

# A véletlen szerepéről

Ferenci Tamás  
tamas.ferenci@medstat.hu

2018. május 16.

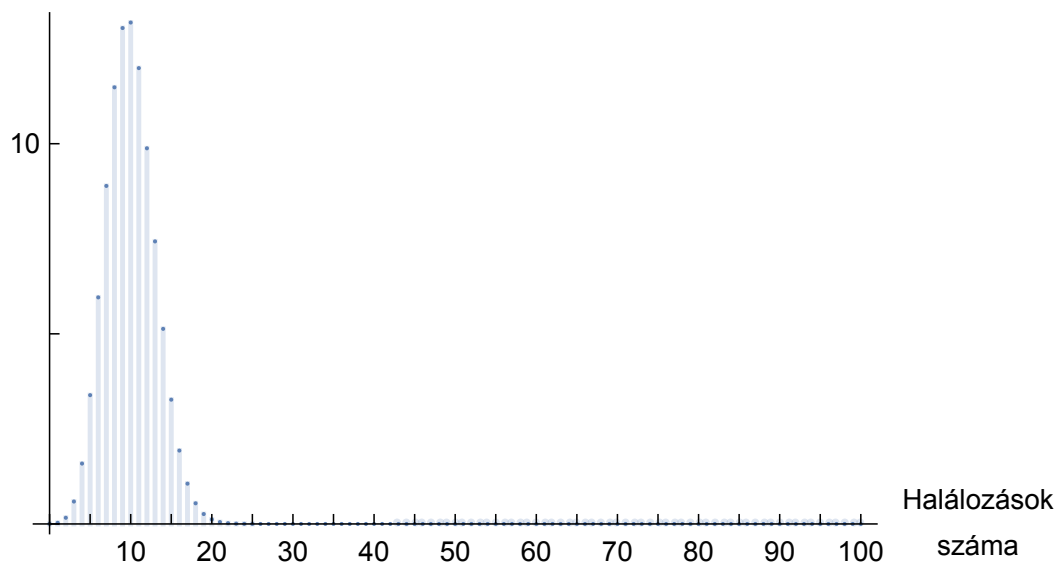
## A véletlen megjelenése az orvosi kutatásokban

- A biológia nem (vagy ritkán) determinisztikus (egyik csoportban sem: sem az exponáltaknál, sem a nem exponáltaknál)
- Még ha hat is a gyógyszer, nem fog *minden* gyógyszert szedőnek csökkenni – és *ugyanannyit* – a vérnyomása (míg a többieknek marad állandó), még ha okoz is ilyen mellékhatást, nem lesz *minden* gyógyszert szedőnek hasmenése (míg a többiek közül senkinek), még ha van is hatása a távvezetéknek, nem lesz *mindenki* rákos a környékén (míg távvezetékétől távol senki) stb.
- Ha ez nem így lenne, az összes kérdés megválaszolható lenne két ember összehasonlításával...!
- A sztochasztikusság azt jelenti, hogy *ingadozni* fog a kimenet
- Az is ingadozás, hogy „a betegek 10% valószínűséggel halnak meg 30 napon belül”!

## A véletlen ingadozás problémaköre

Attól mert 10% a halálozási arány egy csoportban, meghalhat 100-ból 9 vagy épp 11 (sőt, 5 vagy épp 15) is:

Valószínűség [%]



### Egy szemléletes kép

- Egy lehetséges elképzelési mód, ami a későbbiek szempontjából is hasznos lesz, az urnamodell:
  - Van egy nagyon-nagyon-nagyon nagy urnánk, benne a golyók 10%-ára ráírva, hogy „meghalt”, 90%-ára, hogy „túlél”
  - A klinikai kutatás az, hogy a jól megkevert urnából vakon kihúzzunk 100 golyót
- Képzeltben megismételjük az egészet: visszadobjuk a golyókat, jól átkeverjük az urnát, és kezdjük előlről
- Ezt jó sokszor megismételjük: ekkor az előbbi ábra fogja mutatni, hogy milyen arányban fordulnak elő az egyes esetek (a „csupa meghaltot húzzunk”-tól a „csupa túlélőt húzzunk”-ig)

### Egy szemléletes kép

- Azt fogjuk mondani, hogy...
  - ...az urna a *sokaság* („a valóság”), amire kíváncsiak vagyunk, csak épp nem tudjuk közvetlenül megnézni – jelen esetben a 10%
  - ...a kivett 100 golyó a *mintá*, a szemüveg amin keresztül a sokaságot látjuk, csak épp ez nem mindig a valódi képet mutatja – jelen esetben az előbbi ábrán látott eloszlás
- A képzeletbeli újra-mintavétel szemlélete épp azért jó, mert érthetővé teszi ez utóbbit, hogy a mintából kapott érték ingadozik (noha a valódi érték adott, állandó – csak mi nem tudjuk, hogy mi)
- Úgy is fogjuk mondani: a mintából kapott érték mintavételről-mintavételre ingadozik – egyrészt attól függ, hogy mi a valódi érték, másrészt attól, hogy pont hogyan vettük a mintát (nem úgy érteve, hogy bármilyen értelemben „rosszul” vettük a mintát)

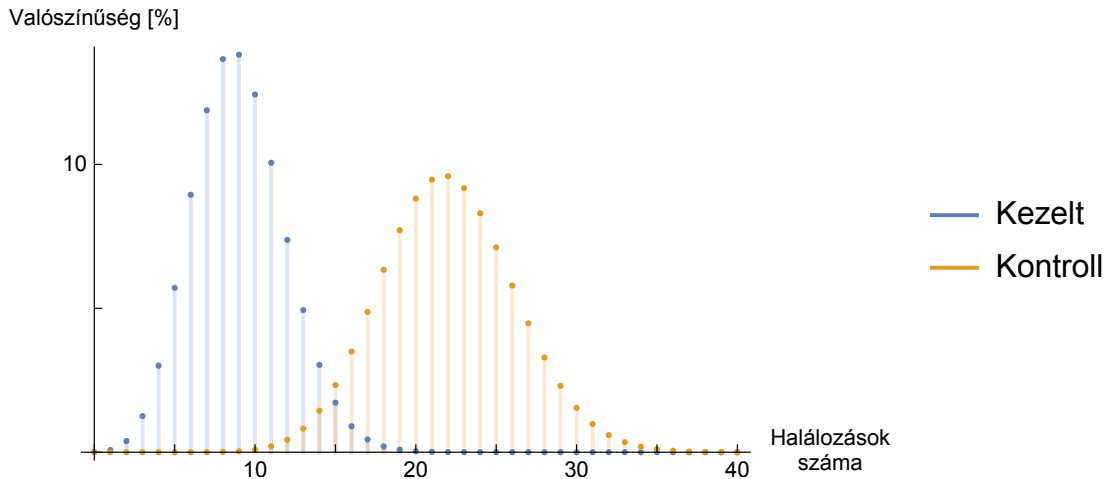
- Ez utóbbi tehát a mintavételi ingadozás, az ebből fakadó hiba a mintavételi hiba
- Kikerülhetetlen, de a jellegéről tudunk nyilatkozni

**A véletlen ingadozás problémaköre**

Akkor, ha a 100 kezeltből 10 halt meg, a 100 kontrollból pedig 20...

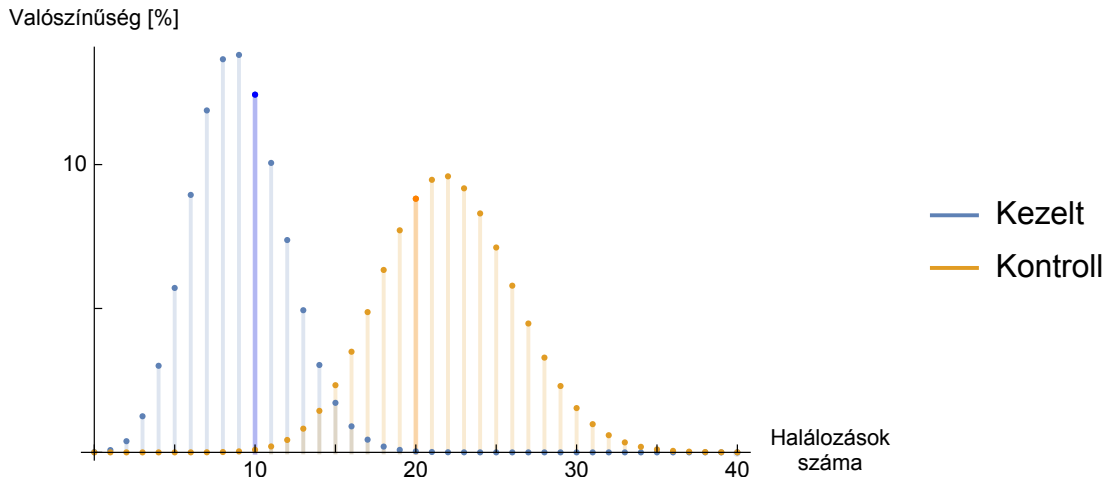
**A véletlen ingadozás problémaköre**

...hat a gyógyszer? (9%-ból 10, 22%-ból 20 a véletlen ingadozás miatt)



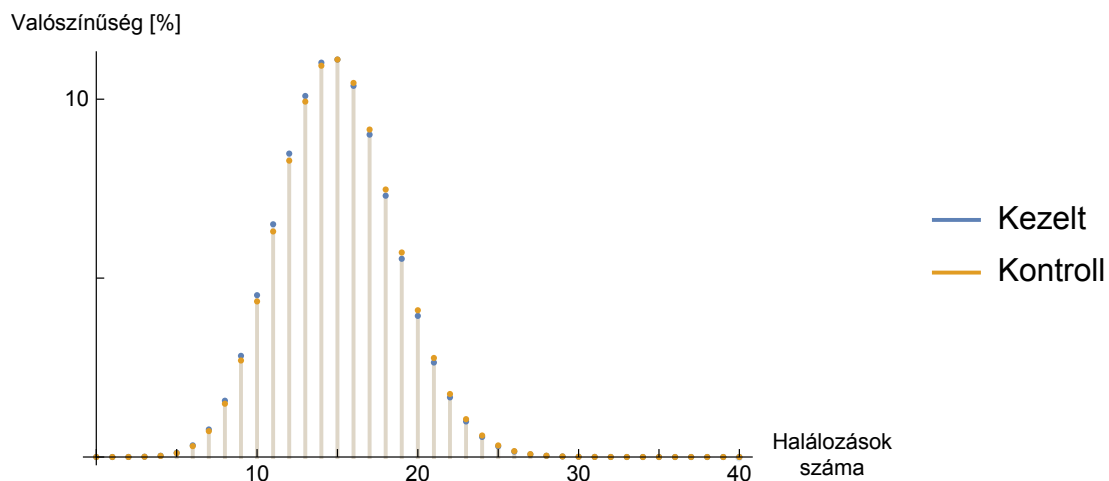
**A véletlen ingadozás problémaköre**

...hat a gyógyszer? (9%-ból 10, 22%-ból 20 a véletlen ingadozás miatt)



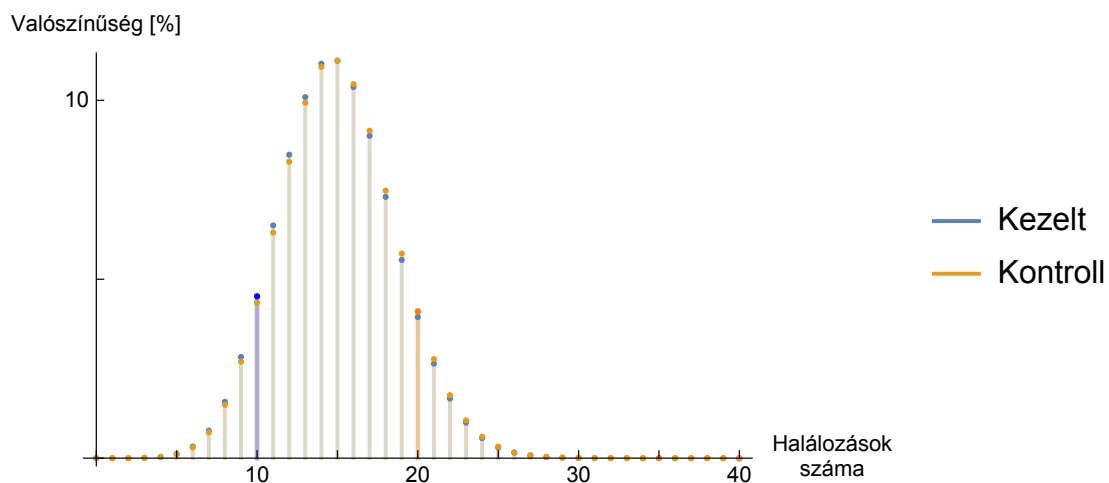
### A véletlen ingadozás problémaköre

...nem hat a gyógyszer? (15%-ból 10, 15%-ból 20 a véletlen ingadozás miatt)



### A véletlen ingadozás problémaköre

...nem hat a gyógyszer? (15%-ból 10, 15%-ból 20 a véletlen ingadozás miatt)



### A véletlen ingadozás problémaköre

- Mi csak annyit látunk, hogy 10 és 20 halálozás – mindkét eset lehetséges!
- Biztosan tehát nem tudunk válaszolni a „hat a gyógyszer?” kérdésre...
- ...de valószínűségi alapon igen!
- Ehhez ismernünk kell az ingadozás matematikai törvényszerűségeit

## Mintavételi ingadozás és a mintanagyság

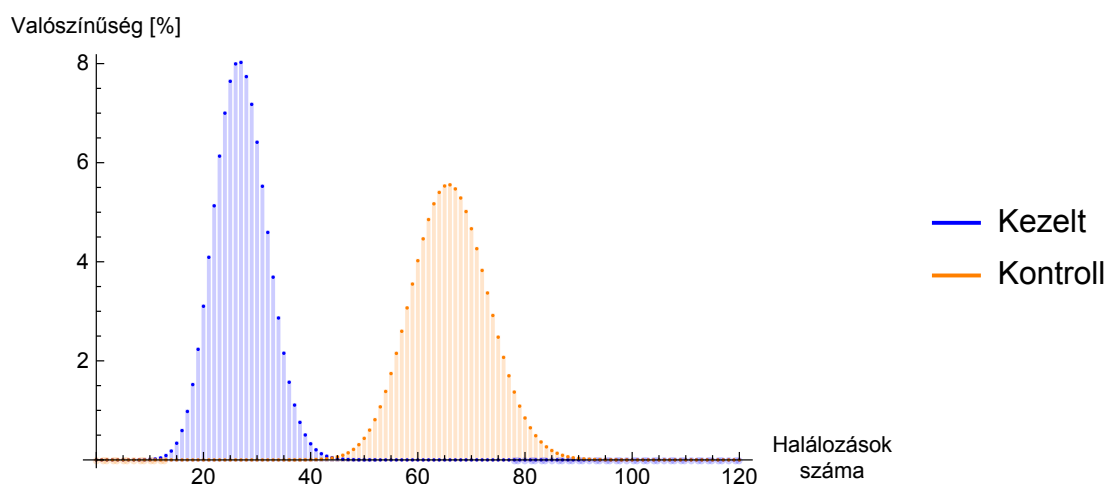
- Elkerülhetetlenül van ingadozás, és időgépünk sincs – mi a megoldás?
- Emberek helyett *csoportokat* kell vizsgálnunk, arra építve, hogy a nem-determinisztikusságból adódó véletlen ingadozások a „kiátlagolódás” révén csökkenjenek!
- Játsszuk még egyszer újra az előbbi eseteket, de most 100-100 beteg helyett 300-300 beteggel
- Azaz minden arány változatlan, csak a számok háromszorozódnak meg

## A véletlen ingadozás problémaköre

Akkor, ha a 300 kezeltből 30 halt meg, a 300 kontrollból pedig 60...

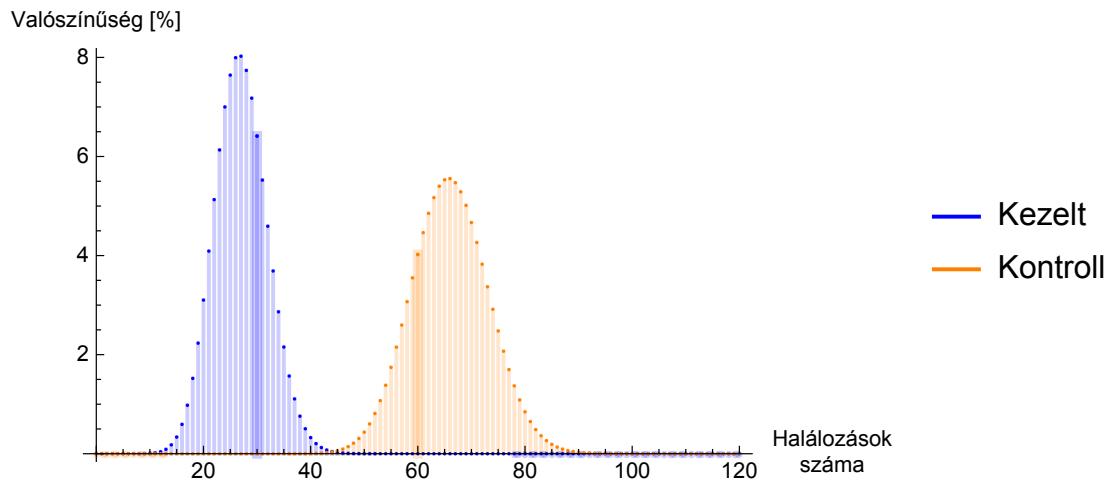
## A véletlen ingadozás problémaköre

...hat a gyógyszer? (9%-ból 30, 22%-ból 60 a véletlen ingadozás miatt)



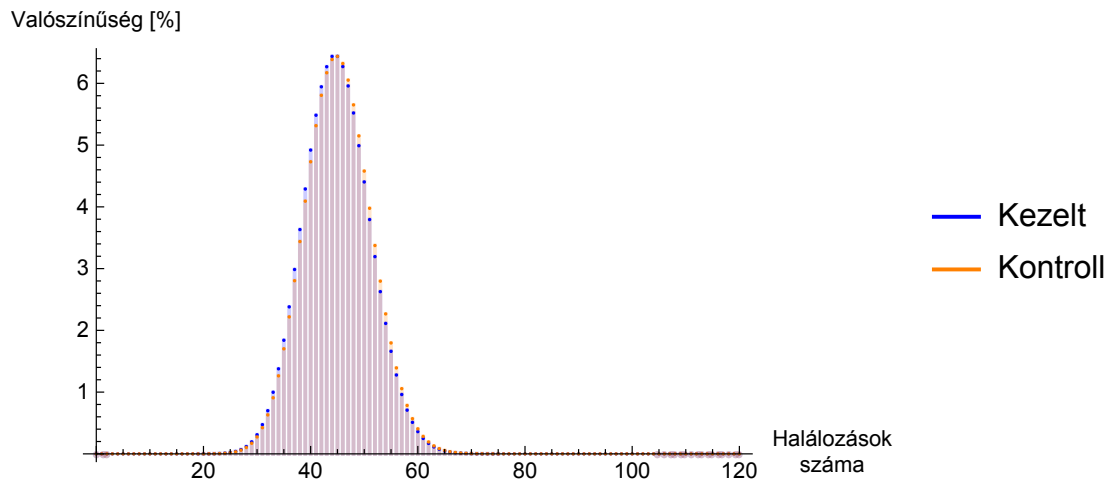
## A véletlen ingadozás problémaköre

...hat a gyógyszer? (9%-ból 30, 22%-ból 60 a véletlen ingadozás miatt)



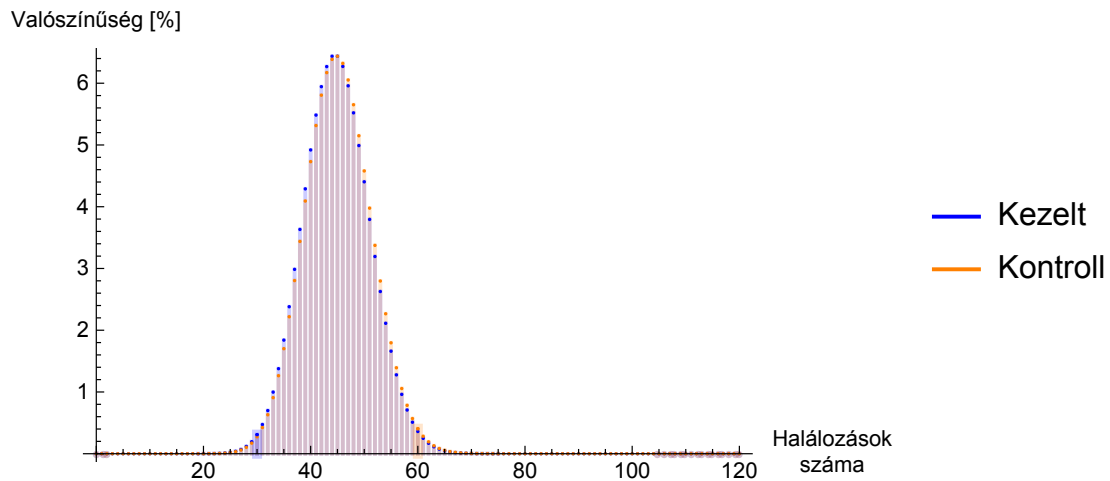
**A véletlen ingadozás problémaköre**

... nem hat a gyógyszer? (15%-ból 30, 15%-ból 60 a véletlen ingadozás miatt)



**A véletlen ingadozás problémaköre**

... nem hat a gyógyszer? (15%-ból 30, 15%-ból 60 a véletlen ingadozás miatt)



### A következtető statisztikai apparátus lényege

- A feladat: tapasztalva egy adott hatást a kutatásban, például mennyire változott a halálzási arány, megmondani, hogy ez betudható-e egyszerűen a mintavételi ingadozásnak, tehát a valóságban nem hat a gyógyszer, vagy olyan nagy, hogy azt feltételezzünk, hogy valódi különbség van mögötte (tényleg hat a gyógyszer)
- Úgy fogjuk mondani: (statisztikai értelemben) lényeges, vagy más szóval *szignifikáns* a tapasztalt hatás
- Persze az, hogy mire mondjuk, hogy „ez még betudható” és mire, hogy „ez már nem”, döntés kérdés: limitet kell valahol húzni
- Ez a limit lesz a kompromisszum: minél szigorúbbak vagyunk, annál valószínűbb, hogy amire azt mondjuk, hogy hat, az tényleg hat a valóságban is, de cserében annál több valóságban hatóra is azt fogjuk mondani, hogy nem hat; e kettő között kell egyensúlyozni
- Ez lesz majd az ún. *szignifikanciaszint*

### Mintavételi ingadozás és a mintanagyság

- Most gondoljunk az előbbi megállapításunkra, hogy ti. a mintavételi hiba, bár soha nem szüntethető meg teljesen, de a mintanagyság növelésével – minden mást változatlanul tartva – csökkenthető
- Tehát minél nagyobb a mintanagyság, adott hatást annál nagyobb valószínűséggel tudunk észrevenni, hiszen az egyre kisebb mintavételi ingadozás mellett egyre biztosabb, hogy az adott, rögzített méretű hatásra már azt mondjuk, hogy nem a mintavételi ingadozásnak tudható be
- A kutatás *erejének* nevezzük annak valószínűségét, hogy ha van egy adott nagyságú hatás, akkor azt észre is vesszük (értsd: szignifikánsnak bizonyul)
- A dolog természetesen fordítva is elmondható, ez csak megfogalmazás kérdése: azt is mondhatjuk, hogy minél nagyobb a mintanagyság, adott valószínűséggel annál kisebb hatást is észre tudunk venni

## Mintanagyság tervezés

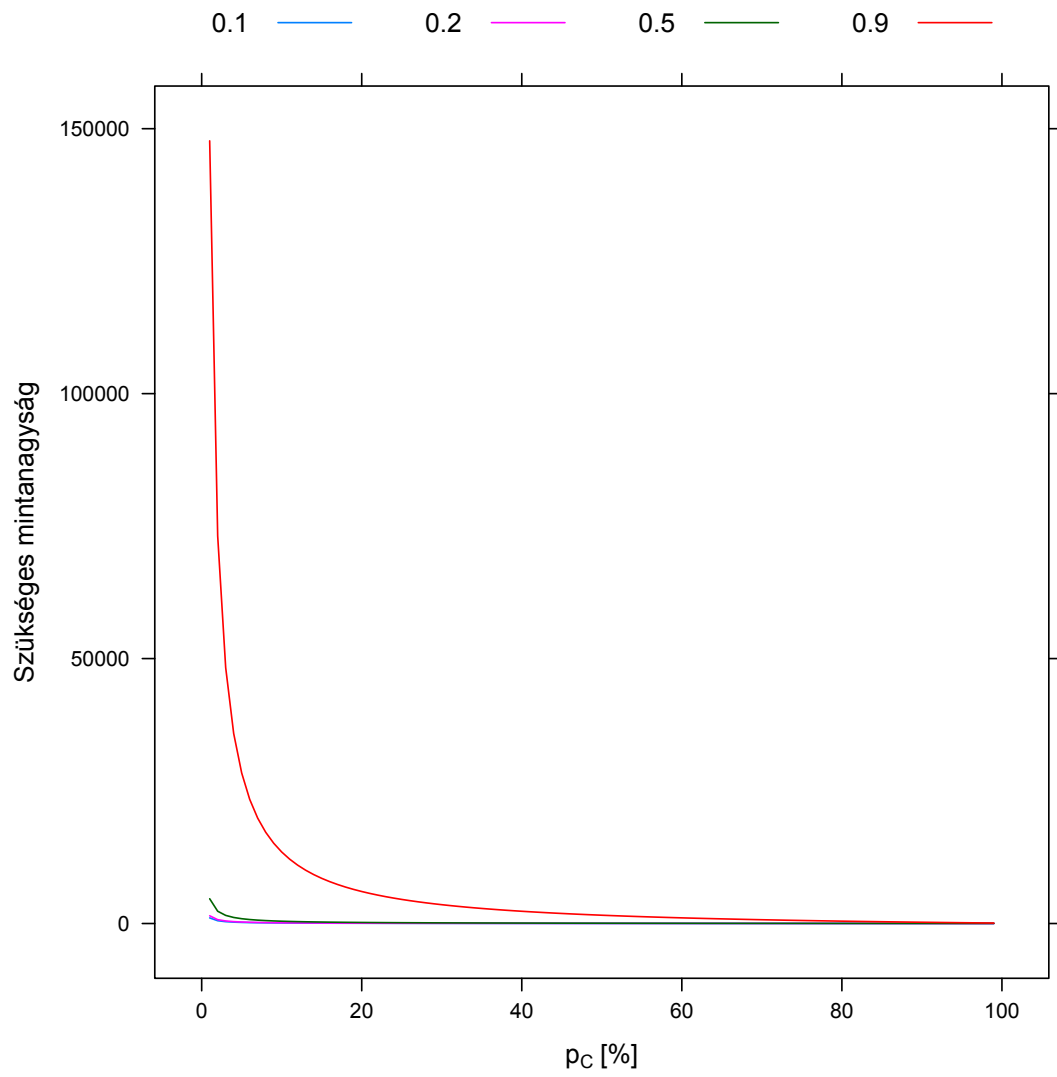
- Ez lehetővé teszi a mintanagyság racionális tervezését
- Ha megadjuk a mintanagyságot és pár egyéb paramétert, kiszámolható, hogy mekkora az erő adott hatás kimutatásához...
- ... fordítsuk ezt most meg!
- Megadva, hogy mekkora hatást akarunk kimutatni, és mekkora erővel, kiszámolható, hogy ehhez mekkora minta kell
- (Furcsa lehet, hogy előre meg kell adnom a hatást, hiszen a kutatás célja épp ennek kiszámolása, de ez olyan mint a csillagászat: minél halványabb csillagot akarok kimutatni, annál nagyobb távcső kell – úgyhogy amikor távcsövet veszünk, kell valamilyen feltételezéssel élni a csillag fényességéről)

## Mintanagyság tervezésének szemléltetése

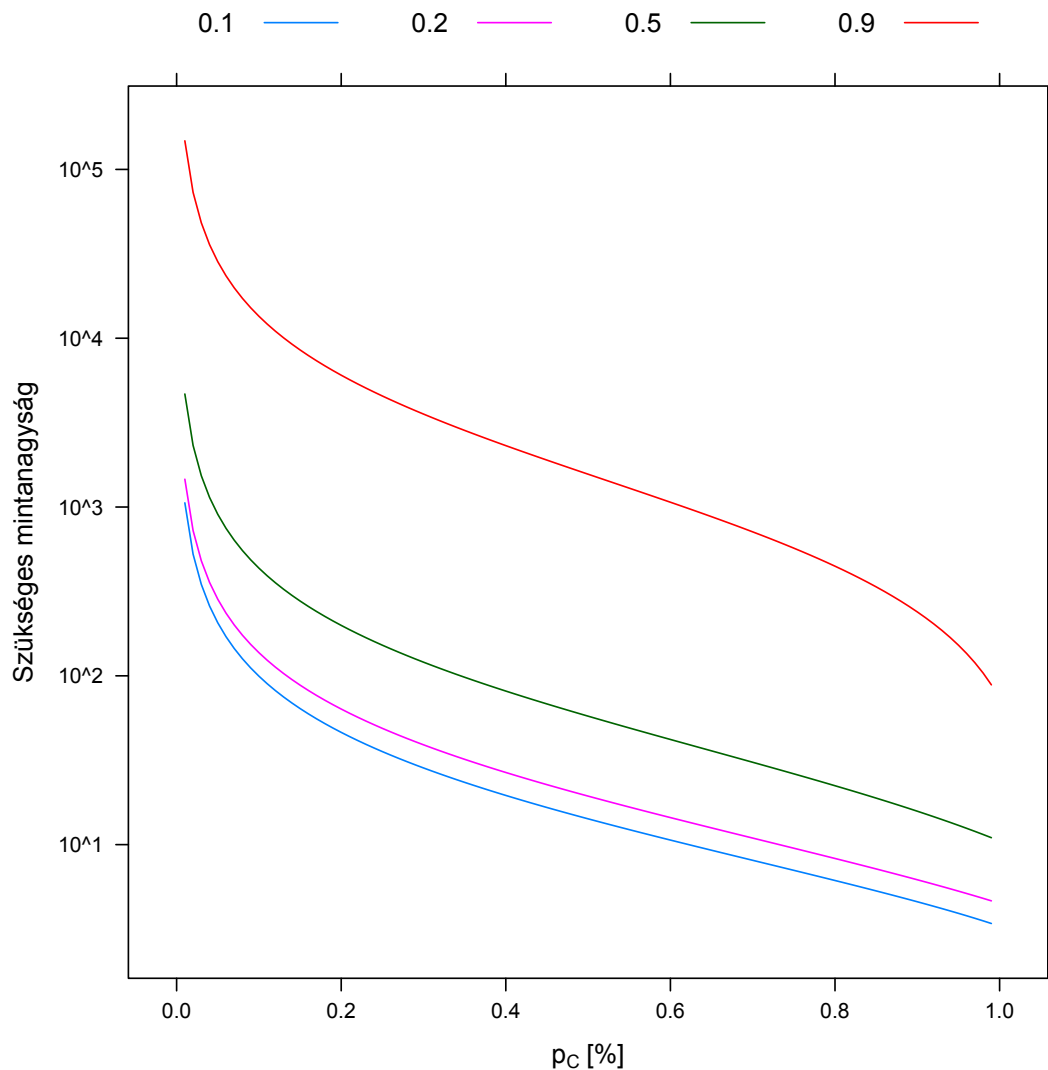
- Kiindulópont: annál könnyebb szignifikánsnak kimutatni valamit, minél nagyobb az abszolút különbség is a csoportok között (még ha a relatív különbség állandó is!)
- Kis szemléltetés: a kontrollcsoportban  $p_C$  a kockázat, a kezelés ezt 10, 20, 50, 80 illetve 90%-kal csökkenti
- Mekkora mintanagyság kell ennek a kimutatásához?
- (Szokásosan  $\alpha = 0,05$  szignifikanciaszint,  $\beta = 0,8$  erő)

## Szükséges mintanagyság





**Sziükséges mintanagyság**



### Mintavételi hiba (variancia)

- A mintavételi hibára azt szokták mondani: *variancia* típusú hiba
- A becült érték eltérhet a valóságtól, de ha sokszor megismételnénk a mintavételt, akkor *átlagosan* nem lenne eltérés: ingadozik ugyan az eredményünk, de legalább a jó érték körül
- ... és minél nagyobb a mintanagyság, annál kevésbé
- Tehát ez a fajta hibázás nem bináris, nem arról van szó, hogy van ilyen vagy nincs, hanem számszerűen lemérhető: *mennyire* ingadozik a mintából kapott érték annak átlaga körül (pl. mekkora a szórása: ez az ún. standard hiba)
- Azt, hogy milyen kicsi ez a variancia szokták a vizsgálat precizitásának is nevezni

- Jól bevált apparátussal – következtető statisztika:  $p$ -értékek, konfidenciaintervallumok stb. – vizsgálható jellemezhető

### Nem-mintavételi hiba (torzítás)

- Minden, ami nem a mintavételi ingadozásból fakad
- ... úgyhogy egyet már nagyon is láttunk eddig is: épp ilyen a confounding is!
- Egy másik tipikus példa: mi van, ha a mintát nem olyan tökéletesen választjuk (azaz nem lehet belőle következtetni a sokaságra)?
- Általános jellemzője az ilyen hibáknak, hogy még ismételt mintavétellel sem tartanak a nullába, azaz a mintából kapott értékeknek az átlaga sem a valódi érték
- Tehát: nem is a jó érték körül ingadozik az eredmény!
- Az ilyen típusú hibákat nevezzük *torzításnak* (lemérhető az átlag és a valódi érték különbségével, ennek neve az ellenség megtévesztése végett szintén torzítás)
- Sokszor keveset beszélnek róla az orvosi gyakorlatban, pedig rettentő fontos tud lenni...

### Variancia és torzítás

- Nagyon nagy általánosságban beszélve a torzítástól félünk jobban
- Lehet a kettőt együtt is mérni: mennyire ingadozunk a jó érték körül
- Ebben benne van mindkét komponens: mennyire ingadozunk az átlag körül (variancia) és az átlag hol van a jó értékhez képest (torzítás)!
- Neve átlagos négyzetes hiba (belátható, hogy értéke a torzítás és a variancia négyzetének összege)

### Hibák vizsgálata

- Mindezek persze elméleti kérdések, olyan értelemben, hogy egyetlen minta alapján úgysem tudjuk megmondani, hogy a hibázás variancia vagy torzítás-e
- Ezért is mondtuk, hogy „képzeltbeli ismételt mintavételek” kontextusában értelmezhető
- Szebben szólva ezek vizsgálatához mind valamilyen modellt kell feltételeznünk

### A nem-mintavételi hibák szerepe

- Ki szokta úgy érezni, hogy egy adott gyógyszergyár által fizetett kísérletben mindig pont az ő gyógyszere lesz a jobb? (Pl. lelkes kutatók végignézték a 4 vezető pszichiátriai újság 1992-2002 közötti összes cikkét: amikor a kísérlet szponzora a vizsgált gyógyszer gyártója volt, akkor 78% volt a sikerarány, függetlenül 48%, kompetitor által fizetettnél 28%...)
- De ez meg hogyan lehet egyáltalán? (hiszen ugyanaz a gyógyszer valakinek a saját terméke, valakinek meg a kompetitoré!)

- „Miért jobb az olanzapin mint a riszperidon, a riszperidon mint a kvetiapin és a kvetiapin mint az olanzapin?”
- Ezek tipikusan nem véletlen ingadozások, nem is direkte meghamisított eredmények, hanem a – többé vagy kevésbé jóhiszemű – nem-mintavételi hibák eredményei!

Kelly RE Jr, Cohen LJ, Semple RJ, et al. Relationship between drug company funding and outcomes of clinical psychiatric research. *Psychol Med.* 2006 Nov;36(11):1647-56. Heres S, Davis J, Maino K, et al. Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: an exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. *Am J Psychiatry.* 2006 Feb;163(2):185-94.

### A mintavételi hibák kézbentartása következtető statisztikával

- Fontos már most hangsúlyozni, hogy a mintavételi hibák kezelésére bevált apparátus ( $p$ -érték, konfidenciaintervallumok stb.) az *égegyadta világon semmit de semmit* nem mond a nem-mintavételi hibákról!
- (Miközben simán lehet, hogy egy orvosi kutatás domináns hibaforrását ez jelenti!)

### Összefoglalva

Mintavételi hiba (elkerülhetetlen de kézbentartható, variancia típusú hiba, mintanagysággal csökken) vs. nem-mintavételi hiba (elkerülhető lehet de nehezen karakterizálható, torzítás típusú hiba, mintanagysággal nem feltétlenül csökken)