Kísérlettervezés alapfogalmak

Rendszermodellezés

Budapesti Műszaki és Gazdaságtudományi Egyetem
Méréstechnika és Információs Rendszerek Tanszék

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

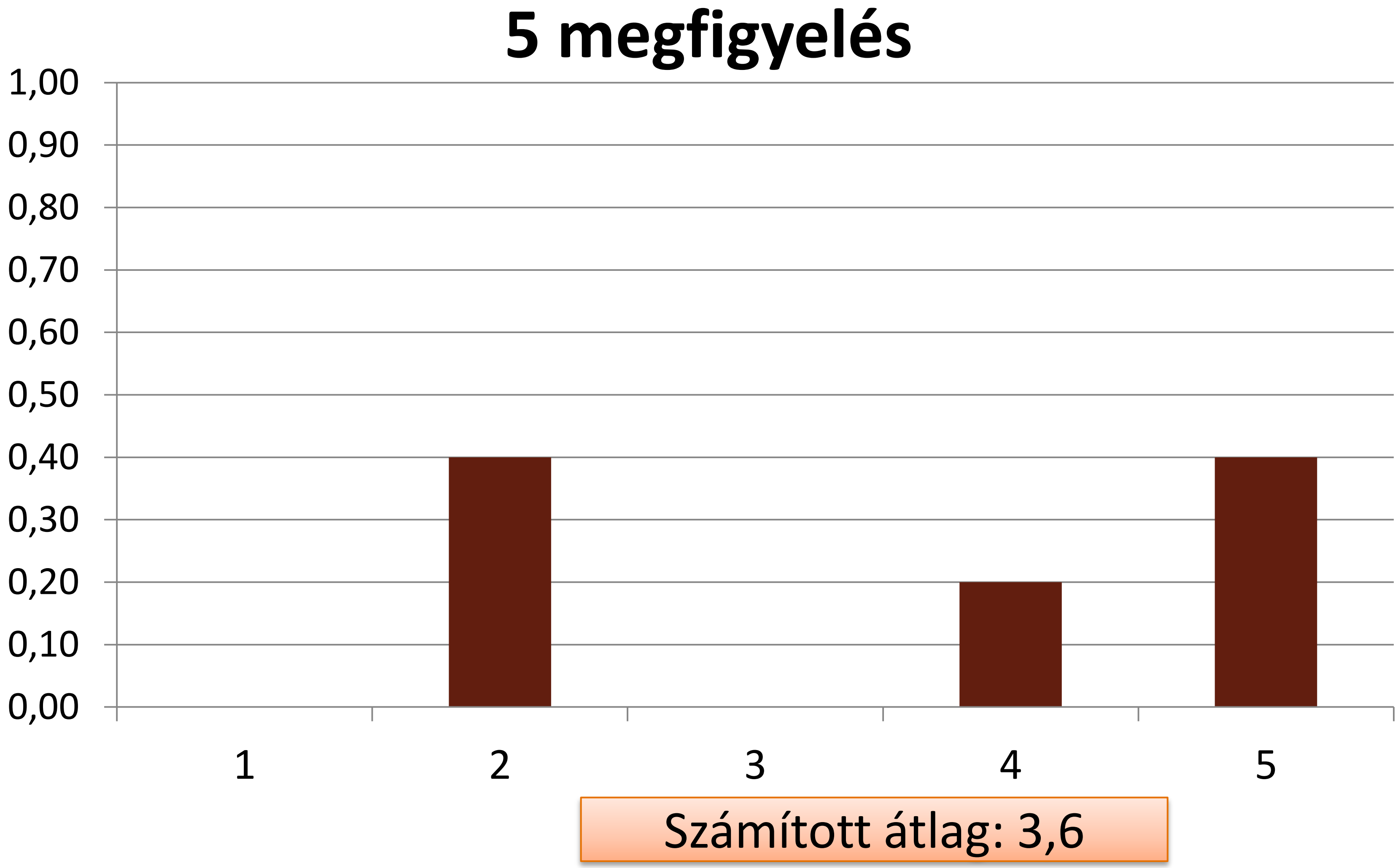
- Mérjük meg egy jellemző értékét! (1..5)

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

Mért érték: 4

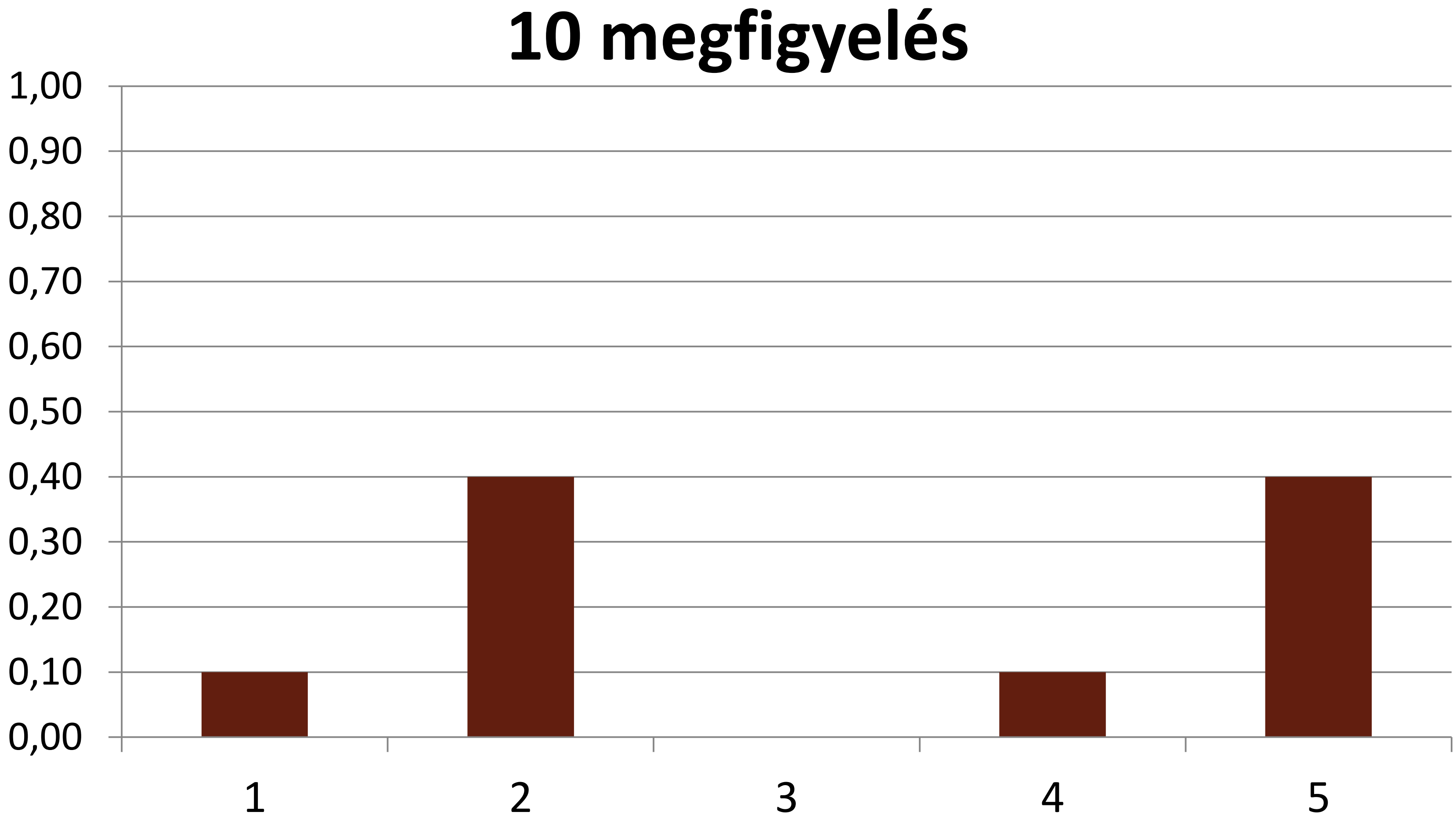
Számszerű jellemzők mérése

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



Számszerű jellemzők mérése

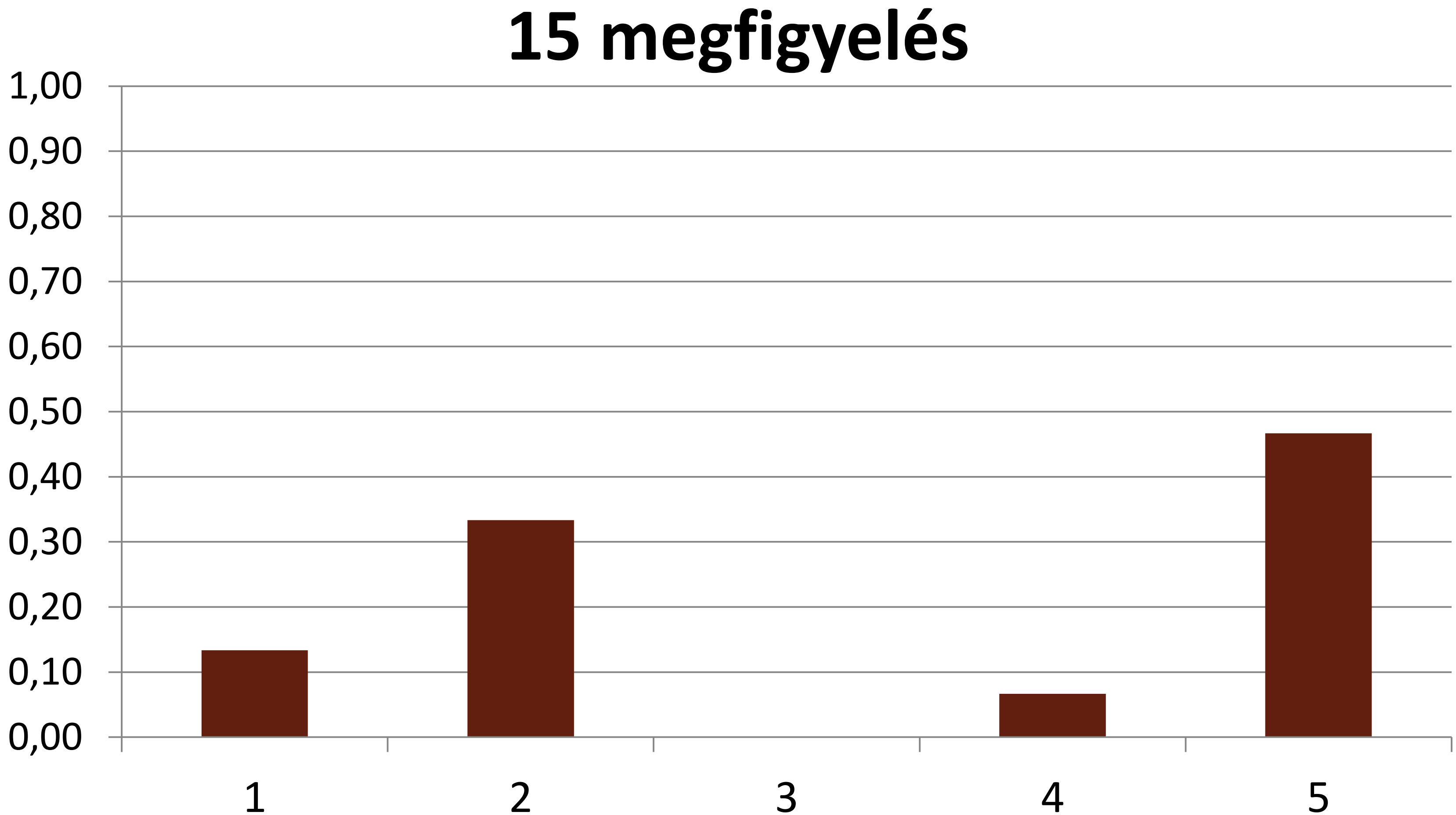
- Mérjük meg egy jellemző értékét! (1..5)



Számított átlag: 3,3

Számszerű jellemzők mérése

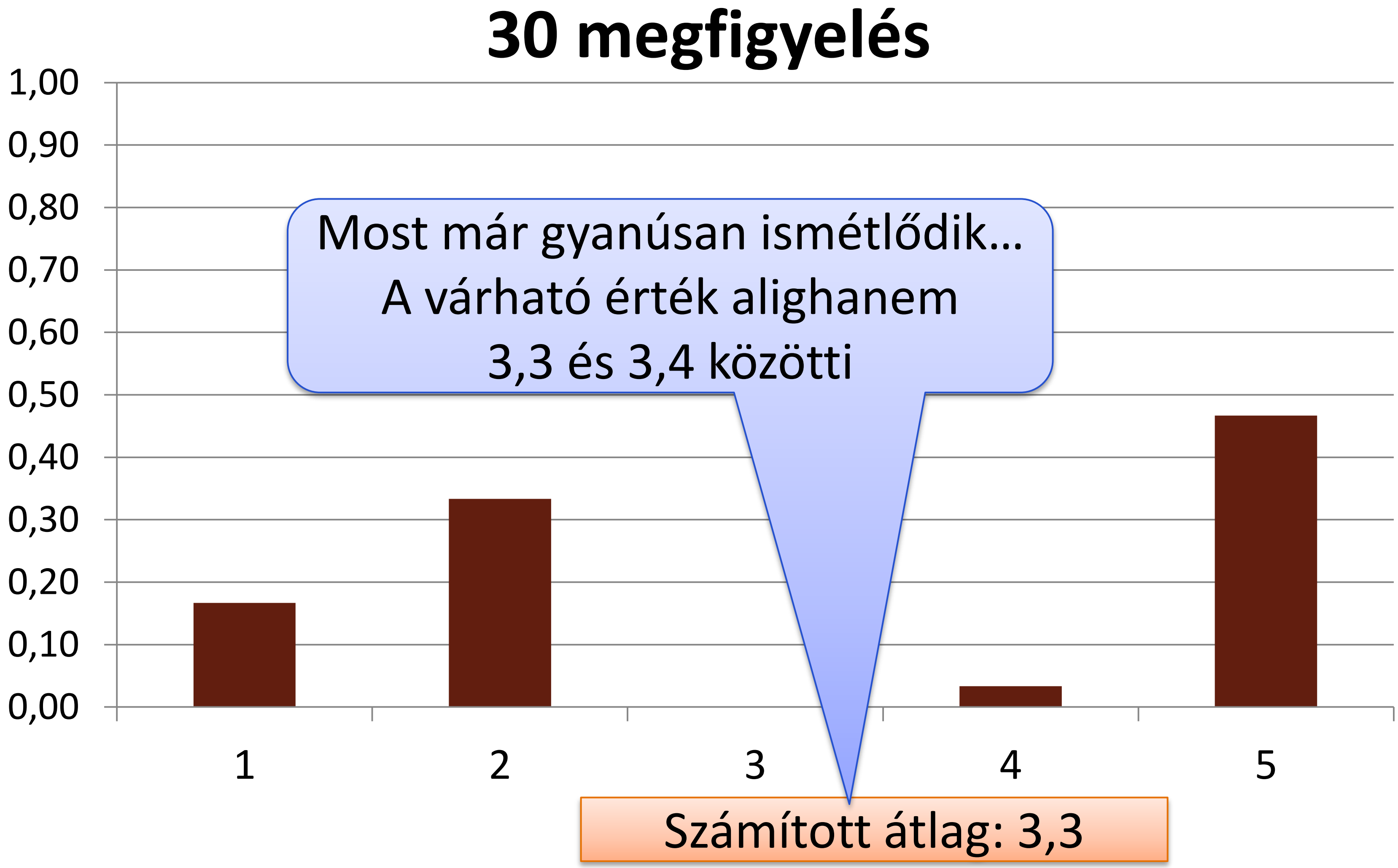
- Mérjük meg egy jellemző értékét! (1..5)



Számított átlag: 3,4

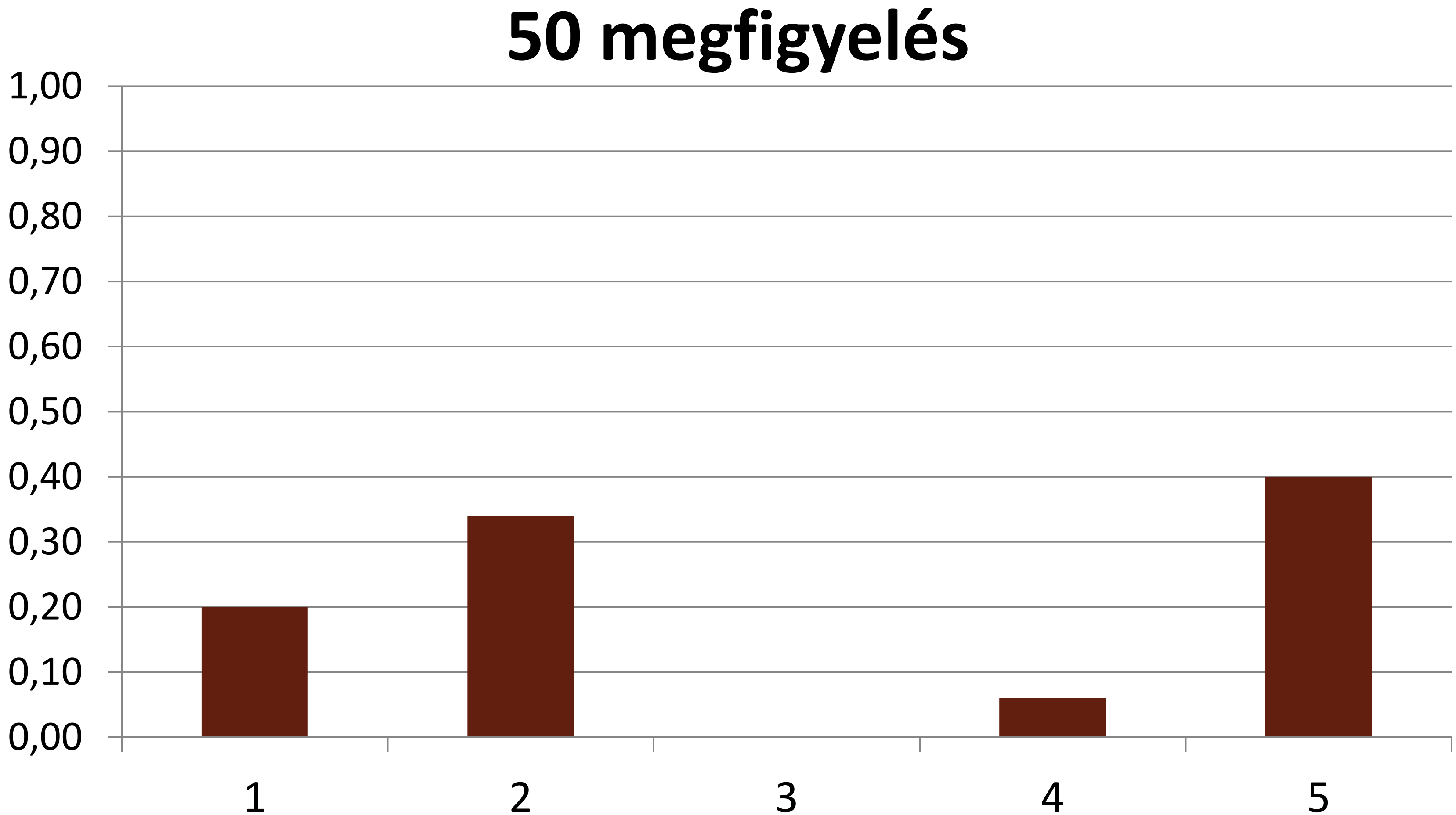
Számszerű jellemzők mérése

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



Számszerű jellemzők mérése

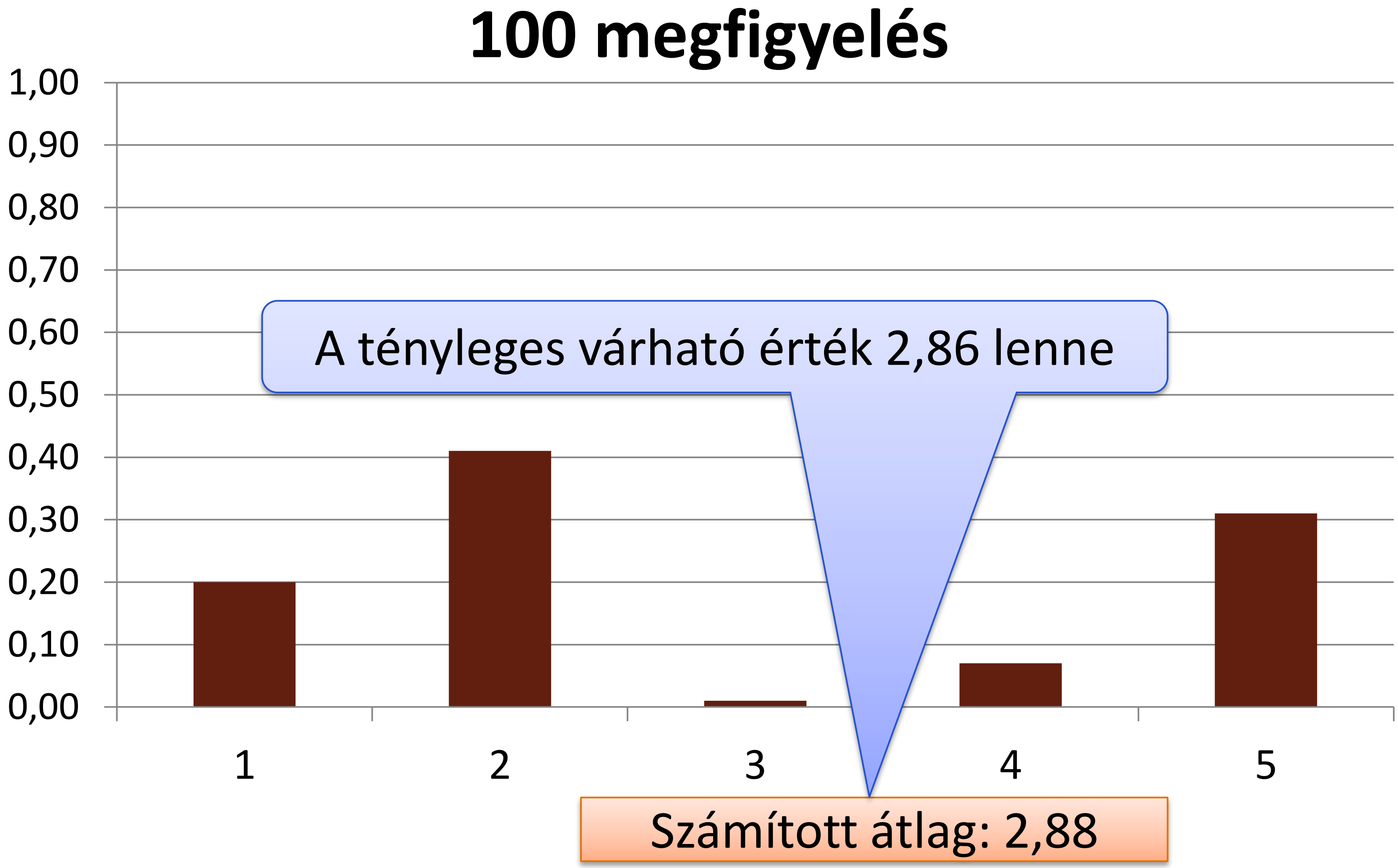
- Mérjük meg egy jellemző értékét! (1..5)



Számított átlag: 3,12

Számszerű jellemzők mérése

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



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

- 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 **d** szórású)
  - tapasztalati átlaga **t**→∞ esetén közelítőleg
  - **normális eloszlású**,
  - μ = ***m*** várható értékkel és σ = $\mathbf{d}/\sqrt{t}$ szórással

Tapasztalati átlag

- Kísérlet = megfigyelések sorozata
- Megfigyelések sorozatának **tapasztalati átlaga**:
  - Egy jellemzőt ***t*** db. független megfigyeléssel mérve,
  - …tt eredmény

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

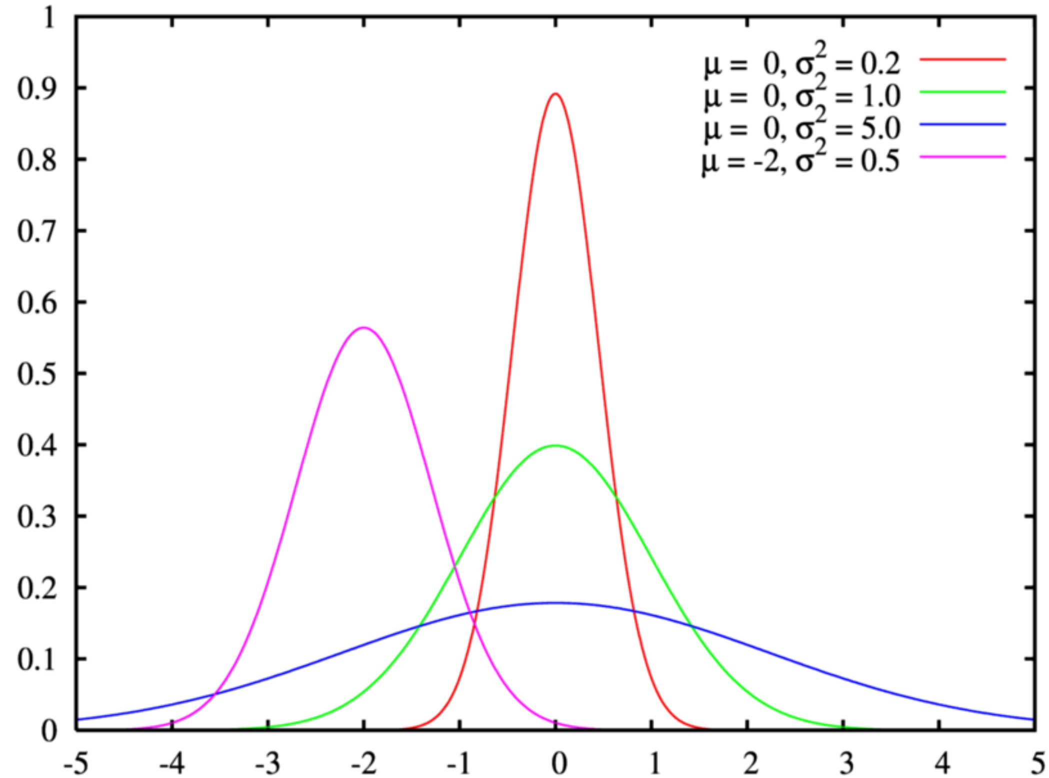
- **C**… következik:
  - (de legyen *véges* ***m*** várható …kű és **d** szórású)
  - tapasztalati átlaga **t**→∞ esetén közelítőleg
  - **normális eloszlású**,
  - μ = ***m*** várható értékkel és $\sigma = {}^{\mathbf{d}}/_{\sqrt{t}}$ szórással

A normális (Gauss) eloszlás

- Valószínűségsűrűség-függvénye: (nem kérdezzük)

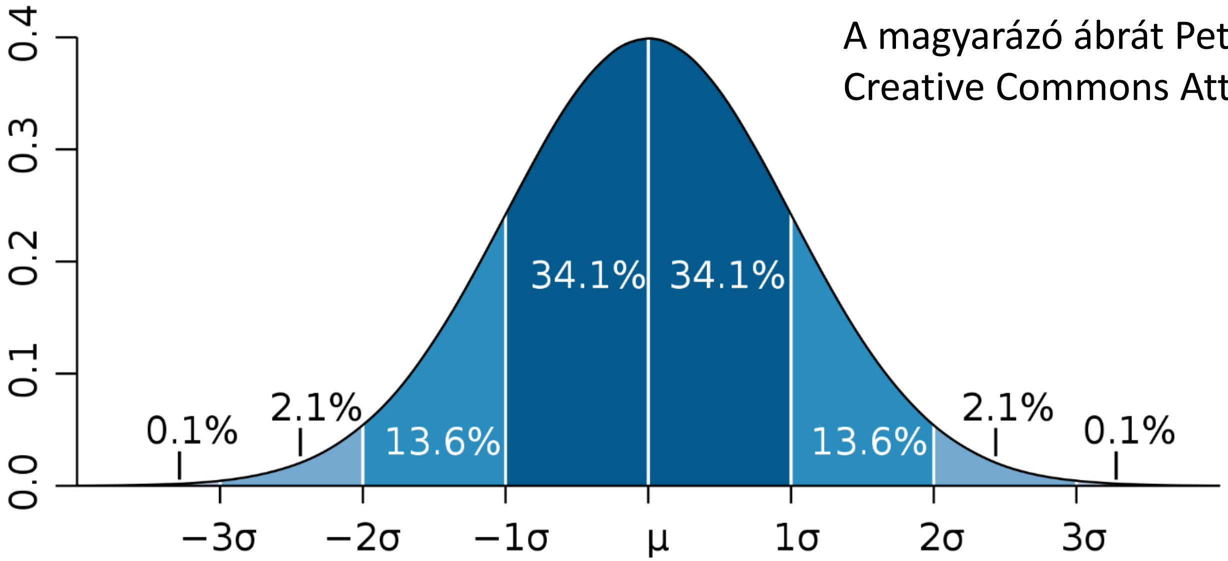
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Paraméterek
  - Várható értéke μ
    - → **m** a mi esetünkben
  - Szórása σ
    - → $\mathbf{d}/\sqrt{t}$ 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 **1σ** messze kerül μ-től
  - az esetek **95%**-ában legfeljebb **2σ** messze kerül μ-től
  - az esetek **99,7%**-ában legfeljebb **3σ** messze kerül μ-től
  - …

Konfidenciaintervallumok

- Ha tehát a tetszőleges eloszlású, ***d*** szórású vizsgált jellemzőről **t** db. (>30) megfigyelést végzünk
- A tapasztalati átlagáról…
  - **68%** biztonsággal kijelenthető, hogy legfeljebb $\mathbf{d}/_{\sqrt{t}}$ pontatlansággal becsli ***m*** értékét
  - **95%** biztonsággal $\mathbf{2d}/_{\sqrt{t}}$ sugarú intervallumba esik
  - **99.7%** biztonsággal $\mathbf{3d}/_{\sqrt{t}}$ sugarú intervallumba esik
- És ***t*** növelésével gyökösen szűkül az intervallum

Konfidenciaintervallumok

- Ha tehát a tetszőleges eloszlású, ***d*** szórású vizsgált jellemzőről **t** db. (>30) megfigyelést végzünk
- A tapasztalati átlagáról…
  - **68%** biztonsággal kijelenthető, hogy
    leg … tatl … t
  - **95%** biztonsággal $\mathbf{2d}/_{\sqrt{t}}$ sugarú intervallumba esik
  - **99.7%** biztonsággal $\mathbf{3d}/_{\sqrt{t}}$ sugarú intervallumba esik
- És ***t*** növelésével gyökösen szűkül az intervallum

Konfidenciaszint

Konfidenciaintervallum sugara (félszélessége)

Kísérlettervezés példa

- A várható értékre 30 megfigyelés
  - Tapasztalati átlag: 2,3 s (jó-e ez? kell még mérni?)
  - Tapasztalati szórás: **d** = 1,1 s
- Cél
  - 99.7%-os konfidenciaintervallum 0,6 s széles legyen
- Kísérlettervezés
  - Elvárt sugár (félszélesség) = 3σ = $\mathbf{3d}/\sqrt{t}$ < 0,3 s
    - (ez a σ 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 ***d*** szórás
- Csak a tapasztalati becslés használható → Gauss/normális helyett Student t-eloszlás
  - (más konfidenciaintervallumok)
- t→∞ esetén Student → normális
- Ökölszabály: t>100 esetén használható a Gauss

Statisztikai próbák (ízelítő)

- Hipotézisek
  - **Nullhipotézis**: a világ X módon működik
  - **Alternatív hipotézis**: a világ nem X módon működik
- Kísérlet eredménye: Z
  - Ha az eltérés X-től nem írható a véletlen számlájára, akkor **szignifikáns** az eredmény
- *p*-próba (leegyszerűsítve)
  - *p* = **Pr**{X esetén a kísérletnek Z lesz az eredménye}
  - Ha $p \ll 1$, akkor a nullhipotézist cáfoltuk
- A nullhipotézis sosem igazolható
  - Vagy szignifikánsan cáfoltuk, vagy nem

Statisztikai próbák (ízelítő)

- Hipotézisek
  - **Nullhipotézis**: a világ X módon működik
  - **Alternatív hipot**… a világ nem X módon működik
- Kísérlet e…: Z
  - …nlájára,

> Pl.
> X és Y jellemző független valószínűségi változók;
> vagy a vállalt SLA be van tartva

- *p*-próba (leegyszerűsítve)
  - *p* = **Pr**{X esetén a kísérletnek Z lesz az eredménye}
  - Ha p << 1, akkor a nullhipotézist cáfoltuk
- A nullhipotézis sosem igazolható
  - Vagy szignifikánsan cáfoltuk, vagy nem

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