Hypothesis test

Questions

(example)

How can we answer?

Is the medicine effective?

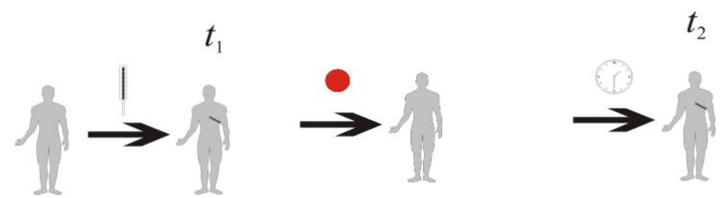
literature

experiments

An example

Qustion: Does the medicine decrease the fewer?

experiment



$$x = \Delta t = t_2 - t_1$$

Hypothesises

The medicine isn't effective.

The medicine is effective.

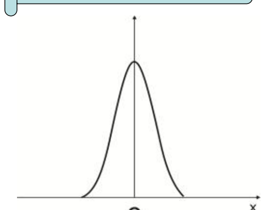
Exclusive statements, enough to test one of them!

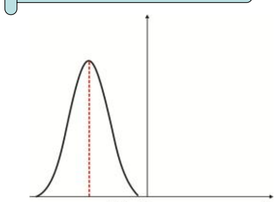
Which is better?

The distribution of the observed quantity

not effective

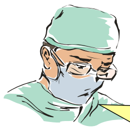
effective





The result of the random effects is that $\mu = 0$.

How much is the effect?



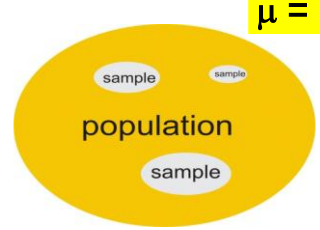
If we know the population!!! (we are able to calculate μ)

result:		conclusion:
$\mu = 0$	→	The medicine isn't effective.
$\mu < 0$	→	The medicine is effective and μ characterizes the effect.

The situation is more difficult

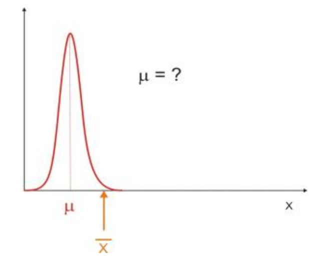
Normally the population is unknown.

μ = ?



The sample differs from the population! (**sampling error**)

E.g. the averages fluctuate around the μ!



What is the reason of the deviation?

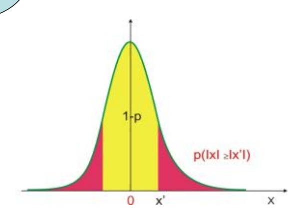
*Sampling error, **random fluctuation**.* (Our hypothesis is right!)

The hypothesis is false (mistake!). The deviation is non-random.

What is the base of the decision?

How much is the chance that the sample derives from the given population?

We must know the parameters of the distibution!



Nullhypothesis: (H_0)

The deviation of the sample or samples from the population or populations is a **random deviation** due to the sampling error. Frequently it is a negative answer. (e.g.: the medicine is not effective.)

Alternative hypothesis: (H_1)

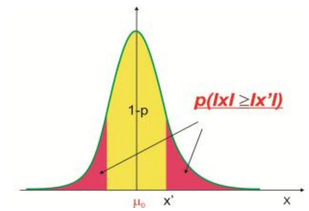
The deviation of the sample or samples from the population or populations is not a **random deviation**. (e.g.: the medicine is effective)

Nullhypothesis

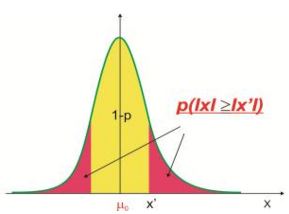
How much is the probability of the random deviation?

In the case of known distribution we are able to determine!

(The shape of the distribution not always gaussian, but known!)



Significant?



If the ***p*** is enough large, may be random, if the ***p*** is enough small we consider the difference being ***significant***!

p is the probability being **random**!

Significance level

Enough large, enough small?

Select a value as limit! This is the significance level.

Symbol: α.
In medical practice this value is frequently 5%.

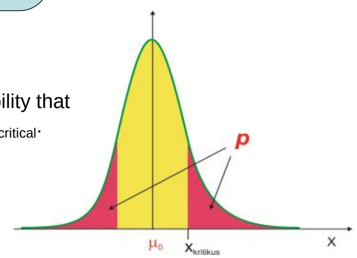
The base of the decision

If the ***p*** is enough small, there is a big chance, that the nullhypothesis is not true. So the alternative hypothesis is more probable.

$x_{critical}$: the value belonging to the significance level

$x_{calculated}$: the value calculated from the sample

p is the probability that $x_{calculated} \geq x_{critical}$.



Decision

- 1. If the probability of the random deviation is small ($p(|x| \geq |x_{crit}|) \leq 5\%$) –we ***reject*** the nullhypothesis.
- 2. If the probability of the random deviation is large ($p(|x| \geq |x_{crit}|) > 5\%$) – we ***accept*** the nullhypothesis.

The answer is newer yes – no or true - false!!!

The possibility of the error

decision: the nullhypothesis is

reality: the nullhypothesis	accepted	rejected
true	**Right decision**	**I. Type error (α)**
false	**II. Type error (β)**	**Right decision**

Test in one sample: 1-sample t-test

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

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 being 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.

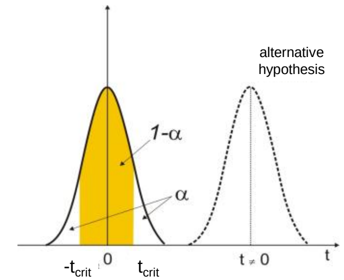
t-value

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

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

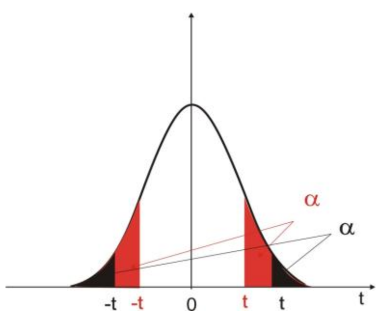
The averages fluctuate around the μ_0 so the *t*-values deviate around the 0.

(providing, that the nullhypothesis is true!)



Why is the t-value is better?

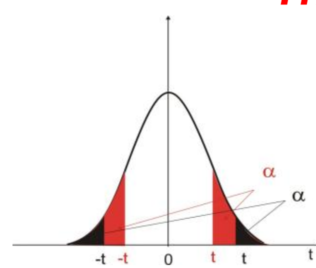
We are able to calculate the probabilities on the base of this distribution!!! (Student- or *t*-distribution)



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

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

The t-table



t-table

significance level

d.f.	probability (p) 0.1	0.05	0.02	0.01
1	6.31	12.7	31.8	63.7
2	2.92	4.3	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.6
5	2.02	2.57	3.37	4.03

d.f.: n-1

Different t_{crit} values belong to different sigficance level.

Degree of freedom (d.f.)

I think 3 numbers! (sample)

The average of them: 8! (information!)

they must be!

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

d.f. = n

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

d.f. = n-1

Decision the base of t-table

t-table

significance level

d.f.	probability (p) 0.1	0.05	0.02	0.01
1	6.31	12.7	31.8	63.7
2	2.92	4.3	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.6
5	2.02	2.57	3.37	4.03

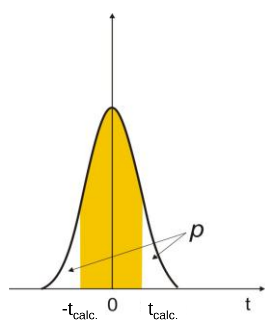
d.f.: n-1

Select an approriate significance level!

If t≥2.78 **reject**, if smaller **accept** the nullhypothesis.

Decision using computer

I am able to integrate!!!



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

Decision

- 1. If the probability of random deviation is small ($p(|t| \geq t_{krit}) \leq 5\%$) – ***reject*** the nullhypothesis.
- 2. If the probability of random deviation is large ($p(|t| \geq t_{krit}) > 5\%$) – ***accept*** the nullhypothesis.

Conditions

- Question: May be a parameter equal to a certain value?
- The variable must have **normal dustribution**.

Test in two goups

Question: May the samples derive from the same population or the parameters of the two populations are the same?

$\mu_1 = \mu_2$?

Nullhypothesis: $\mu_1 = \mu_2$

(but normally $\bar{x}_1 \neq \bar{x}_2$)

2-sample t-test

2-sample t-test

$\bar{x}_1 \neq \bar{x}_2$

?

Known distribution is necessary!

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s^* = \sqrt{\frac{Q_1 + Q_2}{n_1 + n_2 - 2}}$$

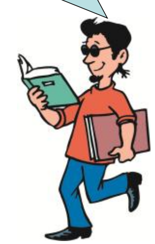
Test

The t-vaue is same!

How much is the d.f.?

d.f. = n_1 +n_2 -2

((n_1-1)+(n_2-1))



Conditions for the test

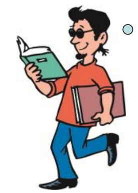
- task: comparison of two **independent** sample.
- The quantity must have **normal distribution**.
- The sd are **same** in the two goups.

The last one is new! How is it possible to prove?

Test for standard deviations

How can I do?

Nullhypothesis: the two sd are the same, the difference is random (sampling error).



It is similar to the hypothesis testing!

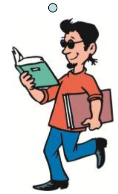
F-test

$$F = \frac{s_1^2}{s_2^2}$$

A so-called F-distribution belongs to the nullhypothesis.

But what sd is in the nominator?

The larger variance is in the denominator! ($F \geq 1$)



Decision

- 1. if the probability of the random deviation is small ($p(F \geq Fc_{rit}) \leq 5\%$) – ***reject*** the nullhypothesis.
- 2. if the probability of the random deviation is large ($p(F \geq F_{crit}) > 5\%$) – ***accept*** the nullhypothesis.

Two or more variables

Correlation and regression

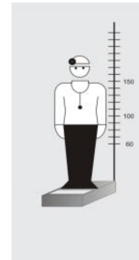
The relationship between two variables.

Method for estimating the relationships among variables.

Correlation of two variables

Example:
Is there any relationship between the height and weigth?

Experiment:



Data pairs:

No.	Height (cm)	Weight (kg)
1	150	61
2	170	70
3	166	75
4	174	70
5	180	72
6	155	50
7	172	65
8	161	59
9	177	81

Graphic representation

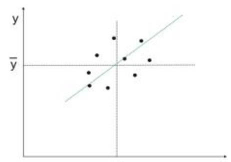
E.g.: height is the **x** and weigth is the **y**.

Possible situations:

A

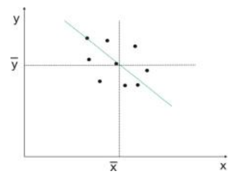
no any tendency

B



Positive tendency

C



Negative tendency

Pearson's correlation coefficient

$$r = \frac{\mathrm{cov}(x,y)}{s_x \cdot s_y} = \frac{Q_{xy}}{\sqrt{Q_x \cdot Q_y}}$$

$$Q_{xy} = \sum_i \left[(x_i - \bar{x})(y_i - \bar{y})\right]$$

$$Q_x = \sum_i (x_i - \bar{x})^2$$

$$Q_y = \sum_i (y_i - \bar{y})^2$$

Possible range for r: $-1 \le r \le 1$

In the population:

r = 0 no correlation,

$r \neq 0$ correlation (strength is proportional to the actual value of r.)

Coefficient of determination

r^2

The coefficient of the determination tells us how strong is the relationship.
Expresses how much percent of the variability of the y values may be accounted by the variability of the independent variable or variables.

Correlation *t*-test

Calculated ***r*** is only an estimation of the ***r*** in the population.
This fluctuates around the theoretical value.
(e.g. $r_{calc} = 0.1$?)

H_0: r = 0! → $t = r\sqrt{\frac{n-2}{1-r^2}}$ → d.f.: n - 2

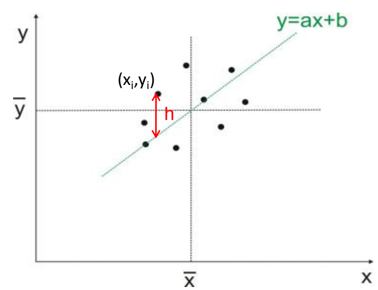
Decision: based on t-value. Look previous cases!

Condition: at least one of the variable has normal distribution.

Linear regression

If the variables have normal distribution, the relationship is linear, and we can describe it with a straightline.

$y_i = ax_i + b + h_i$

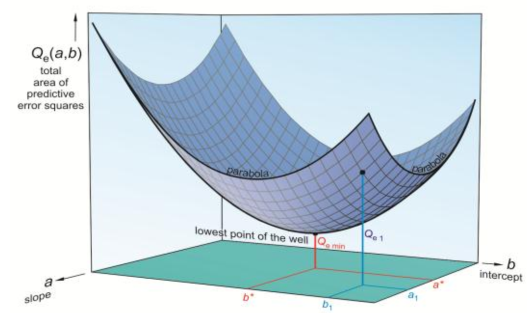


y: dependent variable
x: independent (explanatory) variable
h_i: error term = $y_i - (ax_i+b)$.
(the difference between the actual value and the predicted value.)

Least-squares method

$$Q_e = \sum_i h_i^2 = \sum_i (y_i - (ax_i + b))^2$$

The $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$ measured values.
Unknown are the ***a*** and ***b***!



Which is the best straightline?

Q_e has minimum! →

$$a^* = \frac{Q_{xy}}{Q_{xx}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$b^* = \bar{y} - a^* \bar{x}$$

r^2: coefficient of determination.
How much part of the variance of y is explained by x.

Prediction of insulin sensitivity by BMI.

indep.	regr. coeff.	st. error	t	p	decision
BMI	-0.077	0.018	-4.25	0.0011	significant
r^2	0.6				

Chi-square test (analyzing frequency data)

Example: headache

pill

Effective:
no headache

Not effective:
remaining headache

Experiment

1. group: patients taking the ***medicine***

headache (a)

no headache (b)

2. group: patients taking the ***placebo***

headache (c)

no headache (d)

(a,b,c,d are frequency data)

Contingency table

	headache	no headache	Total
1. group	a	b	a+b
2. group	c	d	c+d
total	a+c	b+d	n

So-called 2 x 2 table.

Nullhypothesis

If the effect is independent from the medicine, we expect:

$$\frac{a}{b} = \frac{c}{d} \longrightarrow a \times d = b \times c$$

Nullhypothesis: the effect is independent from the medicine is due to the placebo effect only.

Chi-Square test for independence.

χ^2-distribution

Shortcut formula for 2 x2 tables:

$$\chi^2 = \frac{n(ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

Nullhypothesis: χ^2-value is 0, the difference is only due to the sampling error.

χ^2-distribution describes the random deviations of the χ^2-value.

Decision

Same, then in the case of *t*-distribution. We use χ^2-distribution.

Expected value is 0, if the null hypothesis is true.

if $\chi^2_{calc} \geq \chi^2_{crit}$ - reject the null hypothesis else accept.

or $p(\chi^2 \geq \chi^2_{calc}) \leq 5\%$ - reject the null hypothesis else accept.

degree of freedom: in this special case = 1.

In general:

d.f.=(r-1)(c-1), where r – no. of rows
c – no. of columns