# ORVOSI STATISZTIKA

Élettan

Anatómia

Kémia

…

Lehet kérdés?

Statisztika

? → !

Az orvos ***döntés***eket hoz!

Mikor jó egy döntés?

Mennyire helyes egy döntés?

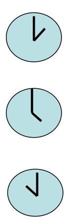
Mekkora a tévedés lehetősége?

# Az orvosi statisztika helye

**Elmélet: matematika**


**Gyakorlat: alkalmazott statisztika (kiemelt példák)**


# Példa: test hőmérséklet



36,7 °C

36,9 °C

36,6 °C

36,7 °C

36,9 °C

36,5 °C

1. Mérési pontatlanság.
2. ***Időbeli ingadozások!!!***

***3. Biológiai változatosság!!!***

A megfigyelt értékek nem állandóak!

Mért érték: 37,0 °C.

Egészséges vagy beteg?

# Egyéb példák

RBC: $4{,}5 \times 10^{12}$ 1/l (<u>$3{,}9$-$5 \times 10^{12}$ 1/l</u>) → normál tartomány?

Az új terápiás eljárás hatékonyabb mint a régi?

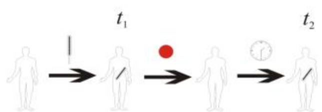
<u>Hogyan igazolható egy lázcsillapító hatásossága?</u>

↓

**Kérdések!**

# Hogyan válaszolható meg?

kísérlet




$x = \Delta t = t_2 - t_1$

**1. $\Delta t < 0$**

A gyógyszer hatásos.

***A gyógyszer nem hatásos***: fluktuáció, a beteg jobban lesz, stb.

**?**

**Következmény**:

A válasz nem egyszerű és nem mindig egyértelmű!

(Nem csak a gyógyszer befolyásolhatja a test hőmérsékletet.)

**2. $\Delta t >= 0$**

A gyógyszer nem hatásos.

***A gyógyszer hatásos***: fluktuáció, nem emelkedik tovább, stb.

# Változók

| változó | tartomány | típus | A változó típusa | |
|---|---|---|---|---|
| magasság | ~50 cm … ~250 cm | valós szám | numerikus | folytonos |
| fogak száma | 0 .. 32 | egész szám | | diszkrét |
| vércsoport | A, B, AB, 0 | betűk | kategóriális | nominális |
| a rák stádiumai | 1 … 4 | egész szám | | ordinális |

**Leíró statisztika!**

# A változó leírása

- Típusa
- Lehetséges értékek
- Az értékek előfordulása

# Numerikus változók

| Név | ***folytonos*** | ***diszkrét*** |
|---|---|---|
| Definíció | Végtelen sok érték lehetséges, egy adott tartományban | Véges számú lehetséges érték |
| Példa | Magasság, testhőmérséklet … | A fogak száma, a gyerekek száma… |

## Kategoriális változók

| Név | ***Nominális*** | ***Ordinális*** |
|---|---|---|
| Definíció | Nincs sorrend | Létezik természetes sorrend |
| Példa | nem, vércsoport … | A betegség súlyossága, a fájdalom nagysága… |

## A lehetséges értékek megadása

- Folytonos : megadjuk a lehetséges tartományt.
  - » pl. magasság: ~60 cm - ~ 250 cm
- Egyéb: felsoroljuk az értékeket, ha lehetséges
  - » pl. vércsoport: A, B, AB, 0

## Előfordulás

***Megfigyelés***: Az értékek előfordulása nem azonos mértékű!

↓

***Kísérlet***: mérés, megfigyelés, kikérdezés...

*Csak olyan esetekkel foglalkozunk, amelyekben a kísérlet megismételhető!*

***Kimenetel***: Egy kísérlet eredménye. (pl.: egy hallgató magassága)

## Populáció

Hány kísérletre van szükség?

↓

Amennyi csak lehetséges.

↓

Ideális eset: pl. az összes ember → ***populáció***

# Minta

Egy kisebb, véges számú hányad a populációból.


$n$: az ***elemek*** száma a mintában.

$x$: a vizsgált mennyiség

$x_i$: egy elem a mintából

# A minta kiválasztása

Alapelv: **véletlen minta.**

Orvosi statisztika: ha nincs egyéb ***kizáró ok***, akkor véletlen legyen a kiválasztás!

# Előfordulás

***Gyakoriság*** ($k$): egy adott érték előfordulásának a száma.

$k_i$: az i-edik érték előfordulása.

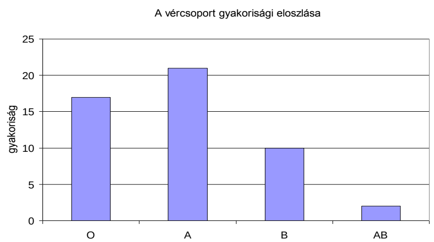
$$n = \sum_i k_i$$

# Gyakoriság eloszlás

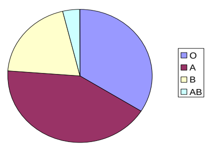
A gyakoriság a változó értékeinek a függvényében.

| Vércsoport | 0 | A | B | AB | összes |
|---|---|---|---|---|---|
| gyakoriság | 17 | 21 | 10 | 2 | 50 |

## Megjelenítés

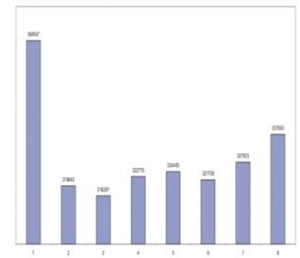


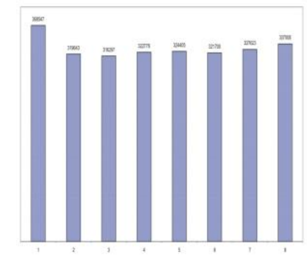

oszlop diagramm




kördiagramm

## A megjelenítés csapdái





## Relatív gyakoriság

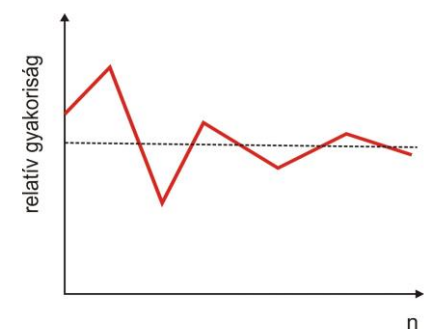
A gyakoriság aránya a teljes elemszámhoz viszonyítva.

$$\sum_i \frac{k_i}{n} = \frac{1}{n}\sum_i k_i = \frac{1}{n} \times n = 1$$

Gyakran százalékos formában adjuk meg:

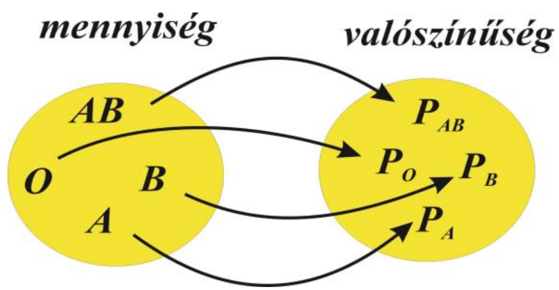
$$\frac{k_i}{n} \times 100\%$$

## Valószínűség (*P*)




A valószínűség a relatív gyakoriság értéke, ha *n* tart a végtelenhez.

## Valószínűség eloszlás




## A valószínűség tulajdonságai

$$0 \le P \le 1 \longrightarrow$$

P = 0 - sohasem fordul elő

P = 1 - mindig előfordul

példa: vércsoport

$$P_A + P_B + P_{AB} + P_0 = 1 \longrightarrow \sum_i P_i = 1$$

(ha, egymást kizáró események)

## Valószínűség és relatív gyakoriság

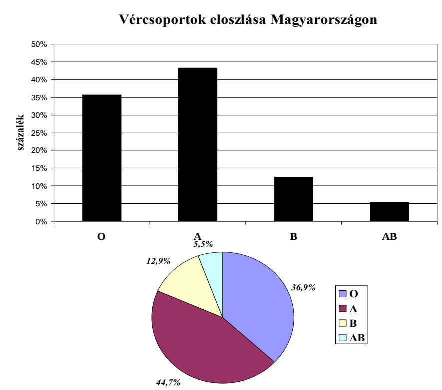
| **Minta** | **Populáció** |
|---|---|
| n véges! | $n = \infty$ |
| relatív gyakoriság | valószínűség |

---

A valószínűség nagyon gyakran nem ismert!

A gyakorlatban a relatív valószínűséget használjuk helyette.

## Megjelenítés



## Folytonos változó

Végtelen sok lehetséges érték!!!

***Osztály***: egy kis intervallum a teljes értéktartományon belül.

***Osztályszélesség***: Az intervallum hossza.

***Gyakoriság***: azon elemek száma, amelyek az adott intervallumba esnek.

↓

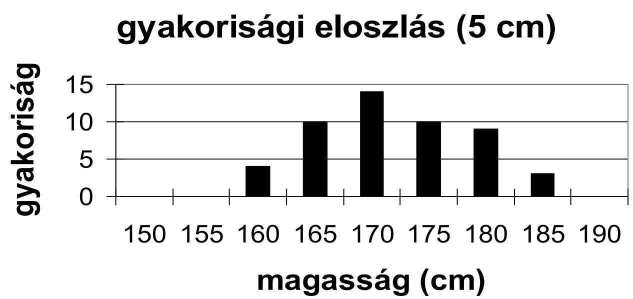
Olyan, mintha diszkrét értékek lennének !

## Egy példa

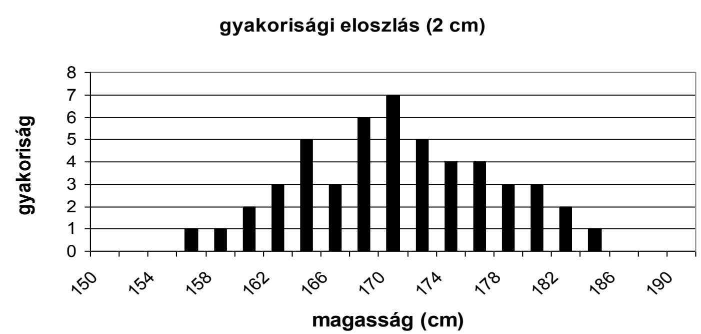
| | |
|---|---|
| 1 | 160 cm |
| 2 | 181 cm |
| 3 | 175 cm |
| 4 | 163 cm |
| 5 | 165 cm |
| 6 | 179 cm |
| 7 | 164 cm |
| 8 | 185 cm |
| 9 | 177 cm |
| 10 | 168 cm |

| osztály | $k_i$ |
|---|---|
| 160-164 | 3 |
| 165-169 | 2 |
| 170-174 | 0 |
| 175-179 | 3 |
| 180-184 | 1 |
| 185-189 | 1 |

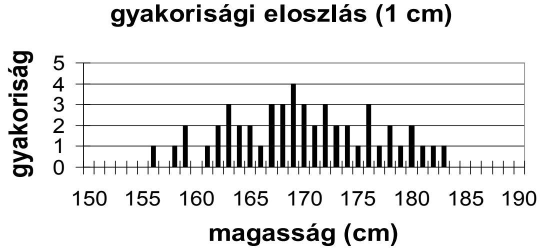
## Megjelenítés



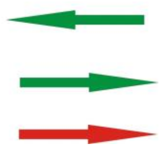

## Finomabb felosztás



# Megjelenítés




osztály-
szélesség



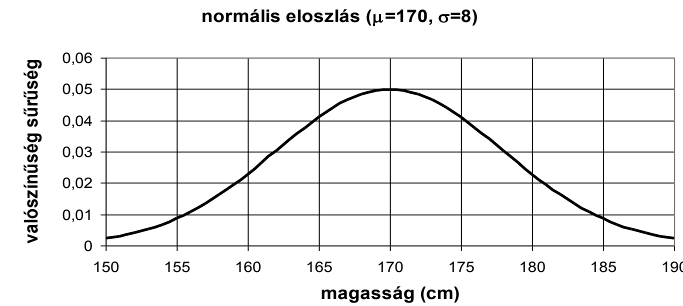
osztályok
száma

elemszám

Valóban teljes leíráshoz akkor jutunk, ha az elemszám végtelen nagy!

# Normális eloszlás

Ha az osztályszélesség végtelenül kicsi és az elemszám végtelenül nagy!




# Elméleti eloszlás

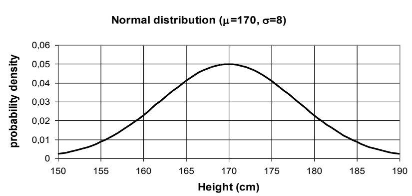
Normális vagy Gauss-eloszlás

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Paraméterek:

μ – várható érték

σ – elméleti szórás




# Elméleti leírás

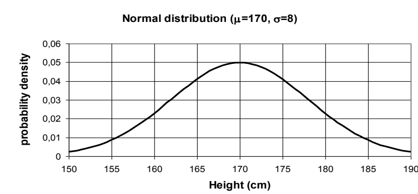
Normális vagy Gauss-eloszlás

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Paraméterek:

μ – várható érték, vagy elméleti átlag

σ – elméleti szórás



# Tulajdonságai, a paraméterek jelentése

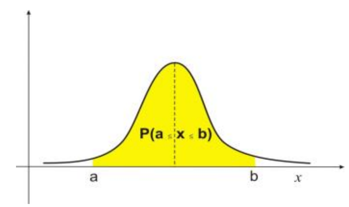
- A -∞-től a + ∞-ig terjed,
- szimmetrikus,
- A görbe alatti terület 1.

$\mu$: a görbe maximumához tartozó érték.

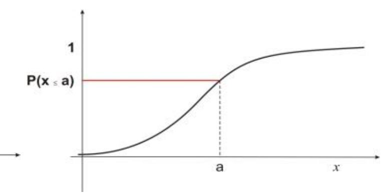
$\sigma$: az adatok átlagos eltérése a $\mu$-től.

# Sűrűségfüggvény, eloszlás függvény

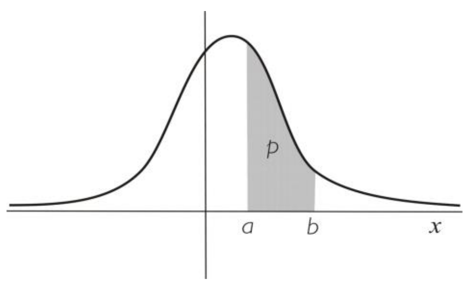
Sűrűségfüggvény




Eloszlásfüggvény




# A valószínűség jelentése



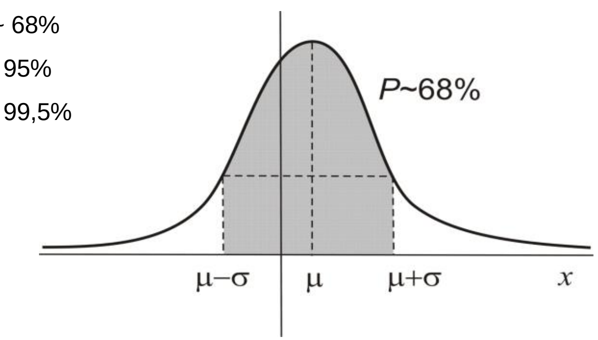

*P* annak a valószínűsége, hogy az x érték az (a,b) intervallumba esik, ill. az adatok P%-a tartozik ehhez az intervallumhoz.

# Elméleti szórás

$(\mu \pm \sigma)$ ~ 68%

$(\mu \pm 2\sigma)$ ~ 95%

$(\mu \pm 3\sigma)$ ~ 99,5%



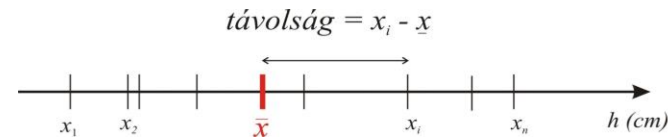
## Normális eloszlás

***Elméleti eloszlás!*** A populáció egészére jellemző. A gyakorlatban általában nem ismerjük a paramétereit.

↓

Általában csak egy vagy több véletlen mintánk van a teljes populációból.

## A μ becslése



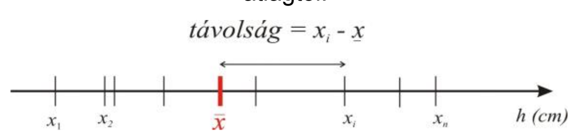

***átlag***: az elemekhez képest középen kell elhelyezkednie.

$$\sum_i (x_i - \bar{x}) = 0 \quad \longrightarrow \quad \bar{x} = \frac{\sum_i x_i}{n}$$

## A σ becslése

σ = Az adatok átlagos eltérése a μ-től.

*s* (tapasztalati szórás) = az elemek átlagos eltérése az átlagtól.




$$Q_x = \sum_i (x_i - \bar{x})^2 \geq 0$$

## A (tapasztalati) szórás

$$s = \sqrt{\frac{Q_x}{n-1}}$$

*s*: az elemek átlagos eltérése az átlagtól.

n-1: a ***szabadsági fok***

$(\bar{x} \pm s)$ ~ 68%

$(\bar{x} \pm 2s)$ ~ 95%

$(\bar{x} \pm 3s)$ ~ 99,5%

Példa: 3 szám átlaga = 12. Melyik ez a három szám?

| Minta | 1. szám | 2. szám | 3. szám |
|---|---|---|---|
| 1. | 8 | 15 | 36-(8+15) = 13 |
| 2. | 3 | 14 | 36-(3+14) = 19 |
| 3. | 10 | 21 | 36-(10+21) = 5 |

## A minta és a populáció kapcsolata

| minta | $n \to \infty$ | populáció |
|---|---|---|
| átlag | $\longrightarrow$ | $\mu$ |
| $s$ | $\longrightarrow$ | $\sigma$ |

## A hét kérdése

Hogyan becsülhetjük meg a $\mu$ és a $\sigma$ értékét?