

## Vizsgálat két csoportban

Kérdés: A két minta származhat-e azonos populációból, vagy a két populáció paraméterei azonosak?

**paraméteres**

$$\mu_1 = \mu_2 ?$$

Nullhipotézis:  $\mu_1 = \mu_2$

kétmintás t-próba

**nem paraméteres**

Nullhipotézis: a két minta azonos populációból származik.

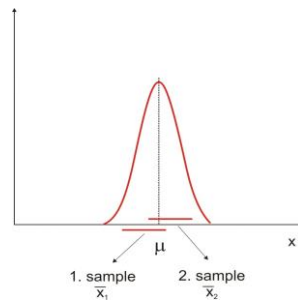
Mann-Whitney U-próba

## Miért nem párosított?

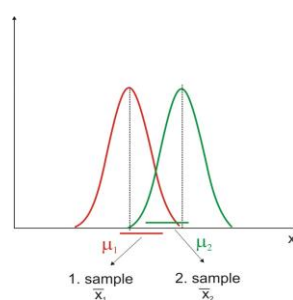
A			B			C		
172	184	12	162	170	8	162	184	22
180	172	-8	165	172	7	165	180	15
165	180	15	172	175	3	172	175	3
184	175	-9	180	180	0	180	172	-8
162	170	8	184	184	0	184	170	-14
	átlag	3.6		átlag	3.6		átlag	3.6
	szórás	11.33		szórás	3.782		szórás	15.11
	st. hiba	5.066		st. hiba	1.691		st. hiba	6.757
	t	0.711		t	2.129		t	0.533
	p	0.517		p	0.1		p	0.622

## Kétmintás t-próba

egy populáció  
(a két átlag eltérése véletlen)



két populáció  
(a két átlag eltérése nem véletlen)



## Standard hiba

$$s_1 = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n_1 - 1}}$$

$$s_2 = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n_2 - 1}}$$

$$s_{\bar{x},1} = \frac{s_1}{\sqrt{n_1}}$$

$$s_{\bar{x},2} = \frac{s_2}{\sqrt{n_2}}$$

**Közös standard hiba:** a két standard hiba súlyozott átlaga.

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{Q_1 + Q_2}{n_1 + n_2 - 2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

## Kétmintás t-próba

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



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



Ismert eloszlású változóra van szükség!



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

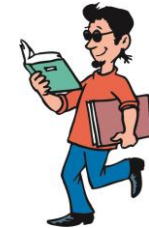
## A próba

A t-érték az t-érték!



Akkor meg tudom csinálni!  
Pardon, mennyi a szabadsági fokok száma?

$$\text{sz.f.} = n_1 + n_2 - 2 \\ ((n_1 - 1) + (n_2 - 1))$$



## A kétmintás t-próba feltétele

- A feladat: két egymástól **független** csoport összehasonlítása.
- A változó **normális eloszlású** legyen.
- A **szórás** a két csoportban **azonos**nak tekinthető.



Ez utóbbi új!  
Hogyan állapítható meg?

## A szórások vizsgálata

Hogyan fogjunk hozzá?



Nullhipotézis: a két szórás azonos, az eltérés véletlen (mintavétel).

De hiszen ez olyan, mint egy hipotézis-vizsgálat!



## Az F-próba

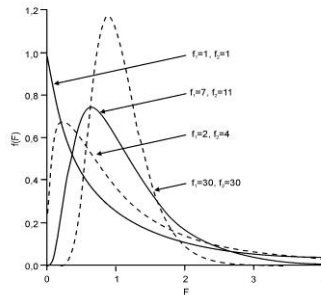
A nullhipotézishez tartozik egy ún. F-eloszlás.



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

Szabadsági fokok:

számláló:  $n_1 - 1$   
nevező:  $n_2 - 1$

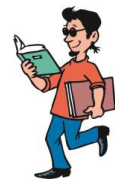


## Az F-próba szabadsági fokai

Számítógéppel számolva, bármelyik lehet. Táblázatot használva, viszont mindig a nagyobb. (Ennek megfelelően kell a sz.f.-okat figyelembe venni)



De melyik variancia legyen a számlálóban?



## A döntés

- 1. Ha a véletlen eltérés valószínűsége kicsi ( $p \leq \alpha$ ) – **elvetjük** a nullhipotézist.
- 2. Ha a véletlen eltérés valószínűsége nagy ( $p > \alpha$ ) – **megtartjuk** a nullhipotézist.

F táblázat

Számítógép: F-próba

## Ha a két szórás nem azonos!

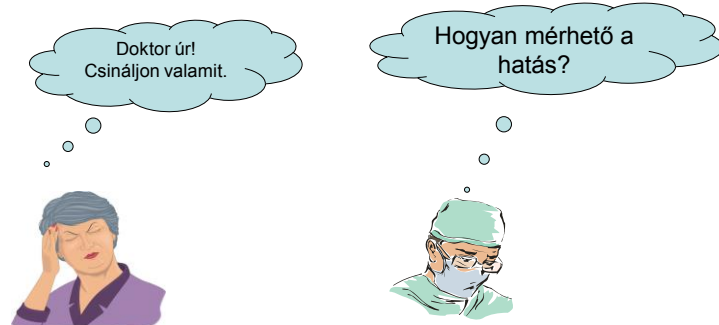
Korrekciók:

A szabadsági fokok korrekciója.

a t-érték korrekciója.

## Mann-Whitney U-próba

Példa: hatásos-e a fejfájás-csillapító?

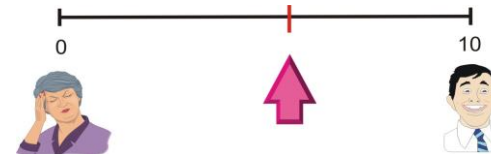


## Kísérlet

I. csoport:  
(eset)  
aszpirint kap

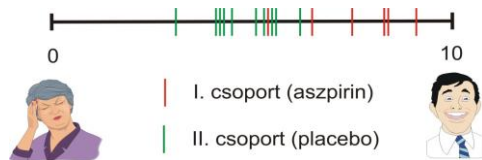


II. csoport:  
(kontroll)  
placebo-t kap  
(hatóanyag nélküli  
tabletta)



Ez egy önkényes,  
folytonos skála.

## Eredmények



érték	3,1	4,1	4,2	4,3	4,5	5,1	5,3	5,4	5,5
rang	1	2	3	4	5	6	7	8	9
érték	5,6	6,2	6,2	6,5	7,5	8,3	8,3	8,4	9,1
rang	10	11,5	11,5	13	14	15,5	15,5	17	18

## A nullhipotézis megfogalmazása

a „gyógyszer” nem  
hatásos.



A két csoport azonos  
populációhoz tartozik.



## A rangok összege

(avagy a kis Gauss esete a tanárral)

Gyerekek! Adjátok össze a számokat 1-től százig.

Miért adjam össze? Könnyebben is kiszámolható!



$$1 + 100 = 101 \\ 2 + 99 = 101 \dots$$

$$\sum_{i=1}^n i = \frac{n}{2} \cdot (n + 1)$$

## A rangok összege

$T$  – a rangok összege az 1. csoportban, véletlen eloszlás esetén a várható értéke:

$$n_1 \cdot \frac{n_1 + n_2 + 1}{2}$$

( $n_1$  elem, amelyek átlaga =  $(n_1 + n_2 + 1)/2$ )

Nullhipotézis: az ettől való eltérés véletlen.

Kis  $n$ : egy U-eloszlás írja le a véletlen eltérés valószínűségét.

## A „nagy átalakítás”

Ha  $n$  elég nagy:

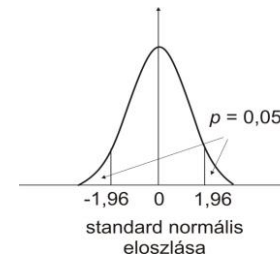
$$z = \frac{T - n_1(n_1 + n_2 + 1)/2}{s}$$

$$s = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}$$

A  $z$  változó standard normális eloszlású.



## Döntés



A kiszámolt  $z$ -érték: 3,24.  
Ez nagyobb, mint az 1,96.

**Következtetés:** a nullhipotézist elvetjük.

Kiszámolt  $p$ -érték  $< 0,1\%$ .  
Következtetés hasonló a fentihez.