# Kísérlettervezés alapfogalmak

## Mérési adatok vizuális elemzése

**Budapest University of Technology and Economics**
**Fault Tolerant Systems Research Group**

# Kísérlettervezés

- Cél: a modell paraméterezése a valóság alapján
  - Vagy absztrakt modell a konkrét modell alapján
- Információt **kísérlet**ek révén szerzünk
  - Pontosan mit szeretnénk tudni?
  - Ehhez milyen megfigyelést, hányszor kell elvégezni?
  - A kapott eredményekből mire lehet következtetni?
- (statisztikai) **kísérlettervezés** (Design of Experiment, DOE)
  - hatékony eljárás a kísérletek tervezésére és elemzésére
  - valós és objektív konklúziók levonásához
- Kísérletterv: még a kísérlet elvégzése előtt

# Kísérlettervezés

- Mire jó?
  - Alternatívák közötti választás
  - Érzékeny paraméterek, kulcsfaktorok
  - Megfelelő célérték, változékonyság csökkentése
  - Robosztussá tétel
- Fontos:
  - Világos cél, egyértelmű eredmények
  - Kis méret, alacsony költség
  - Valós viszonyok

# Számszerű jellemzők mérése

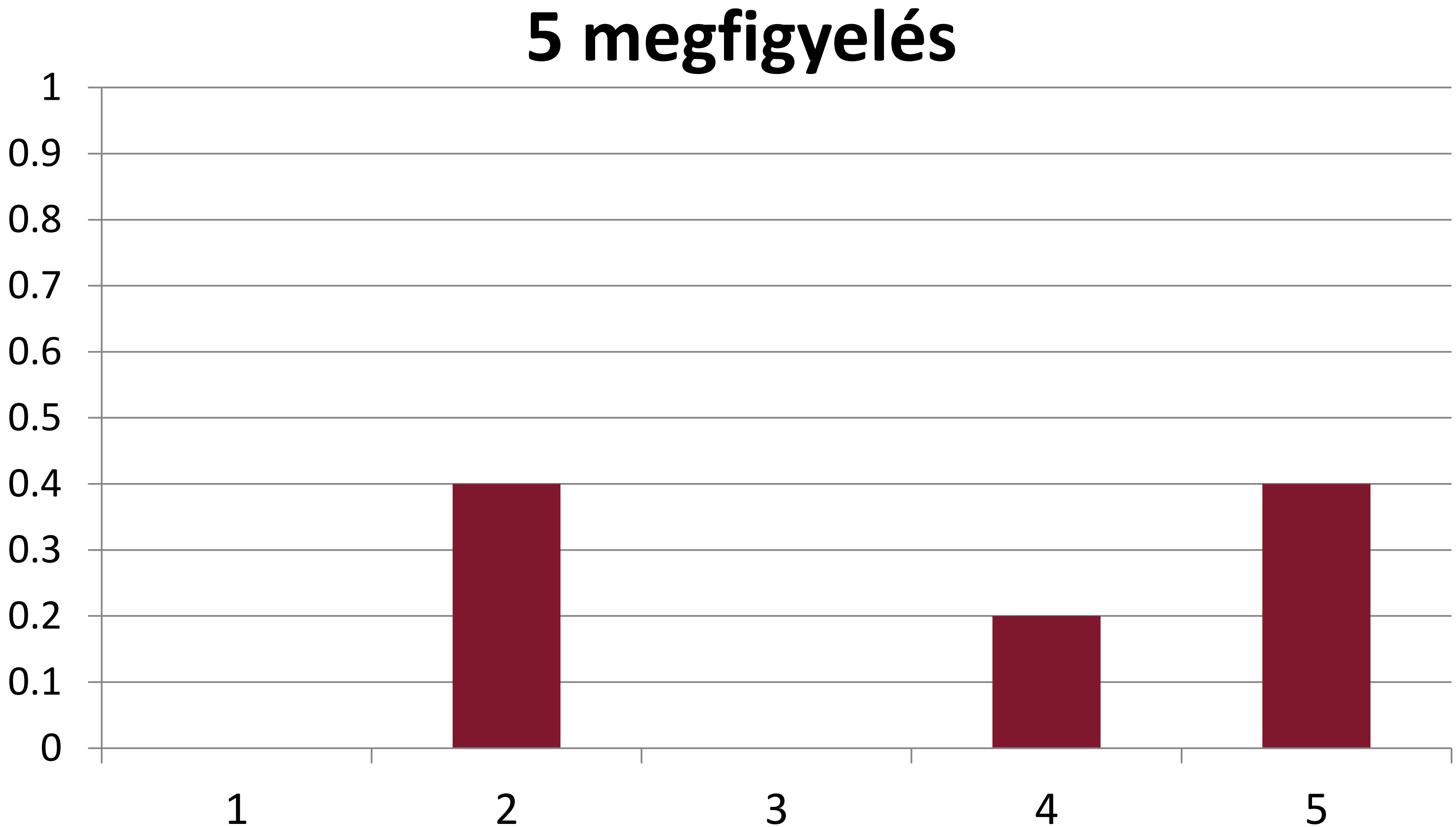
- Mérjük meg egy jellemző értékét! (1..5)


Mi van, ha az érték nem állandó? Véletlenszerű változás → végezzünk több megfigyelést!

Mért érték: 4

# Számszerű jellemzők mérése

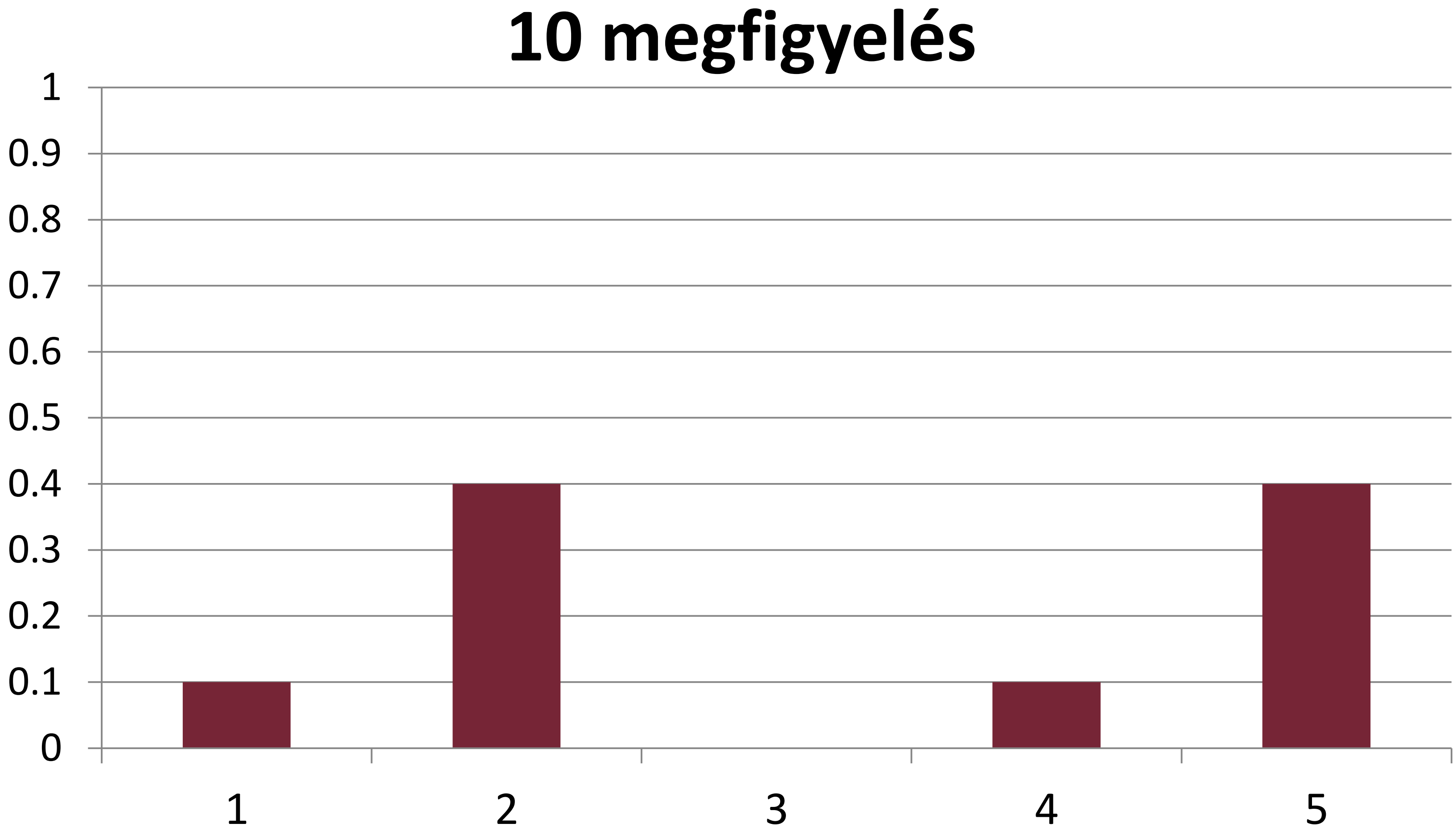
- Mérjük meg egy jellemző értékét! (1..5)




Számított átlag: 3,6

# Számszerű jellemzők mérése

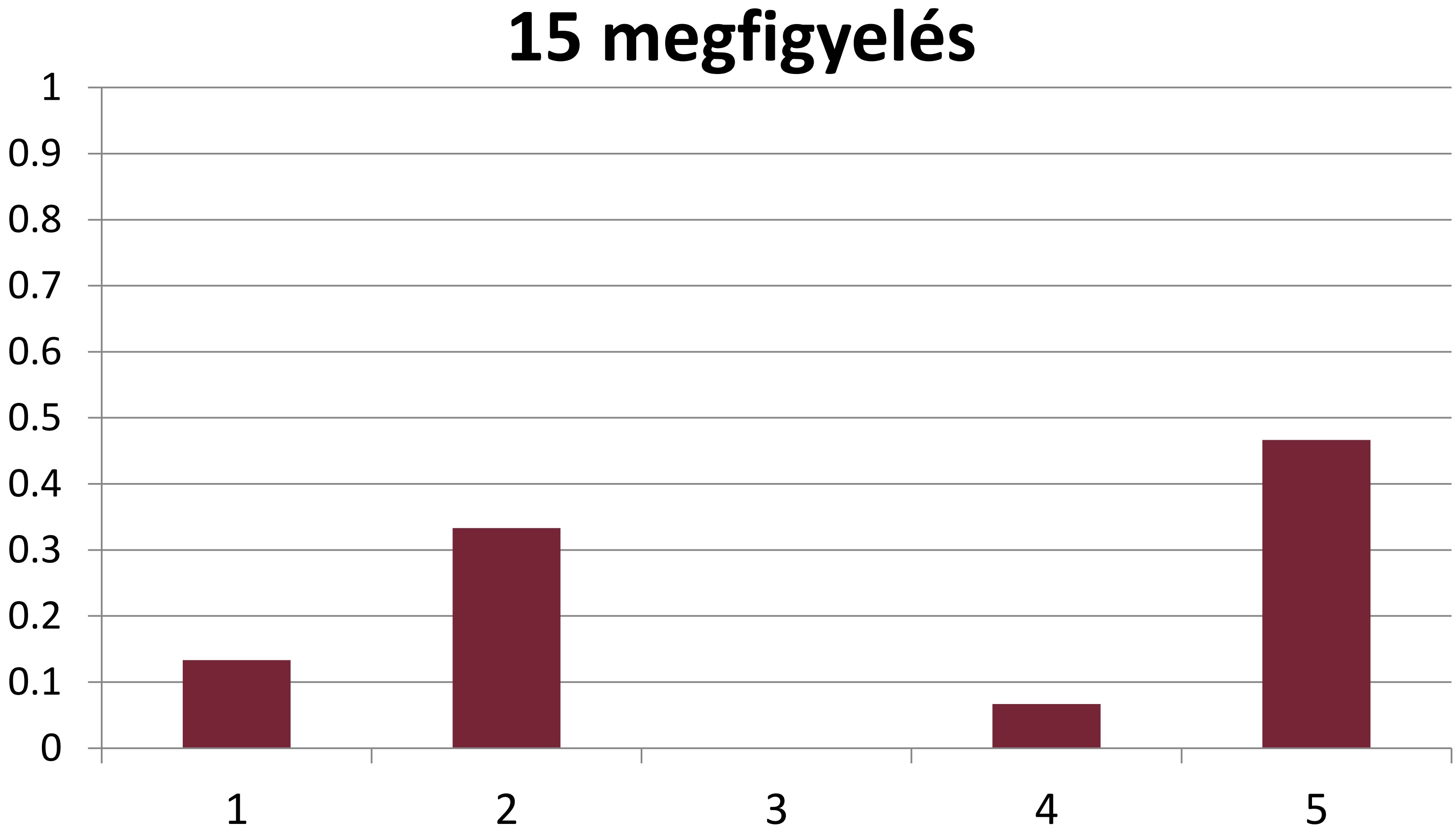
- Mérjük meg egy jellemző értékét! (1..5)




Számított átlag: 3,3

# Számszerű jellemzők mérése

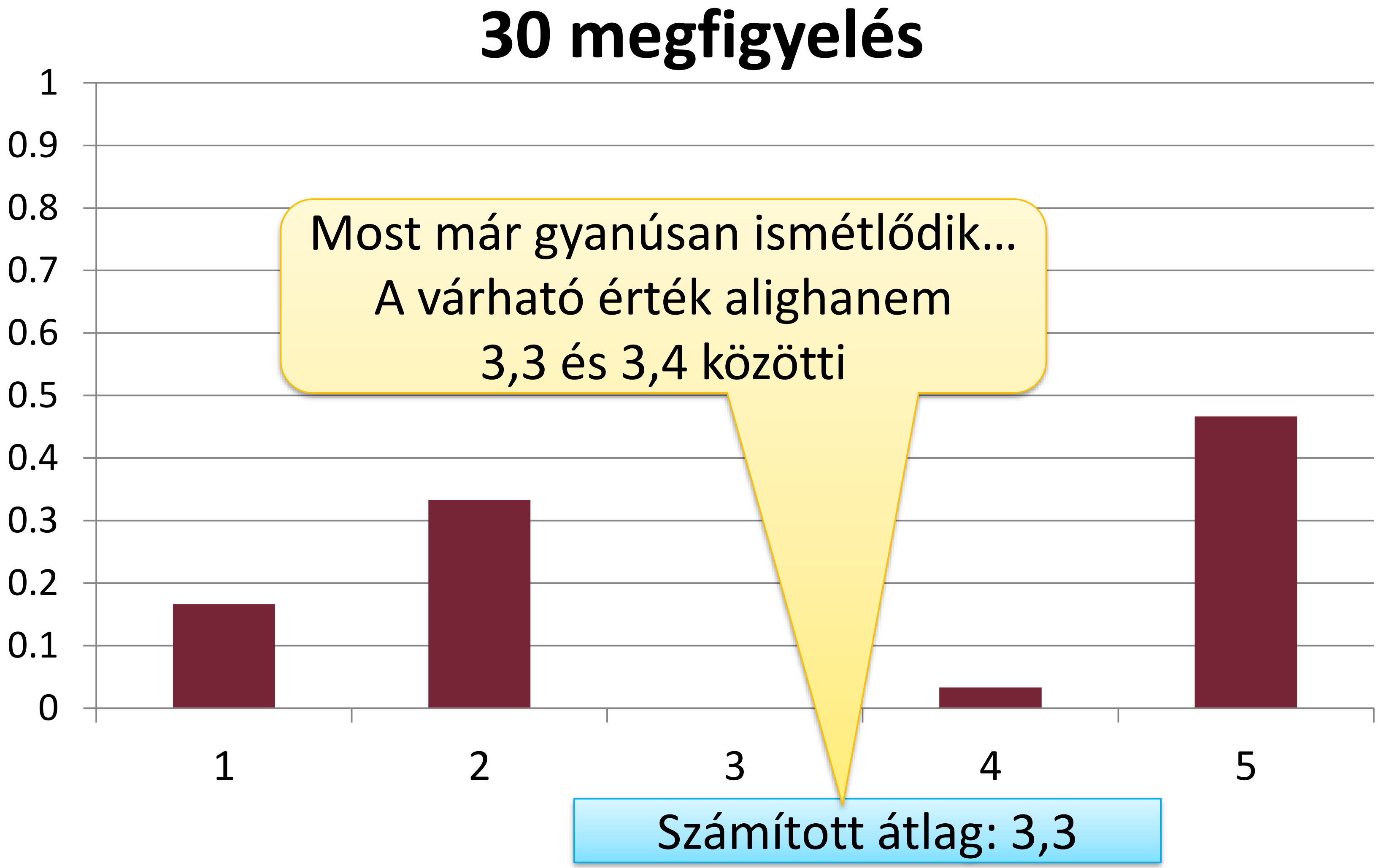
- Mérjük meg egy jellemző értékét! (1..5)




Számított átlag: 3,4

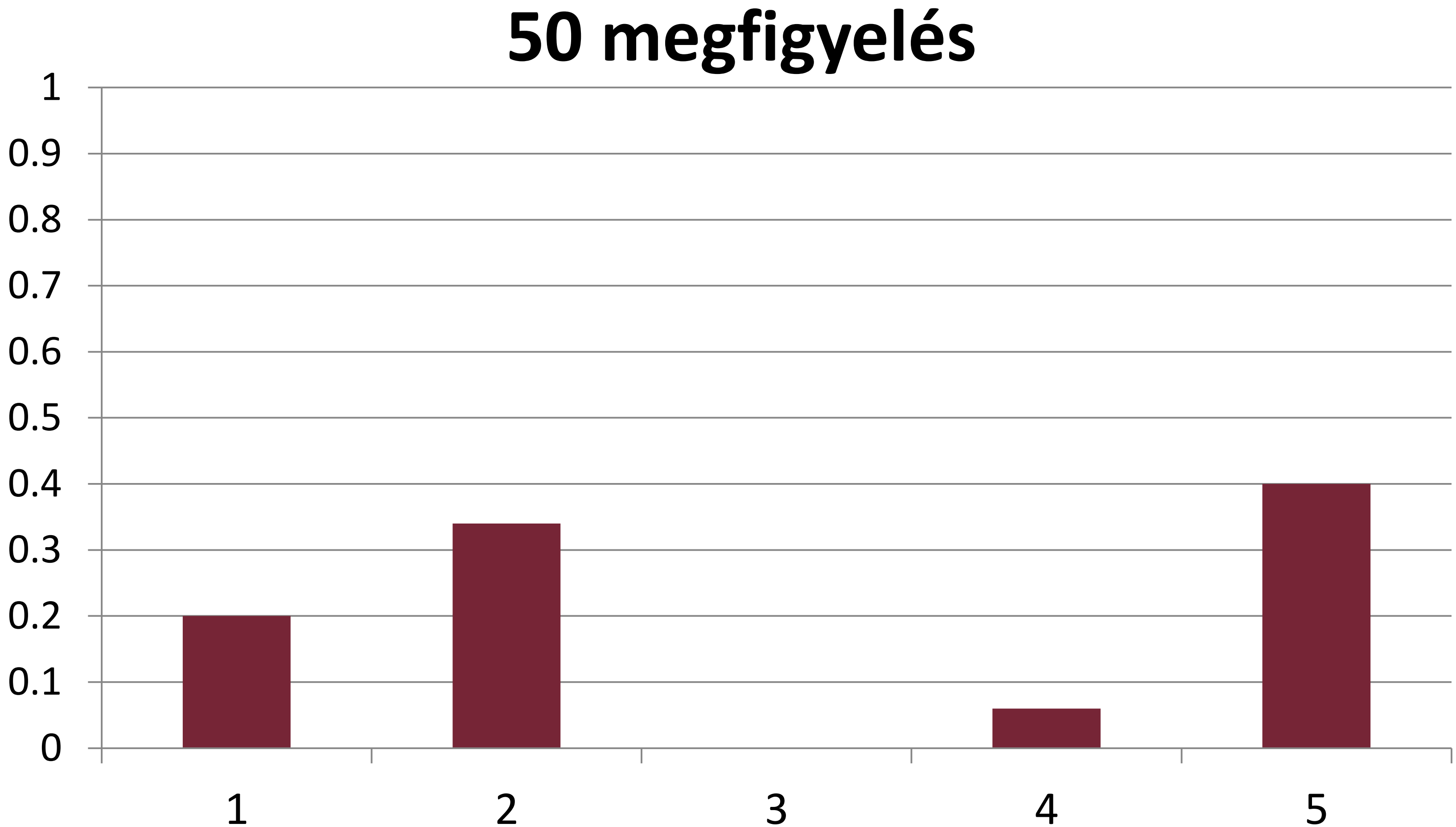
# Számszerű jellemzők mérése

- Mérjük meg egy jellemző értékét! (1..5)



# Számszerű jellemzők mérése

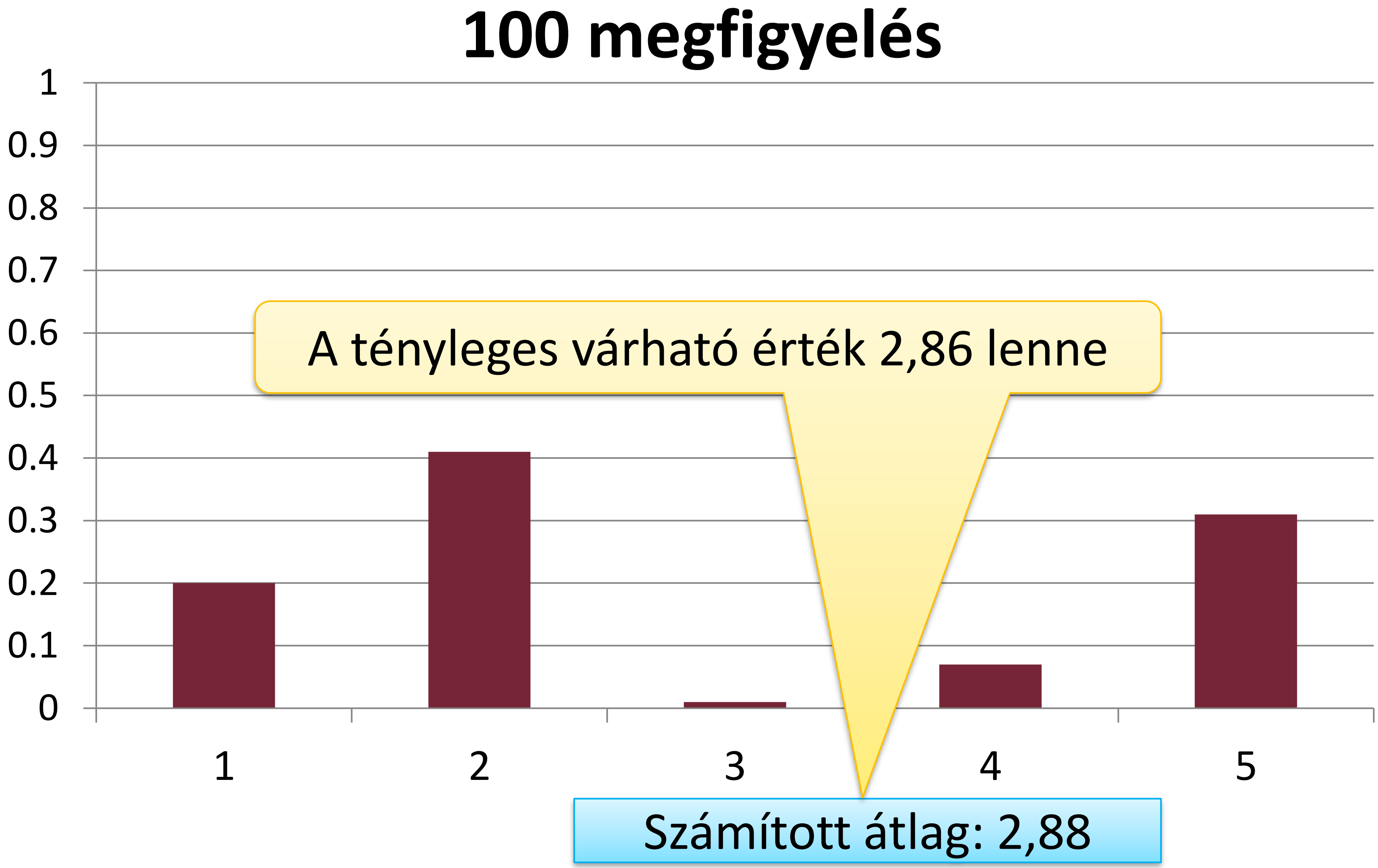
- Mérjük meg egy jellemző értékét! (1..5)




Számított átlag: 3,12

# Számszerű jellemzők mérése

- Mérjük meg egy jellemző értékét! (1..5)



# Ismétlés: tapasztalati átlag, szórás

- Valószínűségi változó: $\mathrm{E}$ (vizsgálandó jelenség)
  - Várható érték: $\mu = \mathrm{E}(X)$ átlagos viselkedés)
  - Szórás: $\sigma = \sqrt{\mathrm{E}(X-\mu)^2}$ (eltérések mértéke)
- Mintavétel: $x_1, x_2, \ldots x_t$ (mérések, megfigyelések)
  - Tapasztalati átlag: $\bar{x} = \dfrac{x_1 + x_2 + \ldots x_t}{t}$
  - Szórásra **nem jó**: $\sqrt{\dfrac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots(x_t-\bar{x})^2}{t}}$
  - Korrigált tapasztalati szórás: $\sqrt{\dfrac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots(x_t-\bar{x})^2}{t-1}}$

# Tapasztalati átlag

- Módszer: számtani átlag képzése
  - Ismételt megfigyelések
  - Egymástól függetlenül
  - Azonos feltételek mellett
- Kérdések
  - Hány megfigyelést kell végezni?
  - A tapasztalati átlag mennyire jellemzi a valódi várható értéket?
- Először tisztázandó:
  - *A tapasztalati átlag eloszlása*

# Tapasztalati átlag eloszlása

- Kísérlet = megfigyelések sorozata
- Megfigyelések sorozatának **tapasztalati átlaga**:
  - Egy jellemzőt ***t*** db független megfigyeléssel mérve,
  - majd a mért értékeket átlagolva kapott eredmény
- **Centrális határeloszlás tételé**ből következik:
  - Tetszőleges eloszlású jellemző
  - (de legyen *véges* ***m*** várható értékű és **s** szórású)
  - tapasztalati átlaga **t**→∞ esetén közelítőleg
  - **normális eloszlású**,
  - $\mu$ = ***m*** várható értékkel és $\sigma$ = $^{\mathbf{s}}/_{\sqrt{t}}$ szórással

# Tapasztalati átlag eloszlása

- Kísérlet = megfigyelések sorozata
- Megfigyelések sorozatának **tapasztalati átlaga**:
  - Egy jellemzőt ***t*** db független megfigyeléssel mérve,
  - …tt eredmény
- **C**… …övetkezik:
  - (de legyen *véges* ***m*** várható …kű és **s** szórású)
  - tapasztalati átlaga **t**→∞ esetén közelítőleg
  - **normális eloszlású**,
  - $\mu$ = ***m*** várható értékkel és $\sigma = {}^{s}/_{\sqrt{t}}$ szórással

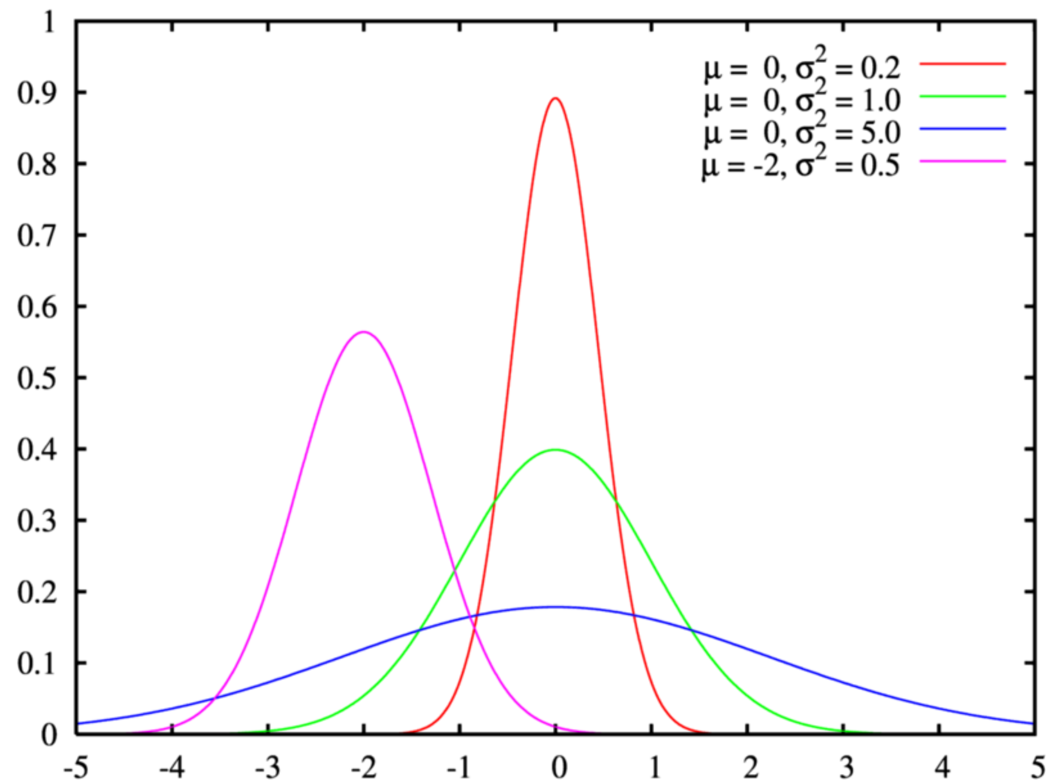
Ökölszabály:
ismert szórásnál t>30,
ismeretlen szórásnál t>100
után kezd elfogadható lenni a közelítés

# A normális (Gauss) eloszlás

- Valószínűségsűrűség-függvénye: (nem kérdezzük)

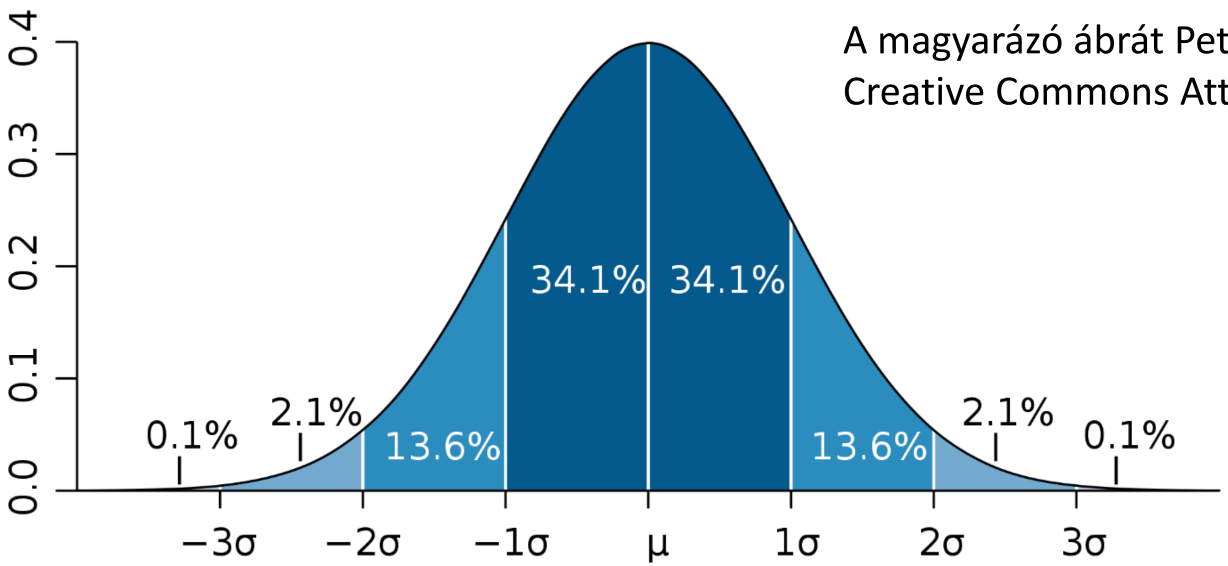
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Paraméterek
  - Várható értéke $\mu$
    - → **m** a mi esetünkben
  - Szórása $\sigma$
    - → $\mathbf{s}/\sqrt{t}$ esetünkben



# A normális (Gauss) eloszlás

- A várható érték körül koncentrálódik




A magyarázó ábrát Petter Strandmark tette közzé Creative Commons Attribution 2.5 Generic licenc alatt

- A normális eloszlású változó…
  - az esetek **68%**-ában legfeljebb **1$\sigma$** messze kerül $\mu$-től
  - az esetek **95%**-ában legfeljebb **2$\sigma$** messze kerül $\mu$-től
  - az esetek **99,7%**-ában legfeljebb **3$\sigma$** messze kerül $\mu$-től
  - ...

# Ismétlés: centrális határeloszlás tétele

- CLT (Central Limit Theorem)
  - A minták statisztikáinak átlaga normális eloszlást követ (bizonyos feltételek mellett).
  - $\bar{x} \sim N\left(mean = \mu, \sigma = \frac{s}{\sqrt{n}}\right)$
    - $\bar{x}$ a mintaátlag
    - $\mu$ a populáció várható értéke
    - $s$ a populáció (empirikus) szórása
    - $n$ a mintaméret

# Konfidenciaintervallumok

- Ha tehát a tetszőleges eloszlású, ***s*** szórású vizsgált jellemzőről ***t*** db (>30) megfigyelést végzünk
- A tapasztalati átlagáról…
  - **68%** biztonsággal kijelenthető, hogy legfeljebb $\boldsymbol{s}/\sqrt{\boldsymbol{t}}$ pontatlansággal becsli ***m*** értékét
  - **95%** biztonsággal $2\boldsymbol{s}/\sqrt{\boldsymbol{t}}$ sugarú intervallumba esik
  - **99,7%** biztonsággal $3\boldsymbol{s}/\sqrt{\boldsymbol{t}}$ sugarú intervallumba esik
- És ***t*** növelésével gyökösen szűkül az intervallum

# Konfidenciaintervallumok

- Ha tehát a tetszőleges eloszlású, $\boldsymbol{s}$ szórású vizsgált jellemzőről $\boldsymbol{t}$ db (>30) megfigyelés égzünk
- A ta gáról…
  - **68%** biztonsággal kijele hogy legfeljebb $\boldsymbol{s}/\sqrt{\boldsymbol{t}}$ pontatl t
  - **95%** biztonsággal $2\boldsymbol{s}/\sqrt{\boldsymbol{t}}$ sugarú intervallumba esik
  - **99,7%** biztonsággal $3\boldsymbol{s}/\sqrt{\boldsymbol{t}}$ sugarú intervallumba esik
- És $\boldsymbol{t}$ növelésével gyökösen szűkül az intervallum

Konfidenciaszint

Egyedi megfigyelés szórása

Konfidenciaintervallum sugara (félszélessége)

# Kísérlettervezés példa

- A várható értékre 30 megfigyelés
  - Tapasztalati átlag: 2,3 s (jó-e ez? kell még mérni?)
  - Tapasztalati szórás: $\boldsymbol{s}$ = 1,1 s
- Cél
  - 99,7%-os konfidenciaintervallum 0,6 s széles legyen
- Kísérlettervezés
  - Elvárt sugár (félszélesség) = 3$\sigma$ = $\boldsymbol{3s}/\sqrt{\boldsymbol{t}}$ < 0,3 s
    - (ez a $\sigma$ az átlag szórása, nem az eredeti mért jellemzőé!)
  - Ezért $t$ = 121 megfigyelés kell legalább
- Hol a csalás?

# Korrekció

- Többnyire a tényleges eloszlás paraméterei *a priori* ismeretlenek (különben minek mérnénk?)
- Így nem használható fel a tényleges ***s*** szórás
- Csak a tapasztalati szórás használható → Gauss/normális helyett Student t-eloszlás
    - (más konfidenciaintervallumok)
- $t \to \infty$ esetén Student → normális
- Ökölszabály: $t$>100 esetén használható a Gauss

# Statisztikai próbák (ízelítő)

- A/B próba
  - ‚A' honlapterv esetén
    - 300 mért látogatóból 4 vásárolt
  - ‚B' honlapterv esetén
    - 50 mért látogatóból 2 vásárolt
  - Mondhatjuk-e, hogy az ‚B' jobb az ‚A'-nál?
    - Szignifikancia: p=25% → több mérés kell még
      - (itt most a magasabb p érték jelzi a szignifikánsabb eredményt)
    - http://web-tracking-guide.com/statistical-significance-calculator.html