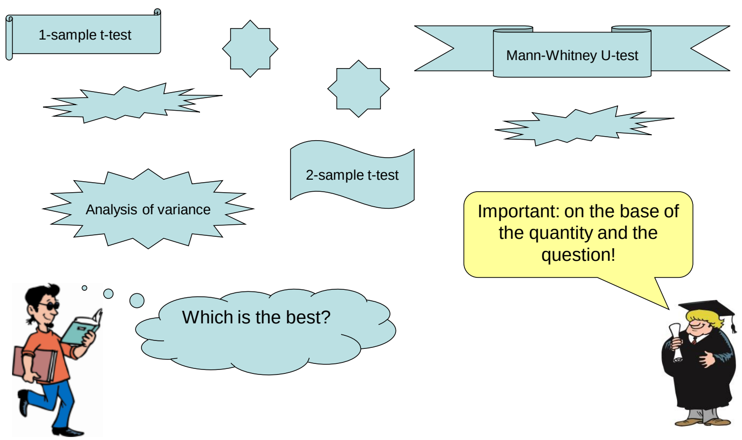
Selection



According to the variable

parametric

Testing the hypothesis on the parameter(s) of a known distribution.
(most frequently the normal distribution).

non-parametric

Testing the hypothesis on the parameter(s) or types of an unknown distribution.

Non-parametric tests

Distribution free methods.

advantage : independent from the distribution.
disadvantage: normally it's power is less than the parametric one.

Ranking tests:
Instead of original values we use the so-called **ranks**.

Ranks

Rank: numerical or ordinal data are replaced by the rank in a series sorted according to a certain rule.

e.g.:
- lieutenant
- major

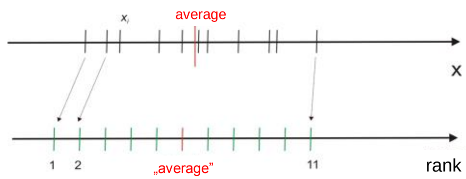
colonel

etc.

ties:
In the case of same values every value are replaced by the average of the ranks.

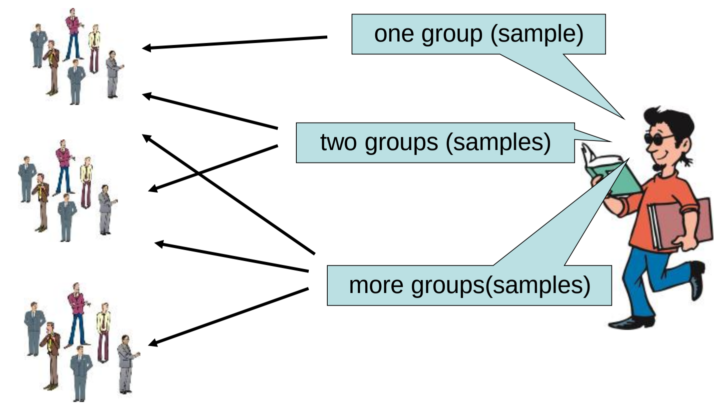
value:	1.2	2	2	3.5	4
rank:	1	2,5	2,5	4	5

the „average" of the ranks is the median



The median plays the role of the average.

according to the question



Summary table

	parametric	non-parametric
one group	1-sample t-test,	Wilcoxon's signed-rank test, sign-test
two groups	2-sample t-test	Mann-Whitney U-test
more groups	ANOVA	Kruskall-Wallis test

Examination in one group

Question: On the base of the sample the parameter of the population may be a given value?

parametric

μ = ?

Nullhypothesis: $\overline{x} = \mu_0$

1-sample t-test

non-parametric

median = ?

Nullhypothesis: the median is a given value

Wilcoxon's signed-rank

1-sample t-test

example: The medicine effective or not?

↓

Nullhypothesis: not! μ_0 = 0. But the average is not 0!

sample	Average
1.	-0.2 °C
2.	-1 °C
3.	-1.5 °C

If the difference is bigger, it seems to be more probable that the difference is non-random.

What is the big difference?

What is the measure of the difference?

Standard error: the average devations of the averages from the μ.

$(\bar{x} \pm s_{\bar{x}})$ ~ 68% - confidence interval.

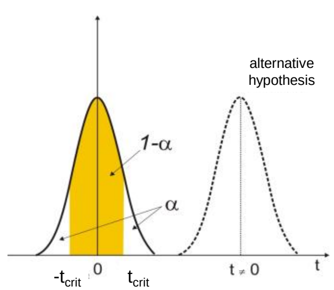
t-value

$$t = \frac{x - \mu_0}{s_x}$$

Compare the difference to the standard error!
(μ_0 very frequently = 0)

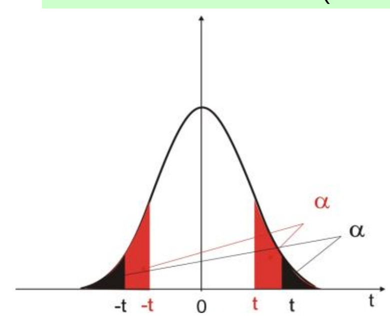
The averages fluctuate around the μ_0 so the *t*-values deviate around the 0.

(Providing, that the nullhypothesis is true!)



What is the advantage of the t-value?

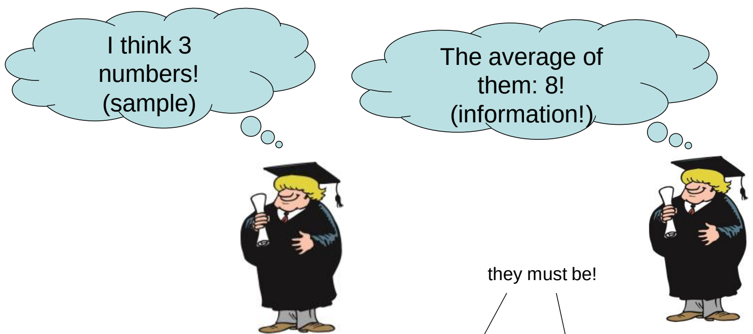
We are able to calculate the probabilities on the base of this distribution!!! (Student- or *t*-distribution)



This describes only the **random deviations** of the *t*-values!

The shape of the distribution depends on the no. of elements.

Degree of freedom (d.f.)



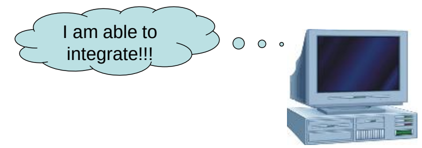
3, 12, 8 or 5, 7, 11 etc.

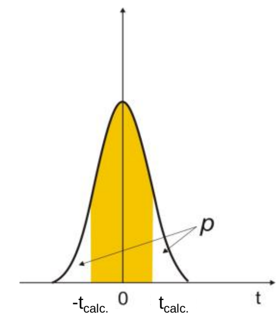
d.f. = n

3, 12, **9** or 5, 7, **12** etc.

d.f. = n-1

Decision using computer





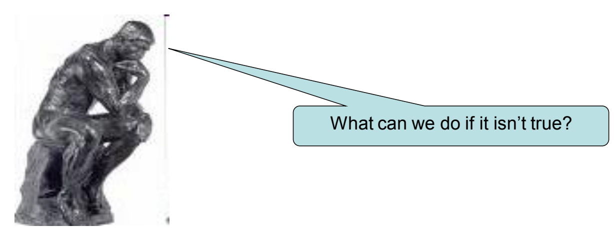
p: probability, that the $t_{calculated}$, is so large randomly.

Decision

- 1. If the probalitiy of the random deviation is small ($p(|t| \geq t_{crit}) \leq \alpha$) – ***reject*** the nullhypothesis.
- 2. If the probalitiy of the random deviation is large ($p(|t| \geq t_{crit}) > \alpha$) – ***accept*** the nullhypothesis.

Condition for 1-sample t-test

- **Task**: Decision about the μ on the base of one sample.
- The variable must have **normal distribution**.



Wilcoxon's signed-rank test

Example: Is there an effect of an entertaining movie on the patients? (The numbers are scores)

n	before	after	Diff.
1	2	2	0
2	0	1	1
3	3	2	-1
4	2	4	2
5	1	3	2
6	3	3	0
7	1	4	3
8	1	5	4
9	5	2	-3
10	4	4	0

Normal distribution?

Ranking

Sort the absolute values of the differences (without 0-s)! Let the sign of ranks be same then the differences! Calculate the averages and sd of signed-ranks!

Diff.	absolute value	rank	Signed-rank
0	0		
1	1	1.5	1.5
-1	1	1.5	-1.5
2	2	3.5	3.5
2	2	3.5	3.5
0	0		
3	3	5.5	5.5
4	4	7	7
-3	3	5.5	-5.5
0	0		

The nullhypothesis

The median = 0!
The deviation is random!

There is no effect of the movie!

H_0: $\mu_0 = 0$

H_1: $\mu_0 \neq 0$

known distribution

$$t = \frac{\overline{R} - 0}{s/\sqrt{n}}$$

If n is enough large!

(Anyway a special, non-discussed distribution must be used!)

$\overline{R}$ - ranks	the average of the signed-
s –	the standard deviation

Rememer!
„average" of the ranks = median

Decision

This is known!!!

Of course! This is similar to the 1-sample t-test!!!

Paired t-test

If the data may be paired according to a rule!

Observation on the same person, paired organ (e.g. kidney).

Rare, on the base of viewpoints (age, profession, etc.).

Look at: decreasing the fewer.

Experimental design

„real" 1-sample t-test

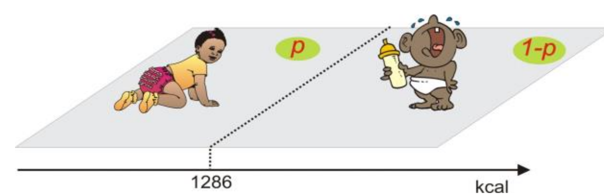
Is it possible that the μ is equal to a value?

$$t = \frac{x - \mu_0}{s_x}$$

Rare case.

sign-test

Example: Energy uptake in the population of 2-year old children.
Question: May be the median (This derives from an another test) is 1286 kcal?
Nullhypothesis: median = 1286 kcal (deviation is random).



Test

Small sample

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

binomial distribution

Large sample

$$z = \frac{|x - np| - 1/2}{\sqrt{np(1-p)}}$$

standard normal distribution

x – no. of children below 1286 kcal.
n – no. of children in test.
p – probabilty, that randomly smaller (look at: binomial distribution)

Decision

Calculate the probability of the random deviation. (binomial, or standard normal distribution)

End of this part!