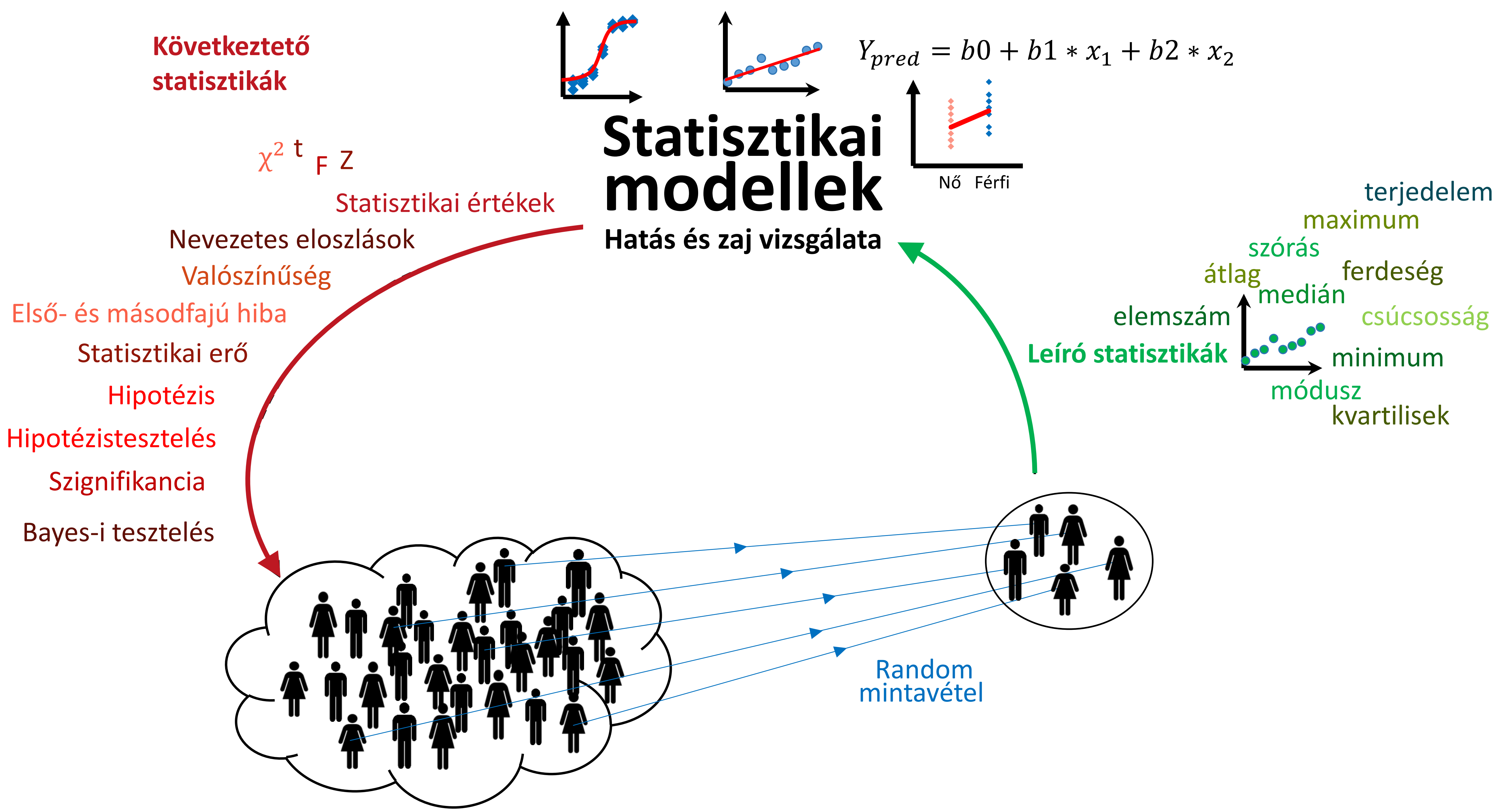
stathelp.hu

Készítette: Soltész-Várhelyi Klára

Következtető statisztika



Következtető statisztikák
χ^2 t F Z
Statisztikai értékek
Nevezetes eloszlások
Valószínűség
Első- és másodfajú hiba
Statisztikai erő
Hipotézis
Hipotézistesztelés
Szignifikancia
Bayes-i tesztelés
Statisztikai modellek
Hatás és zaj vizsgálata
$Y_{pred} = b0 + b1 * x_1 + b2 * x_2$
Nő Férfi
terjedelem
maximum
szórás
átlag
ferdeség
medián
elemszám
csúcsosság
Leíró statisztikák
minimum
módusz
kvartilisek
Random mintavétel

Leíró és következtető statisztika

- **Leíró statisztika** (descriptive stats)
  - Célja a mért adathalmaz jellemzése, az rendelkezésre álló információ tömörítése
  - A mért mintát írja le, következtetéseket nem tartalmaz

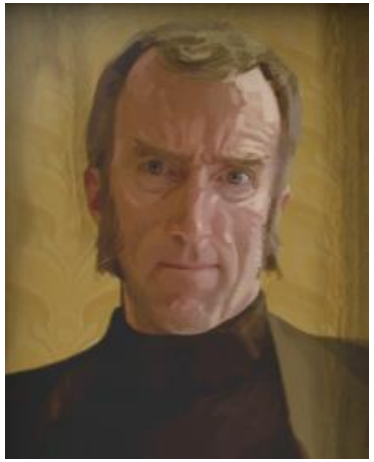
- **Következtető statisztika** (inferential stats)
  - Célja a populációból választott mintából a populációra való (vissza)következtetés
  - Valószínűségekkel dolgozik
  - Fogalmak:
    - Statisztikai modellek és paraméterek becslése
    - Statisztikai értékek
    - Eloszlások, és a hozzájuk tartozó valószínűségek
    - (Konfidencia intervallumokról újra)
    - Klasszikus hipotézisvizsgálat, és ami ehhez kell: hipotézis, első és másodfajú hiba, hiba becslése, szignifikancia, elemszámbecslés
    - Bayes-i statisztika
    - (A hatásnagyságot itt beszéljük meg, bár nem szigorú értelemben véve nem tartozik a következetető statisztikák közé)

- A kettő között nincs éles határ

Statisztikai modellek

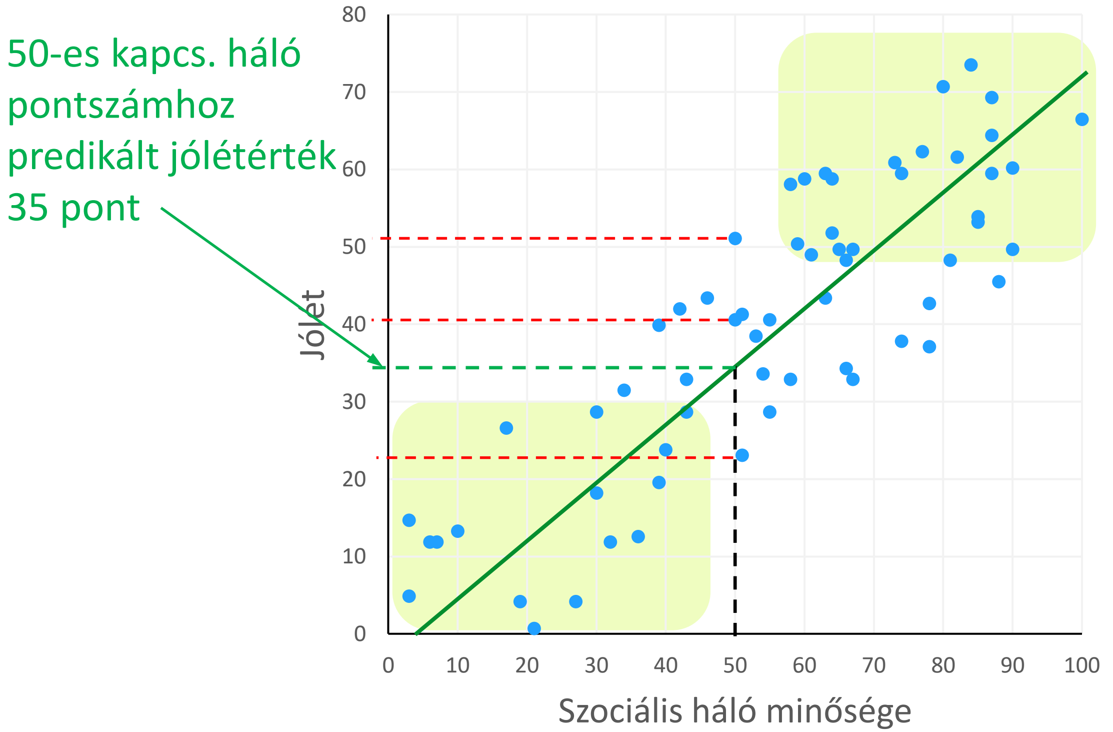
„I'll be honest, we're throwing science at the wall here to see what sticks.”

Cave Johnson



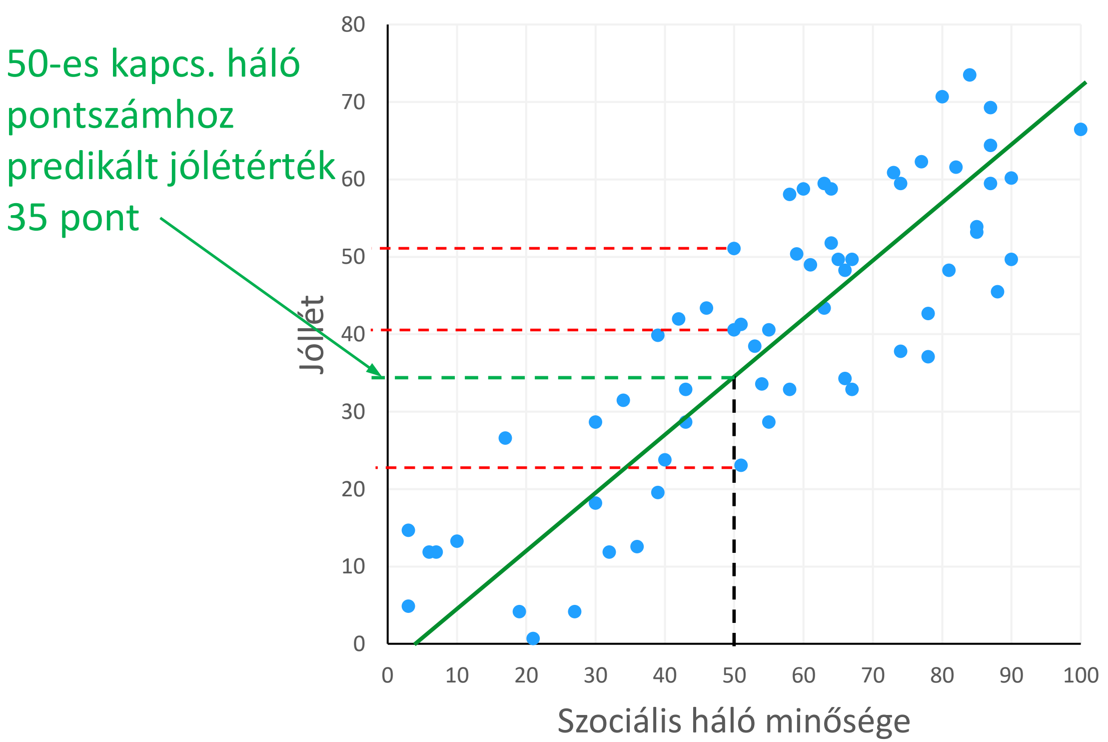
Modellek

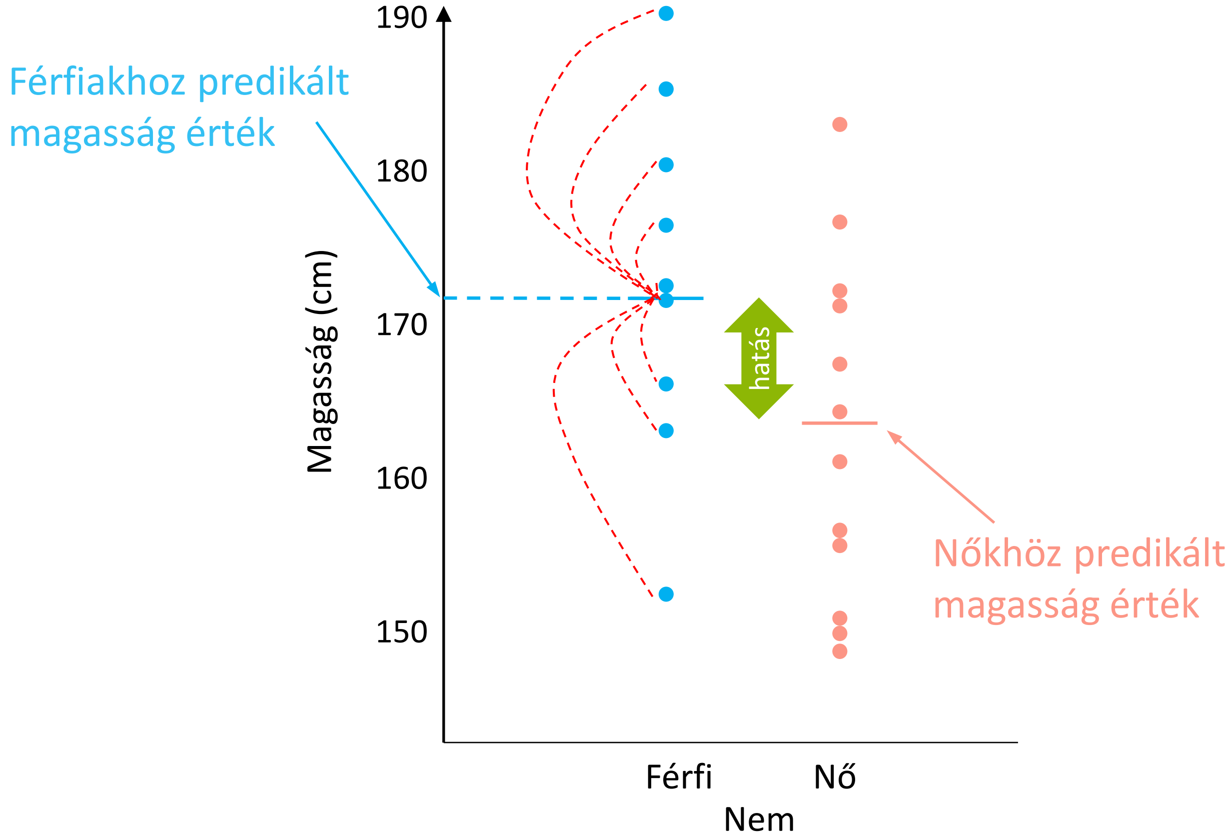
- A statisztika
  - A világ összefüggéseit modellezi, hogy a modellek alapján becsléseket tehessen
  - Ellenőrzi, a felállított modellek milyen valószínűséggel helytállóak a populcióban



Modellek

- A statisztika
  - A világ összefüggéseit modellezi, hogy a modellek alapján becsléseket tehessen
  - Ellenőrzi, a felállított modellek milyen valószínűséggel helytállóak a populcióban





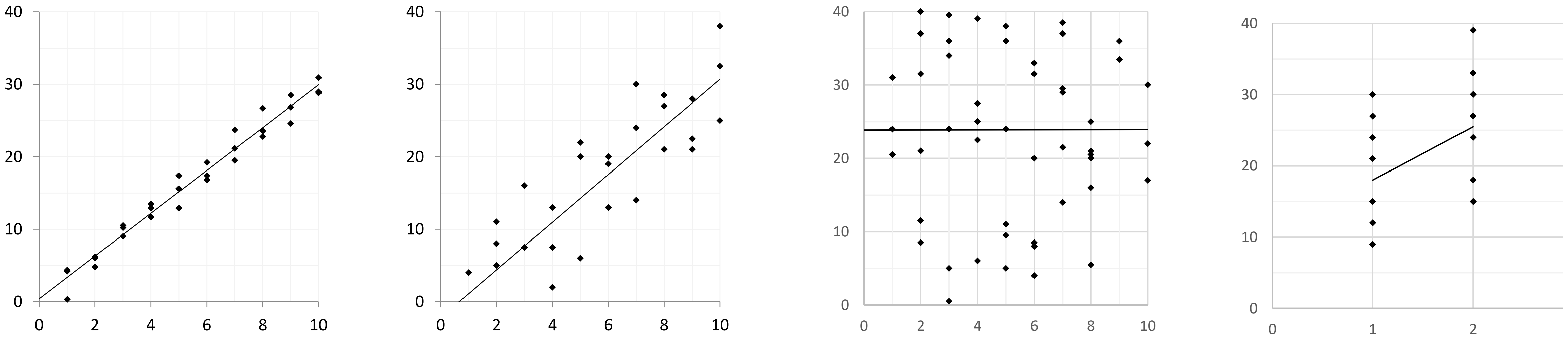
Modellek

Modellek a mért adatokat valamilyen modellel írjuk le

- **Lineáris modell**: a két változó összefüggése egy egyenessel jellemezve
- A statisztika azt ellenőrzi, hogy a felállított modellel milyen pontos becslést tudunk adni

Jól megválasztott modell esetén *gyakorlatban* egyenrangúként használt kifejezések:

- Mekkora a hatás nagysága
- Milyen erős a változók közötti kapcsolat
- Milyen mértékben magyarázza az egyik változó a másikat
- Milyen pontos becslést tudunk az egyik változóból a másikra adni
- Milyen mértékben írja le a modell a mért értékeket



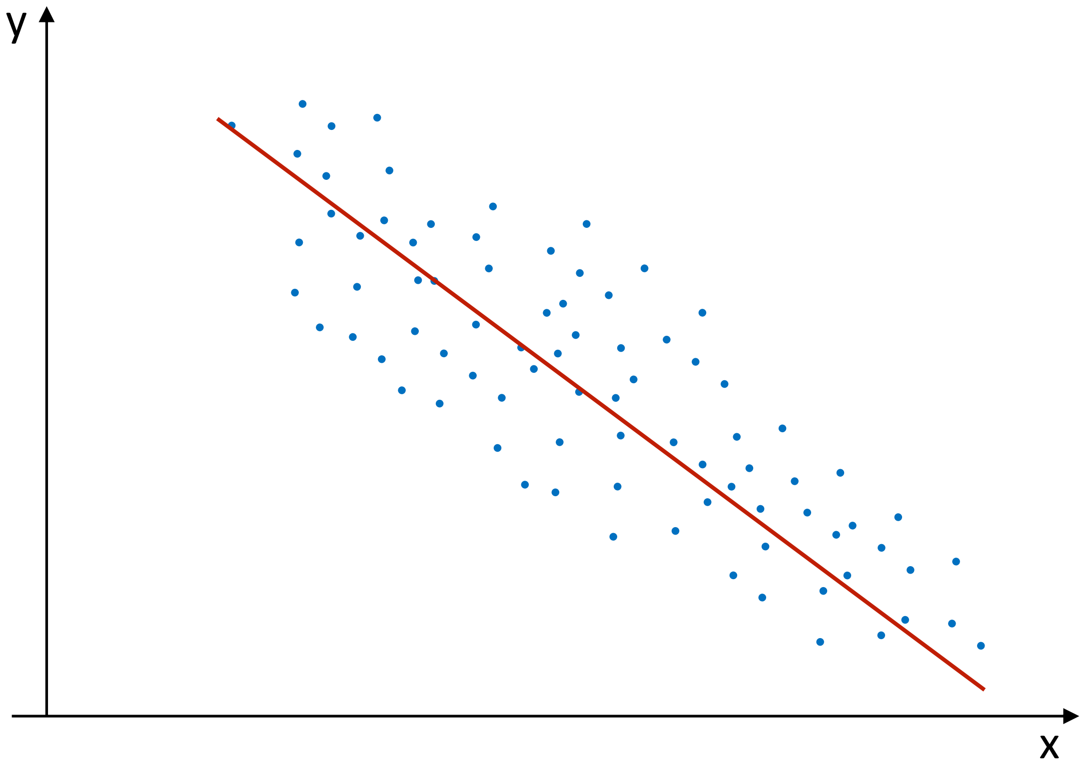
Kapcsolat & különbség

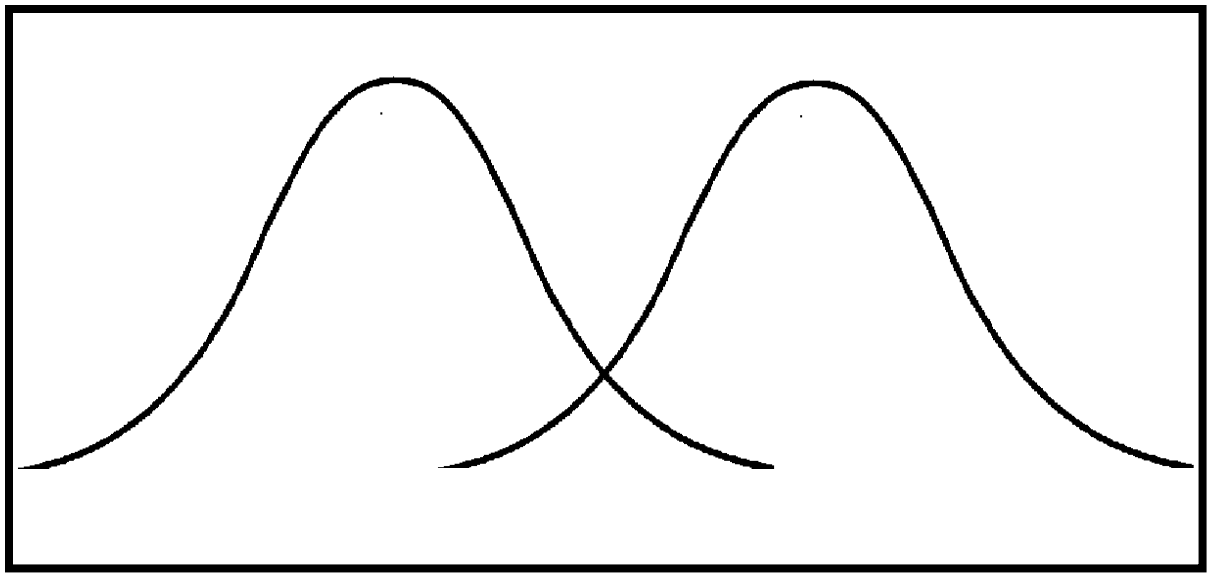
Az egyszerűség kedvéért gyakran megkülönböztetünk **kapcsolat**ot és **különbség**et vizsgáló próbákat.

Ez a megkülönböztetés a statisztikai modellezés szempontjából **nem létezik**.

Mindkettő a hatásokat vizsgálja,
és minden próba azt vizsgálja, mennyire lehet az egyik változóval (vagy változókkal) a másika(ka)t magyarázni.

Kapcsolat & különbség





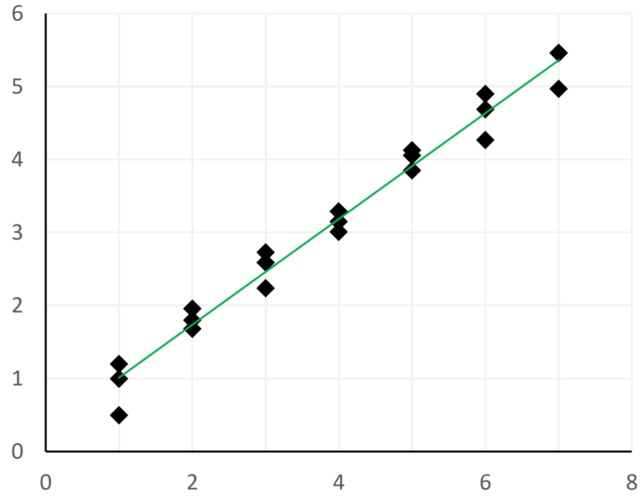
Ugyanaz a hipotézis többféleképp is megfogalmazható:

Különbség: A férfiak és a nők magassága között különbség van.

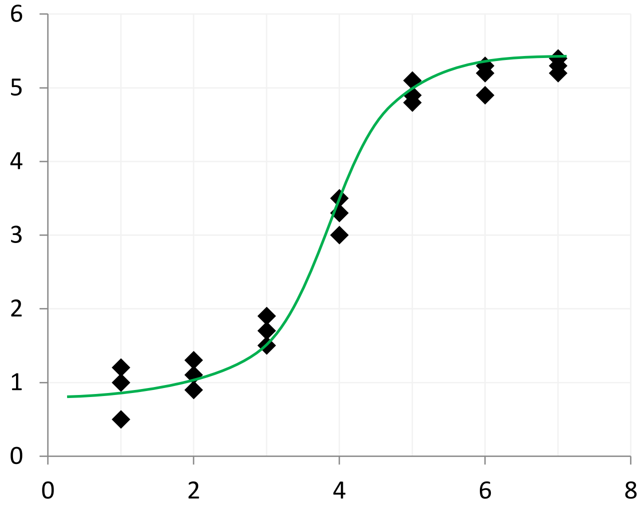
Kapcsolat: A nem és a magasság összefügg egymással.

Általános: A nem részben magyarázza a magasságbeli változatosságot.

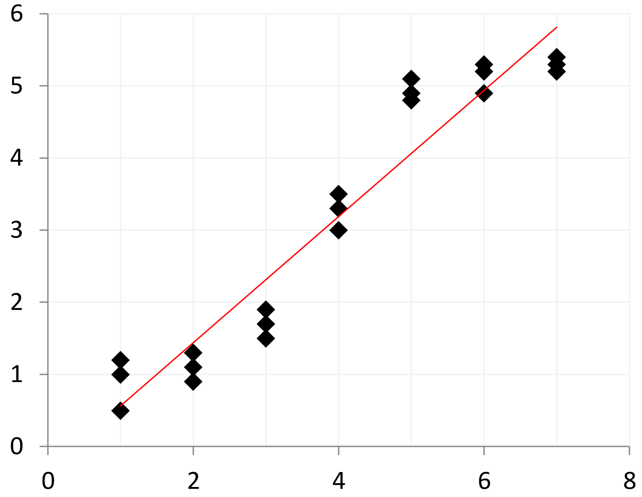
1. Fontos a jól megválasztott modell.



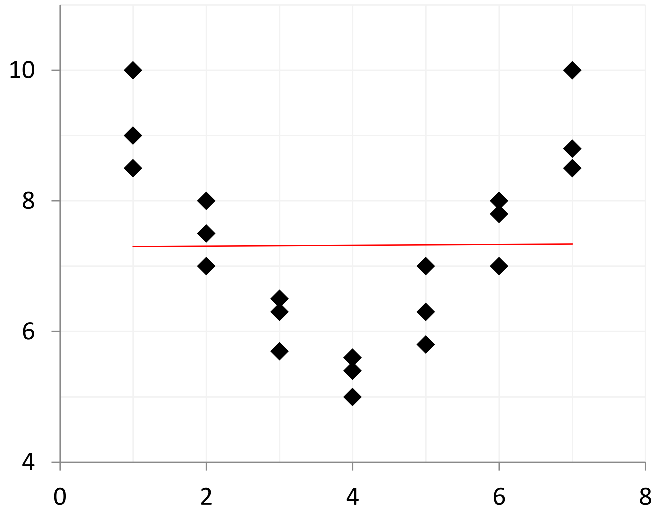
2. A legtöbb összefüggés, amivel találkozni fogunk, viszonylag jól leírható egy lineáris modellel.



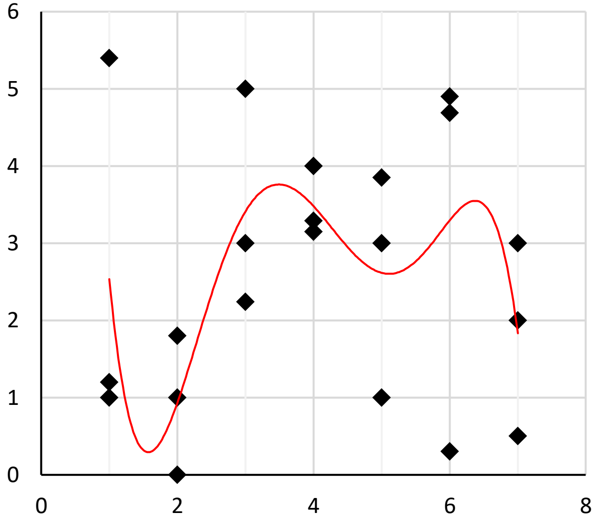
3. De attól még a világ nem csak lineáris összefüggésekből áll.



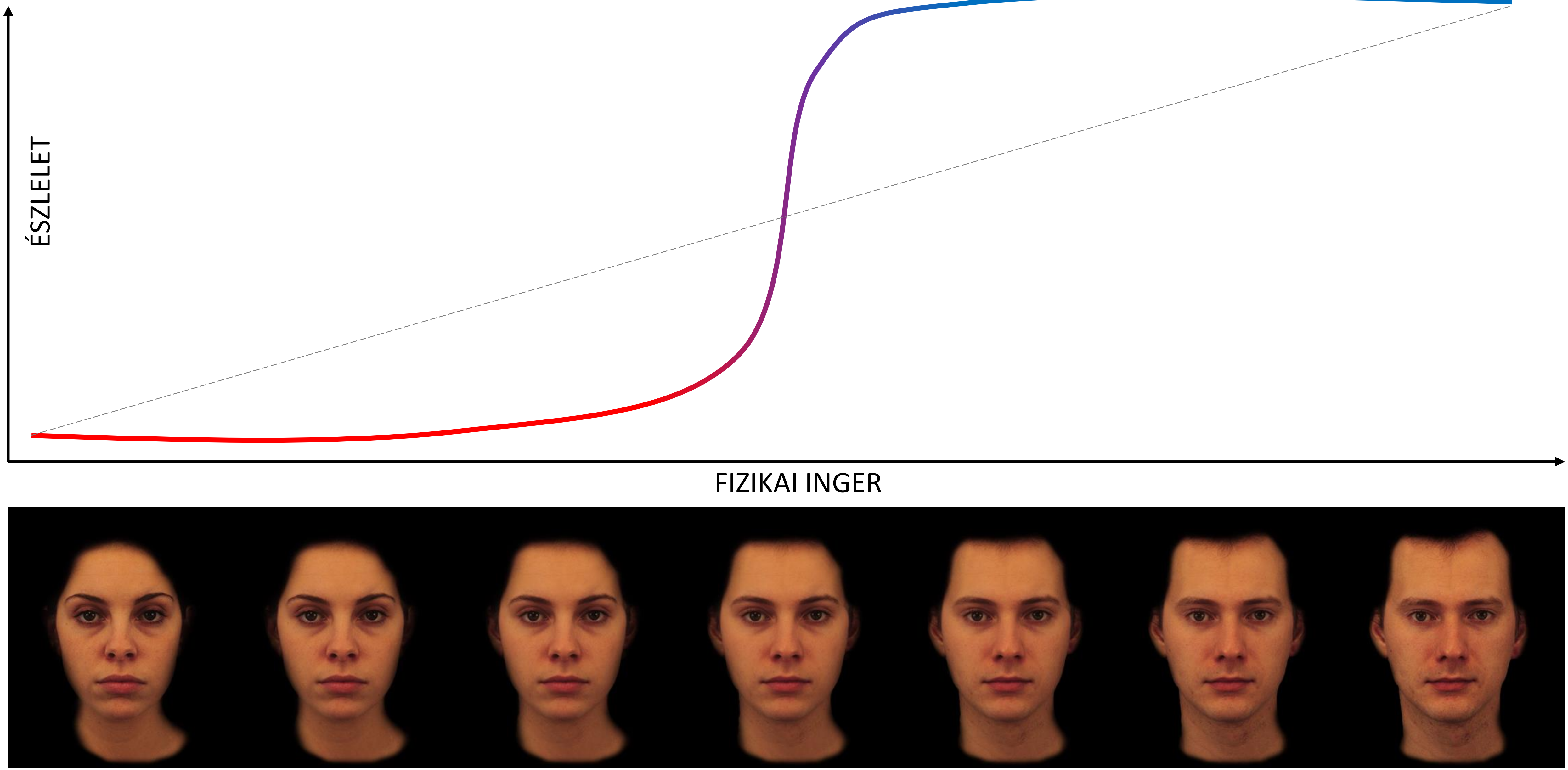
4. Ha egy nem lineáris összefüggést lineáris modellel írsz le, alul fogod becsülni a hatást.



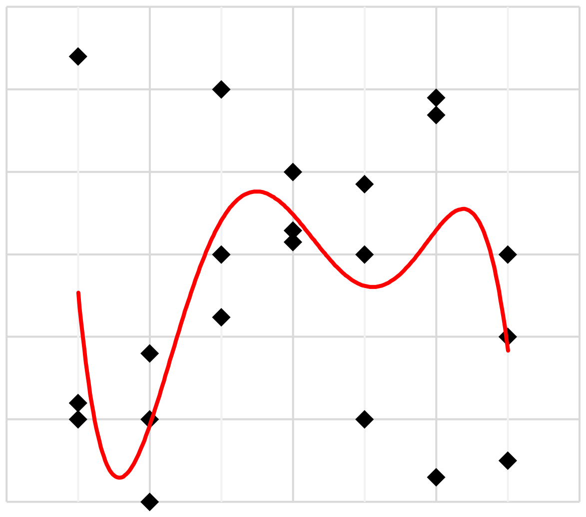
5. Sőt tévesen akár arra a következtetésre is juthatsz, hogy nincs összefüggés a két változó között.
Ezért egy lineáris modellt használó próba kikérése előtt ellenőrizni kell, megfelelő-e a lineáris modell!



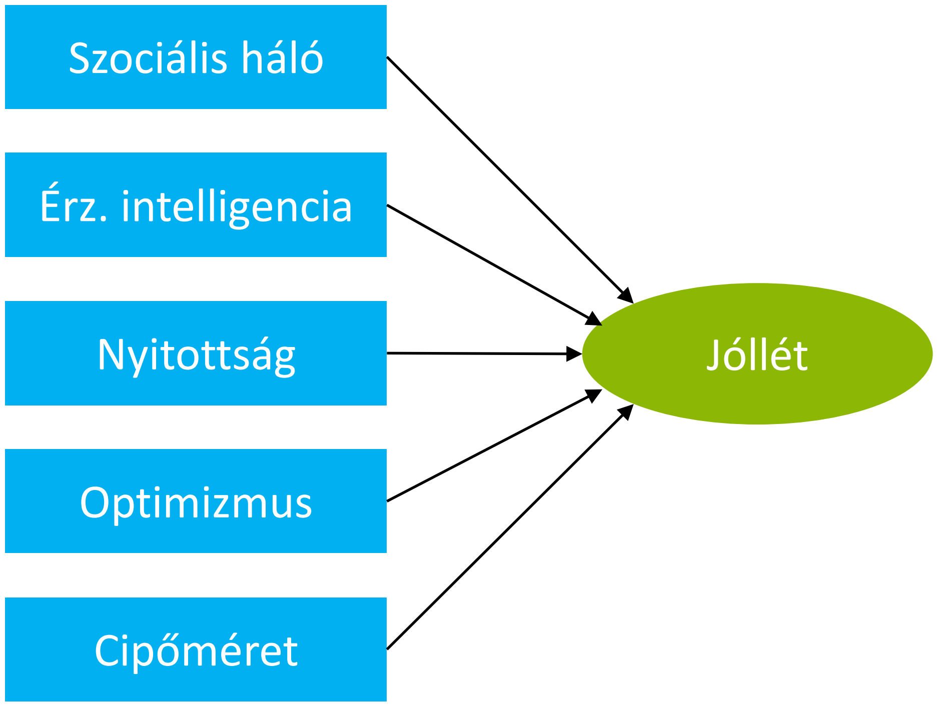
6. Nem szabad azonban túlzásba esni. Egy túlillesztett görbe (overfitting) csak a mintádra lesz igaz, nem lesz általánosítható a populációra.



Eredeti arcok: E.Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.; Morfolta: Soltész-Várhelyi Klára

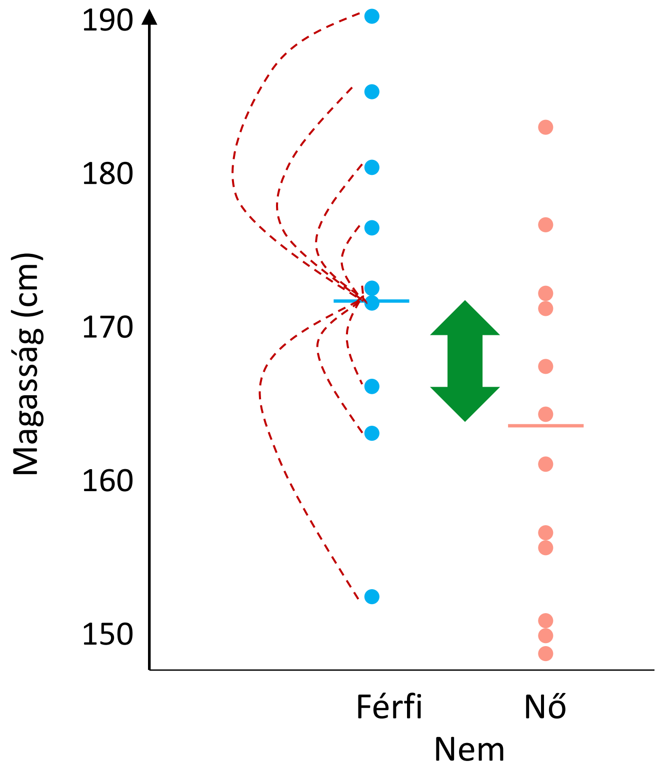


Nem szabad azonban túlzásba esni. Egy túlillesztett görbe (overfitting) csak a mintádra lesz igaz, nem lesz általánosítható a populációra.



$$Y_{pred} = b0 + b1*X1 + b2*X2 + b3*X3 + b4*X4 + b5*X5$$

Statisztikai érték



A függő változó teljes varianciája – a függő változóban lévő változatosság.

A teljes variancia egy elemzés szempontjából két részből áll:

Szisztematikus változatosság
általunk előidézett vagy általunk kontrollált tényező miatt van,
a modellünkben számolunk vele (nem, kísérleti csoportok, stb. hatása)

Nem-szisztematikus változatosság
olyan változatosság, mely okáról nem tudunk számot adni,
nem emeljük be a modellbe

$$statisztikai\ érték = \frac{átlagos\ változatosság, amit\ a\ modell\ magyaráz}{átlagos\ változatosság, amit\ a\ modell\ nem\ magyaráz} = \frac{hatás}{hiba}$$

(Ez egyenlőre egy kicsit pongyola meghatározás, de a statisztikai érték *jelentésének* megértéséhez elég. A fogalmak ennél pontosabb definíciójától egyelőre eltekintünk, majd a lineáris regresszió alatt visszatérünk rá.)

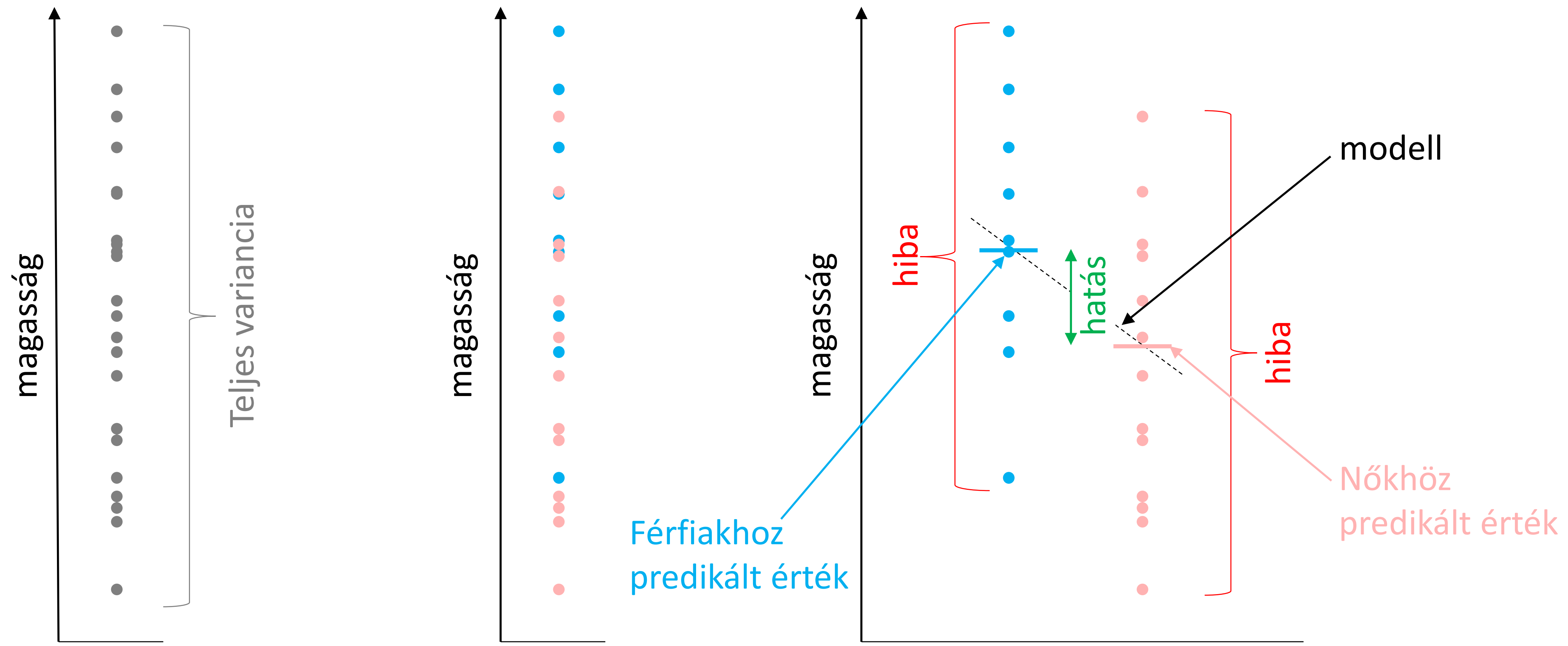
A függő (kimeneti) változó teljes varianciája: A magasságnak van egy varianciája, változatossága.

A prediktor változó szerepe: A függő változó varianciát szeretnénk egy magyarázó (prediktor) változóval, például a nemmel magyarázni.

Modell: Látjuk, hogy a férfiak átlagosan magasabbak, mint a nők. Erre felállítunk egy modellt.

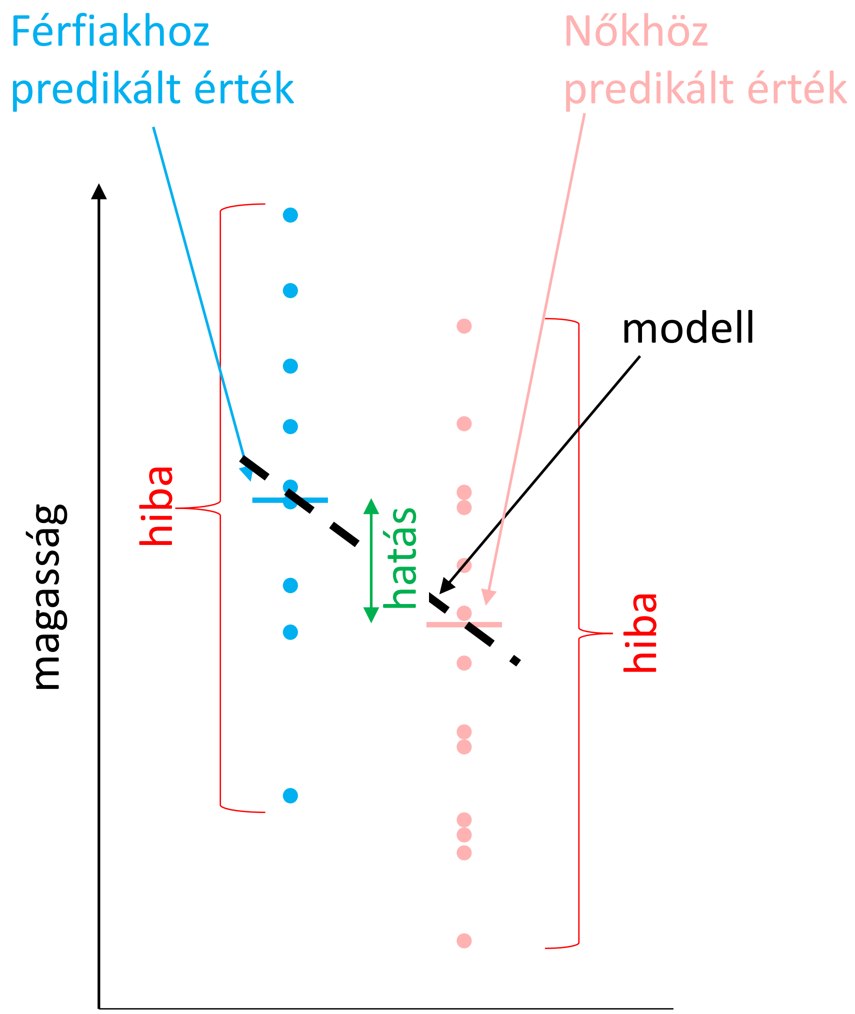
Hatás: A férfi és női minta átlaga közötti különbség, az a varianciarész, melyet a magasságból a nem megmagyaráz, melyről a modellel számot tudunk adni.

Hiba: A mintákon belüli változatosság pedig a hiba, az a varianciarész, melyet a prediktor (a nem) nem magyaráz.



Példa

- Egy egyszerű példaként nézzük meg, hogyan számolódik a t-próba t-érték! Számoljuk ki különbözik-e férfiak és nők magassága!
  - (a t-próba képlete többféle lehet a minta tulajdonságainak függvényében)
- Van egy férfi és női mintánk:
  - Elemszámok: $n_f = 9$ és $n_n = 13$ így a teljes minta: $N = 22$
  - Átlagok: $M_f = 175$ és $M_n = 164$
  - Szórások: $SD_f = 7{,}5$ $SD_n = 7{,}6$
  - Szabadságfok: $n_f - 1 + n_n - 1 = N - 2 = 20$
- Hatás: amit a magasságból a nem megmagyaráz (D, mint difference)
  - $D_{Model} = M_f - M_n = 175 - 164 = 11$
- Hiba: amit a magasság nem magyaráz
  - $SE = \sqrt{\frac{{SD_f}^2}{n_f} + \frac{{SD_n}^2}{n_n}} = \sqrt{\frac{7{,}5^2}{9} + \frac{7{,}6^2}{13}} = 3{,}270$
- t-érték: $t = \frac{hatás}{hiba} = \frac{11}{3.270} = 3{,}3639$
- A kapott t-értékhez és szabadságfokhoz a t-táblázatban kereshetjük ki a szignifikancia értéket: p = 0,003



Példa

- t-érték: $t = \frac{hatás}{hiba} = \frac{M_1 - M_2}{\sqrt{\frac{SD_1^{\,2}}{n_1} + \frac{SD_2^{\,2}}{n_2}}}$ t és df → p

- Vesd össze az első előadás a populáció és minta kapcsolatából következtetéseivel! **Mitől függ, mennyire tudsz kimutatni egy hatást?**
  - A **hatás nagyság**ától:

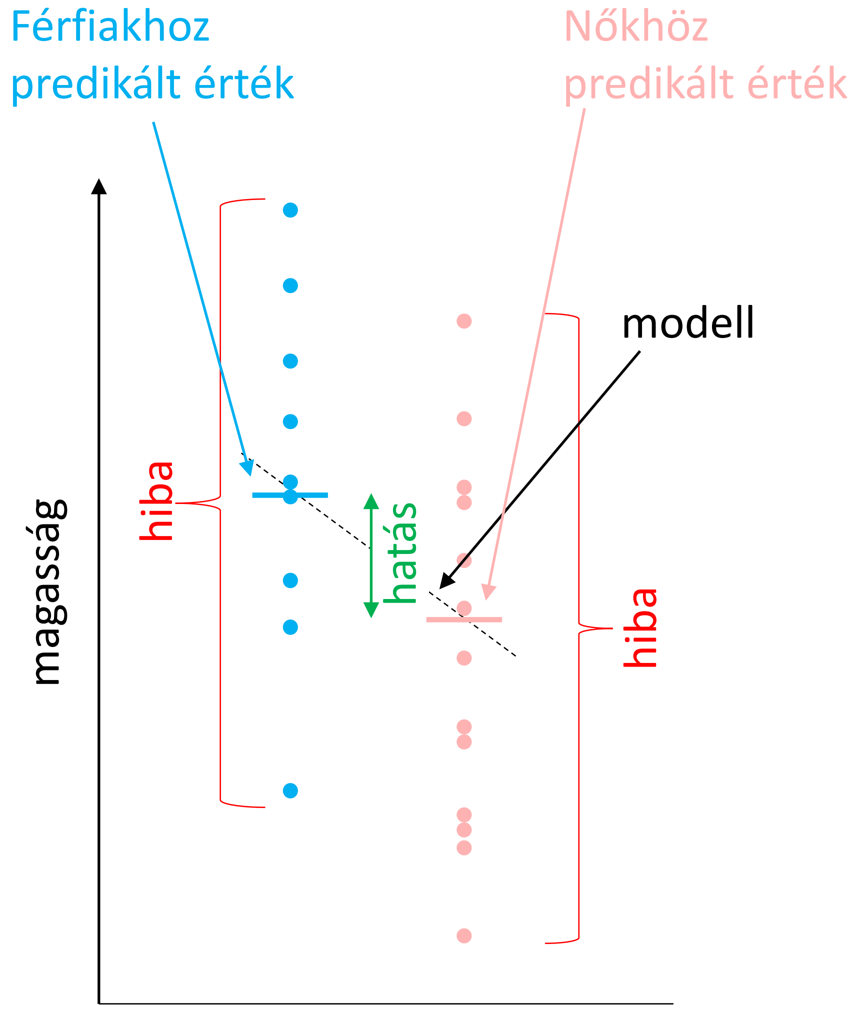
    Evidens, minél nagyobb a hatás, az a variancia, amit meg tudsz magyarázni (különbség férfiak és nők magassága között), annál könnyebb lesz azt kimutatni.
  - A **zaj, hiba** nagyságától:

    Minél nagyobb a zaj, az a variancia, amit nem tudsz magyarázni (a csoporton belüli szórás), annál inkább elveszik a zajban a hatás, annál nehezebb lesz a hatást kimutatni.

    Éppen ezért az olyan kísérleti elrendezések, melyek képesek lecsökkenteni a mérés zaját (ilyenek pl. az összefüggő mintás próbák – később megtanuljuk miért) érzékenyebbek lesznek
  - Az **elemszám**tól

    Az elemszám a nevezőben szerepel osztóként, tehát minél nagyobb az elemszám, annál kisebb a nevező, ergo nagyobb a t értéke

    Az elemszám a szabadságfokkal is összefüggésben van, és a nagyobb szabadságfokhoz csúcsosabb t-eloszlás tartozik, a kritériumszint közelében kisebbek a p-értékek ergo könnyebb szignifikáns eredményre jutni.



Hipotézis tesztelés

„The six most confusing words in statistics: failed to reject the null hypothesis”

Rebecca G. Bettencourt

Kutatói kérdés és hipotézis

Kutatói kérdés

Szabadabb nyelvezettel – számodra érthető módon megfogalmazva, milyen kérdésre keresel választ.
Például a kérdésekre:

- Eltér férfiak és nők magassága? Igaz az, hogy a férfiak magasabbak, mint a nők?
- Van összefüggés magasság és súly között?

Hipotézis

Egyértelműen igennel vagy nemmel megválaszolható állítás (lásd hipotézis-írás szabályai).
Példa hipotézisre:

- Férfiak és nők magassága között különbség van. A férfiak magasabbak, mint a nők.
- Magasság és súly között pozitív összefüggés van.

Null és alternatív hipotézis

Tesztelni viszont a hatás létezését nem tudjuk. Helyette amit tesztelni tudunk, az a **hatás hiánya**.

Null hipotézis

Nincs hatása a kísérleti manipulációnak / **nincs kapcsolat** a változók között / **nincs különbség** a minták között.

Példa nullhipotézisekre:

- Ha az a hipotézised, hogy férfiak és nők különböző magasságúak,
  akkor a null hipotézis az, hogy férfiak és nők magassága között nincs különbség,
  (mert) **férfiak és nők a magasság szempontjából egyazon populációból származnak**
- Ha az a hipotézised, hogy a magasság és súly összefüggésben van,
  akkor a null hipotézis az, hogy magasság és súly között nincs kapcsolat,
  (mert) **a magasság és súly a populációban egymástól független (ortogonális) tulajdonságok.**

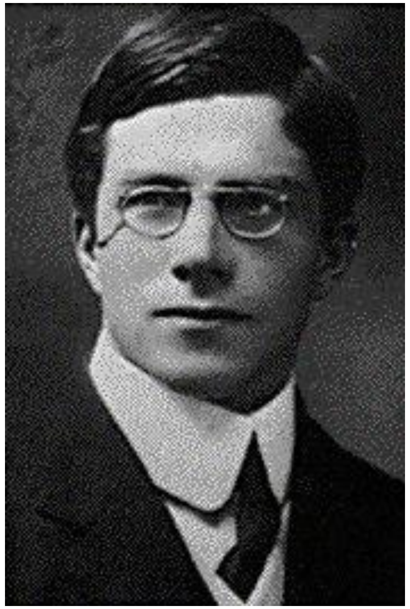
Ez az, amit **tesztelni tudunk**, amihez a **eloszlásgörbe tartozik**.

Alternatív hipotézis

Az adott kísérleti elrendezésnek **hatása van** a vizsgált változóra / a vizsgált változók között **kapcsolat van** / a minták között **különbség van**.

Ha a null hipotézis valószínűsége túl kicsi, akkor elvetjük, és helyette elfogadjuk az alternatív hipotézist igaznak.

Muriel Bristol



Az ifjú Ronald Fisher

Fisher és a teakóstoló nő
193X-ben egy teázgatás során a Rothamsted Kutatóközpont fiatal algakutatónője, Muriel azt állította, hogy meg tudja állapítani, hogy az angol teájába a tej vagy tea került először kitöltésre.

Kutatói kérdés
Tényleg meg tudja állapítani, hogy a tej vagy tea került először a csészébe?

Hipotézis
Muriel meg képes állapítani, hogy…

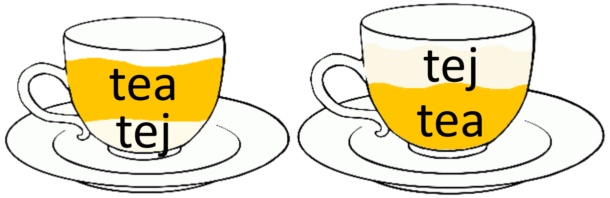
Null hipotézis
Muriel nem képes megállapítani, hogy…

Kísérlet
adjunk neki teákat, amikbe *véletlenszerűen* a tejet vagy teát öntjük először, és döntse el, szerinte melyik került először a csészébe.
(a történet szerint 6-8 csésze teát kapott, mi a szemléletesség kedvéért adjuk egy kicsit többet, 20 csészét!)

Muriel Bristol

Ha igaz a null hipotézis, és csak találgat, akkor is előfordulhat, hogy eltalál néhány csészét (ez véletlen szerepe).

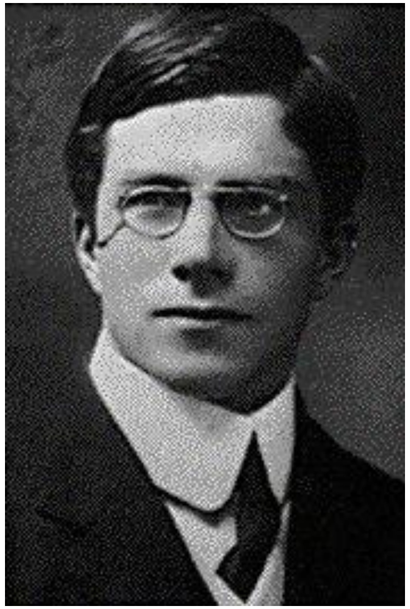
Megvan a maga valószínűsége annak, hogy egyet sem talál el, annak, hogy egyet, kettőt, stb… Annak valószínűsége, hogy N csésze teából X darabot talál el:



$$p_{X,N} = \frac{\binom{N}{X}}{2^N} = \frac{\left(\frac{N!}{X!*(N-X)!}\right)}{2^N}$$

Például annak valószínűsége, hogy a 20 csészéből 5-öt talál el,

$$p_{5,20} = \frac{\binom{20}{5}}{2^{20}} = \frac{\left(\frac{20!}{5!*(20-5)!}\right)}{2^{20}} = 1{,}479\%$$



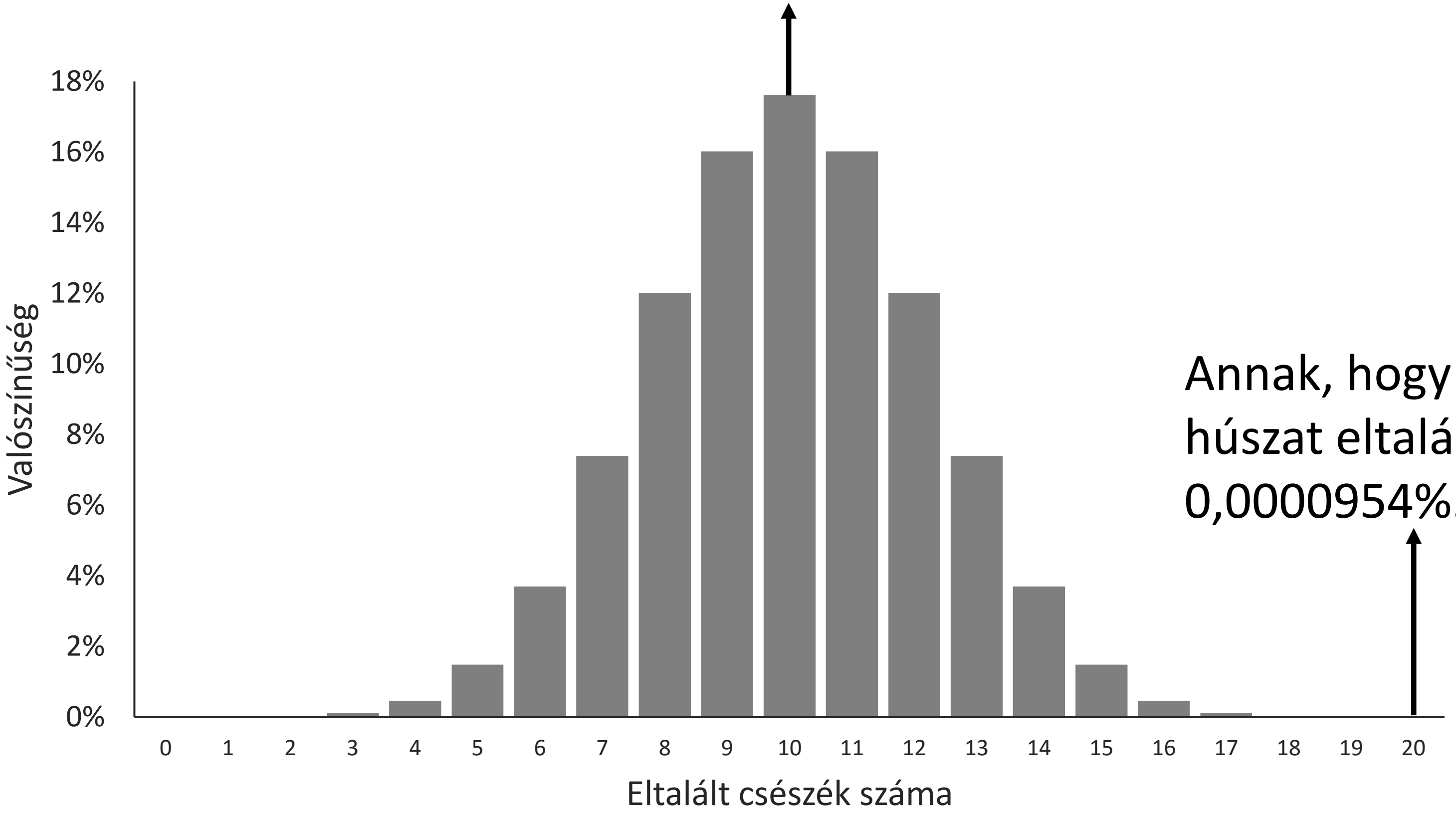
Az ifjú Ronald Fisher

Muriel Bristol

Ha igaz a null hipotézis, és csak találgat, akkor is előfordulhat, hogy eltalál néhány csészét (ez véletlen szerepe).

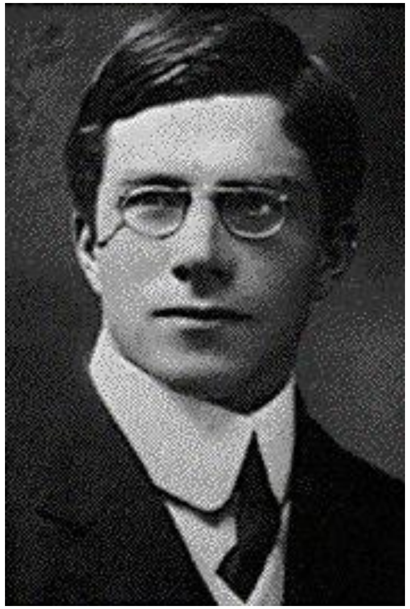
A valószínűségek felrajzolhatóak egy eloszlásgörbén.

Ha csak találgat, a legnagyobb valószínűsége (17,620%) annak van, hogy a csészék felét találja el.



Teák eltalálási valószínűsége véletlen találgatás esetén

Annak, hogy mind a húszat eltalálja, összesen 0,0000954%.

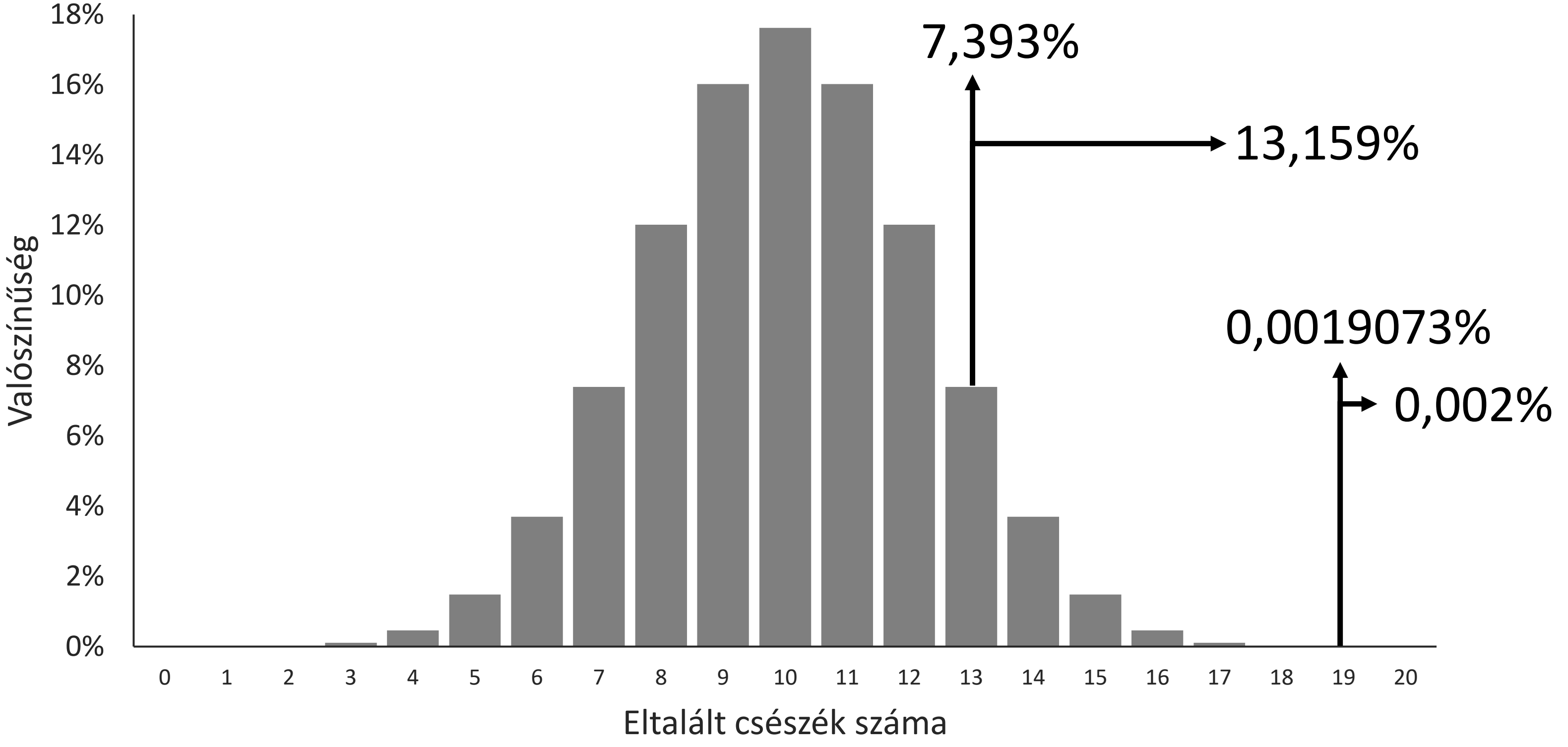


Az ifjú Ronald Fisher

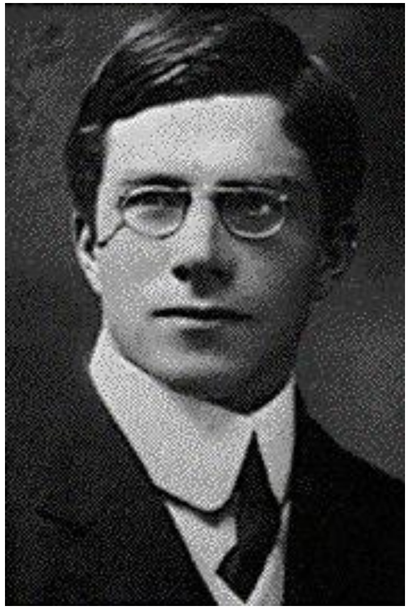
Muriel Bristol

Most nézzük meg a kísérlet eredményét!

- Ha Muriel csak 13 csészét talált volna el a húszból, akkor nem hinnénk el neki, hogy van teakóstoló képessége, hiszen véletlen találgatással is 7,393% esélye volt, hogy eltalál 13 csészét, és annak, hogy *legalább* 13-at, 13,159%.
- De Muriel eltalált 19-et a húszból! Véletlen találgatással ennek a valószínűsége csak 0,0019073%, és a *legalább* 19 találatnak pedig 0,002%!



Teák eltalálási valószínűsége véletlen találgatás esetén

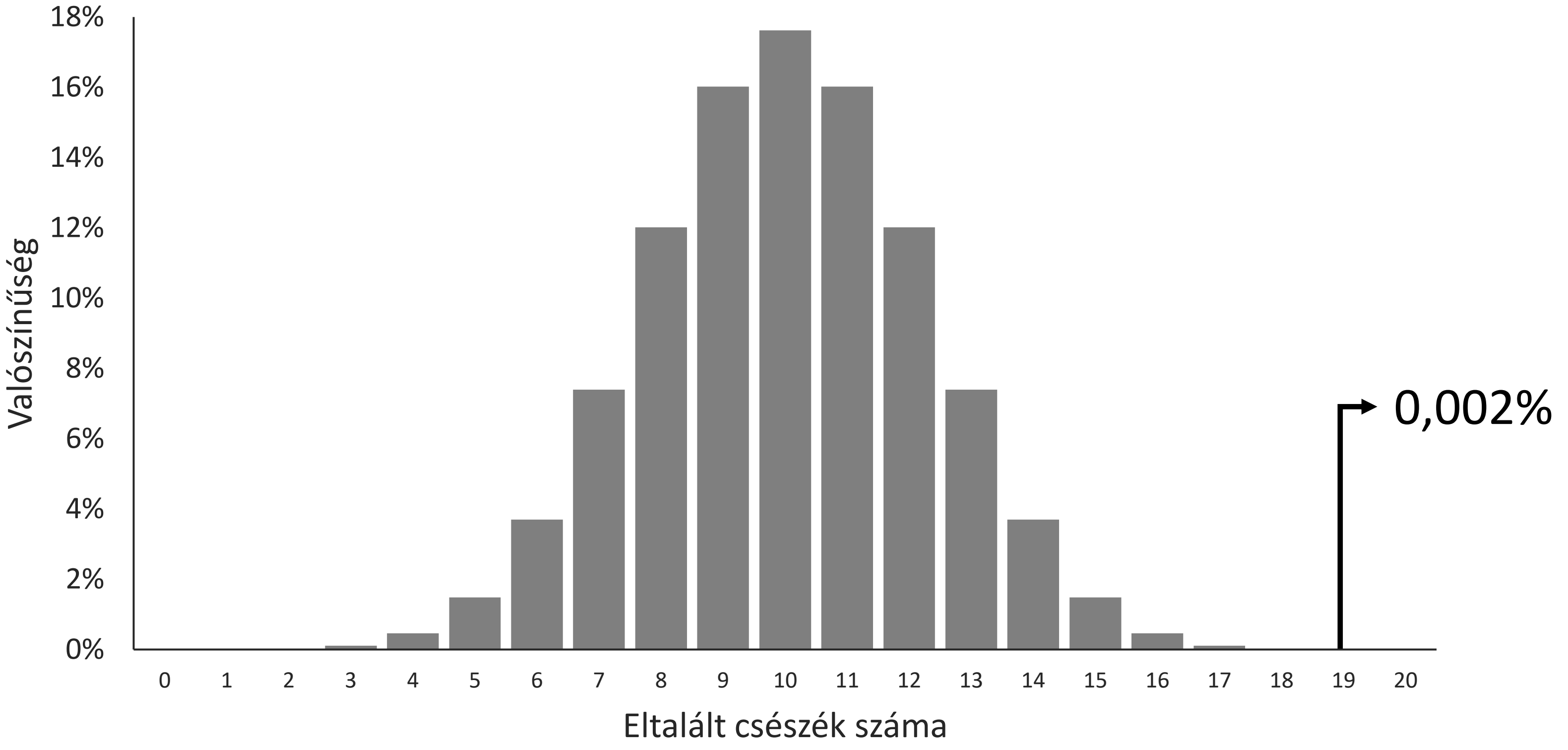


Az ifjú Ronald Fisher

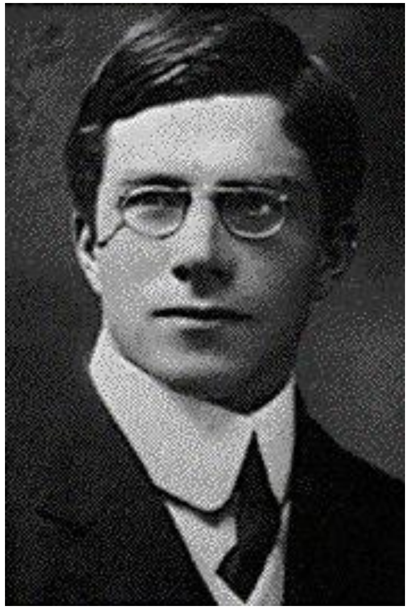
Muriel Bristol

Mivel a nullhipotézis esetén annak a valószínűsége, hogy Muriel ilyen jó teljesítményt nyújtson, elképesztően alacsony,

- ezért elvetjük a nullhipotézist (nincs tea-kóstoló képessége),
- és helyette elfogadjuk az alternatív hipotézist, azaz elfogadjuk azt, hogy Muriel tényleg meg tudja állapítani a tej és tea sorrendjét az angol teában.



Teák eltalálási valószínűsége véletlen találgatás esetén



Az ifjú Ronald Fisher

Statisztikai érték – szignifikancia érték – szignifikancia szint

Nézzük ugyanezt a gondolatmenetet egy független mintás t-próbára!

Kutatói kérdés
Van különbség férfiak és nők magassága között? Igaz az, hogy a férfiak magasabbak, mint a nők?

Hipotézis
Férfiak és nők magassága között különbség van, a férfiak magasabbak, mint a nők.

Null hipotézis
Férfiak és nők magassága között nincs különbség, a magasság szempontjából férfiak és nők ugyanabból a populációból származnak.

Kísérlet
Lásd korábban: random kiválasztott 9 férfi és 13 nő magasságát mértük le.

A független t-próbához használt **statisztikai érték** a t-érték.

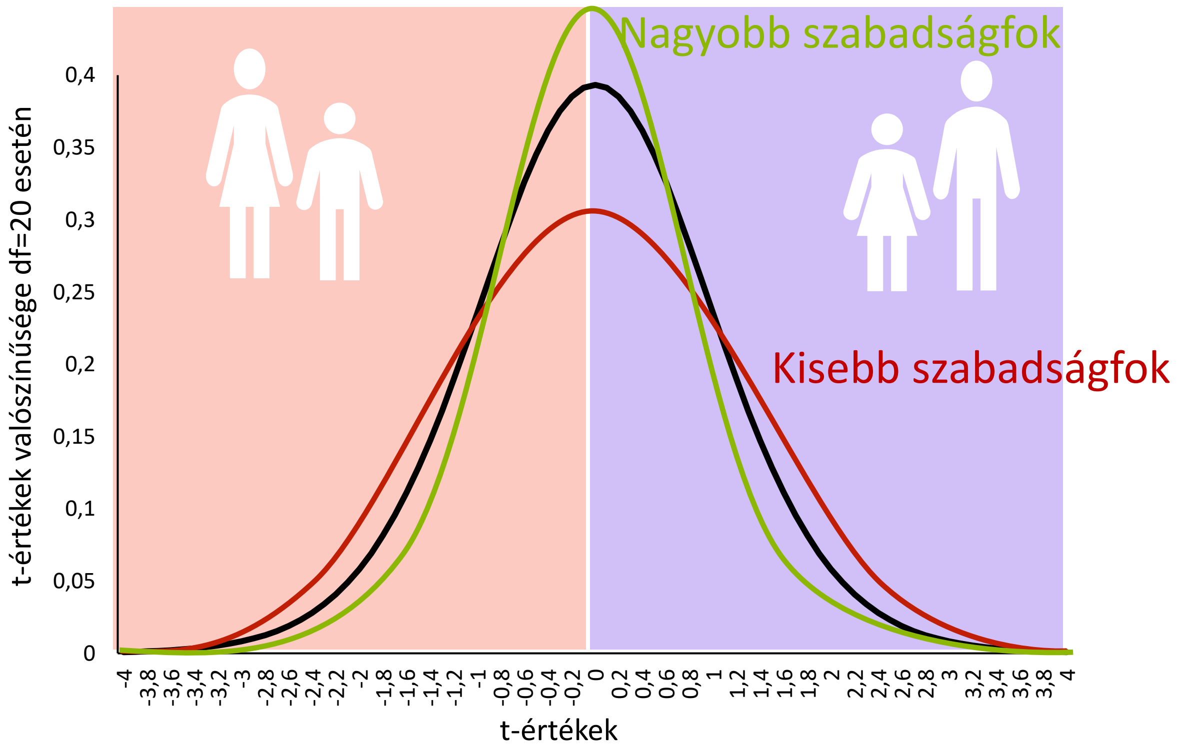
$$t = \frac{hatás}{hiba} = \frac{M_1 - M_2}{\sqrt{\frac{{SD_1}^2}{n_1} + \frac{{SD_2}^2}{n_2}}}$$

Az **eloszlásgörbén** azt ábrázoljuk, hogy milyen valószínűségek várhatók a statisztikai értékekhez a null hipotézis esetén.

Azaz mennyi egy adott statisztikai érték (t-érték / a férfi és női *minta* magassága közötti különbség) valószínűsége akkor, ha a *populáció*ban nincsen eltérés férfiak és nők magassága között.

Null hipotézis esetén (azaz ha populációban nincs különbség férfiak és nők magassága között) melyik statisztikai értéknek a legnagyobb a valószínűsége?

A t-eloszlás görbéjének alakja a szabadságfoktól függ.



Most számoljuk ki a mintáink alapján mekkora a tényleges t-érték (számolást lásd korábban)

$$t = \frac{hatás}{hiba} = \frac{175 - 164}{\sqrt{\frac{7{,}5^2}{9} + \frac{7{,}6^2}{13}}} = 3{,}3639$$

Keressük ki a statisztikai értékünkhöz (t-értékhez) tartozó valószínűségét!

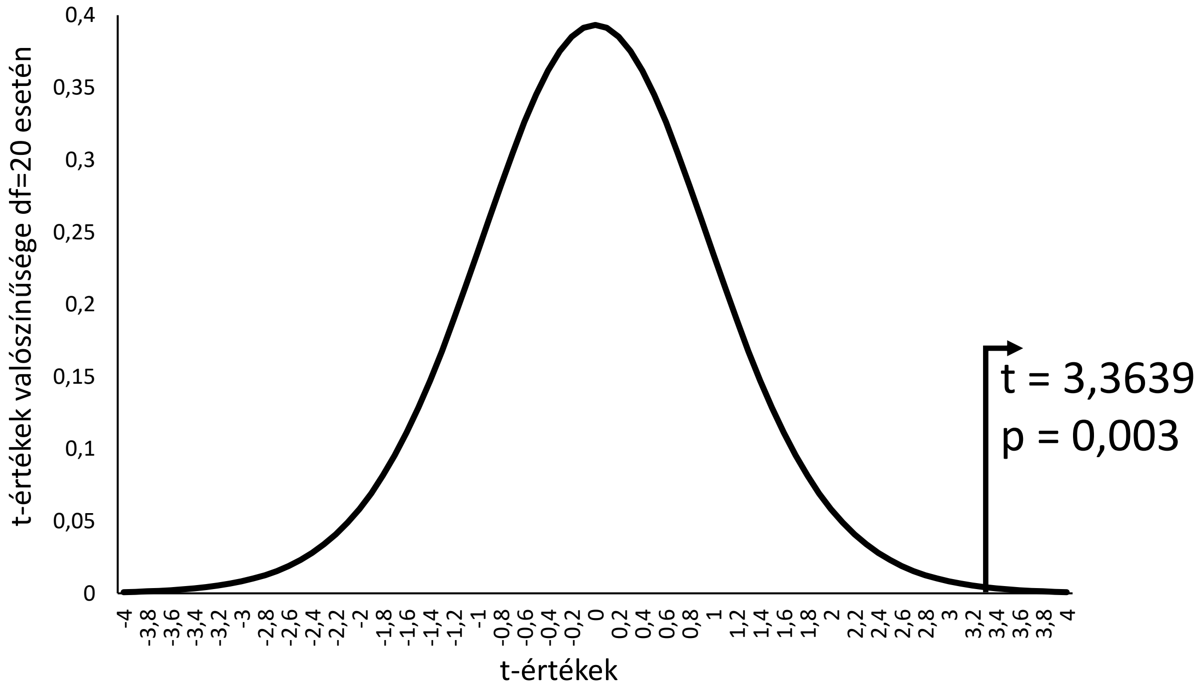
p-érték (**szignifikancia érték**): Adott t-érték vagy annál nagyobb érték előfordulási valószínűsége, ha a null hipotézis igaz.

Azaz annak *valószínűsége*,

hogy egy olyan *populáció*ból, melyben *nincsen* eltérés férfiak és nők magassága között,

a *véletlen mintavételezés* során sikerül egy olyan férfi és női mintát kiválasztani, mely között legalább t = 3,3639 mértékű különbség van,

p = 0,003, azaz 0,3%.



Ha a null hipotézis teljesülését feltételezve a kapott statisztikai érték előfordulási valószínűsége nagyon alacsony, akkor elvetjük a null hipotézist, és elfogadjuk az alternatív hipotézist.

Mit tekintsünk „nagyon alacsony”-nak?

Fisher definíciója alapján legyen a kritérium 5%.

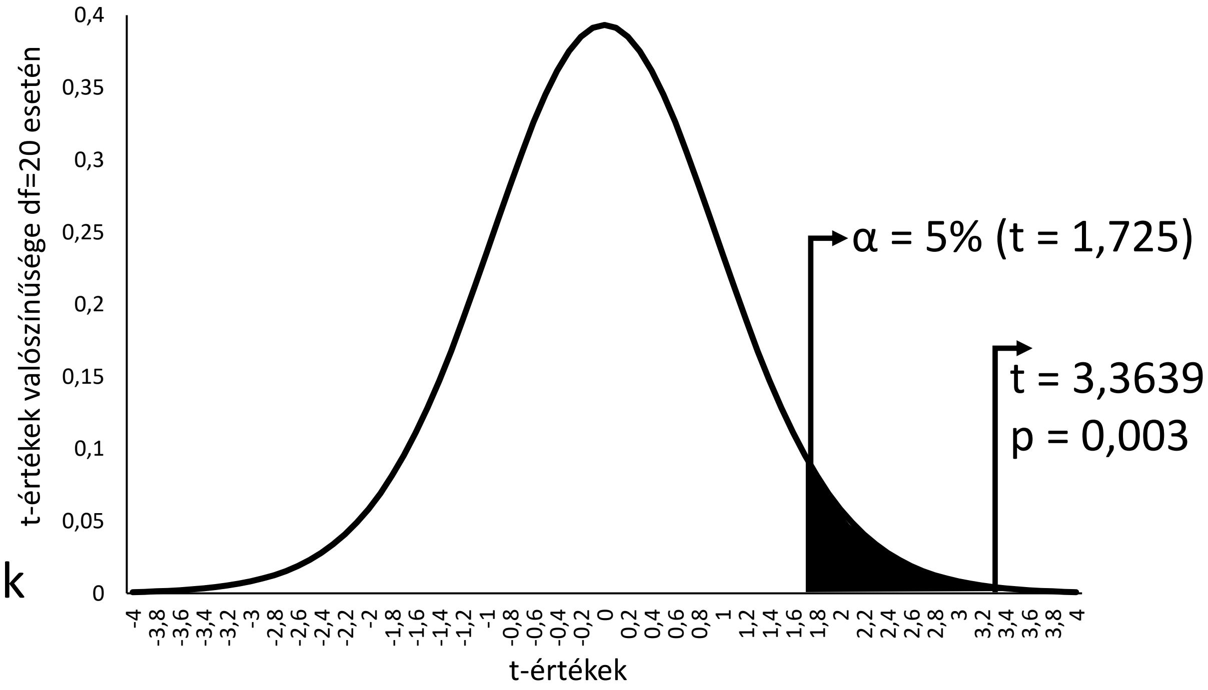
Szignifikancia szint (**α-szint**): az a valószínűség, aminél a p-nek kisebbnek kell lenni ahhoz, hogy elvessük a null hipotézist.

Döntés klasszikus hipotézis-tesztelésnél:
Csak ha maximum 5% esély van arra, hogy a minta véletlenül alakult úgy, hogy úgy tűnik van valamilyen hatás, pedig valójában nincsen, akkor hisszük el a hatás létezését.

$$\boldsymbol{p < .05}$$

Szignifikancia érték (p-érték)
annak valószínűsége, hogy az észlelt hatás a véletlen műve

Szignifikancia szint
kritérium szint, aminél a p-nek kisebbnek kell lennie ahhoz, hogy elvessük a nullhipotézist



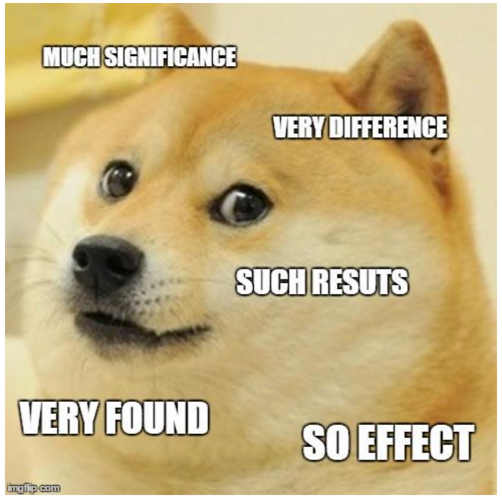
Szignifikancia – gyakori tévedések

„Elég a p < .05-öt kiírni"

„A szignifikancia szintet még a statisztikák elvégzése előtt állapítottam meg, ezért az érdekes csak, hogy p ennél kisebb vagy nagyobb, nem az, hogy mennyi" – **elavult elgondolás,** mert:

- egyik eredmény $p = .051$ -> $p > .050$
- másik $p = .049$ $p < .050$
- A két p-érték egymáshoz nagyon közel, ha viszont csak úgy jelenítjük meg, hogy kisebb vagy nagyobb 5%-nál, értelmezésük teljesen más lesz.

APA formátum szerint ki kell írni a p értékét 3 tizedesjegyig!



„Nagyon szignifikáns hatást találtam"

A szignifikancia értékéből a *hatás általánosíthatóságára* lehet következtetni, *nem a hatás nagyságára*, hiszen a szignifikancia értéke elemszámfüggő (is).

Attól, hogy valami szignifikáns még nem biztos, hogy jelentős.

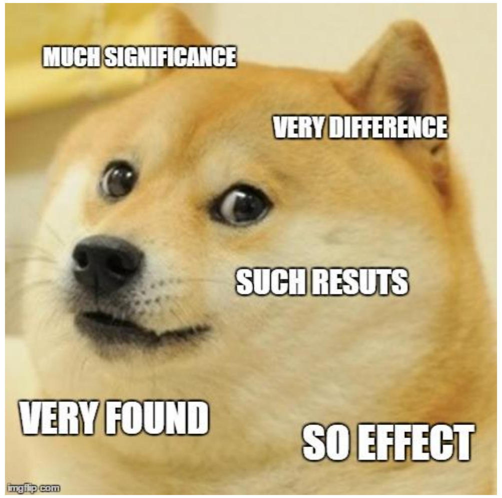
Nagyon pici hatás is lehet szignifikáns, ha például elég emberrel vettem fel.

Szignifikancia – gyakori tévedések

„Nem szignifikáns a különbség, tehát a két minta egyenlő”

Ha nem tudom elvetni a nullhipotézist, attól még **NEM** lesz igaz
Lehet, hogy nincs hatás, vagy annyira gyenge, hogy nem tudom megmondani, a véletlen műve-e.
A nem-szignifikáns eredmény jelentése:

(1) A populációban vélhetőleg nincs hatás
(2) túl nagy a zaj vagy túl kicsi a minta, hogy el merjem vetni a null hipotézist



„Szignifikáns különbséget találtam, tehát tuti különböznek”

Valószínűségekkel számolunk. Ha valami szignifikáns, az csak azt jelenti, hogy kicsi az esélye, hogy nincs hatás, és mégis ilyen adatokat sikerült gyűjtenem, de soha nem biztos a hatás létezése

Első- és másodfajú hiba

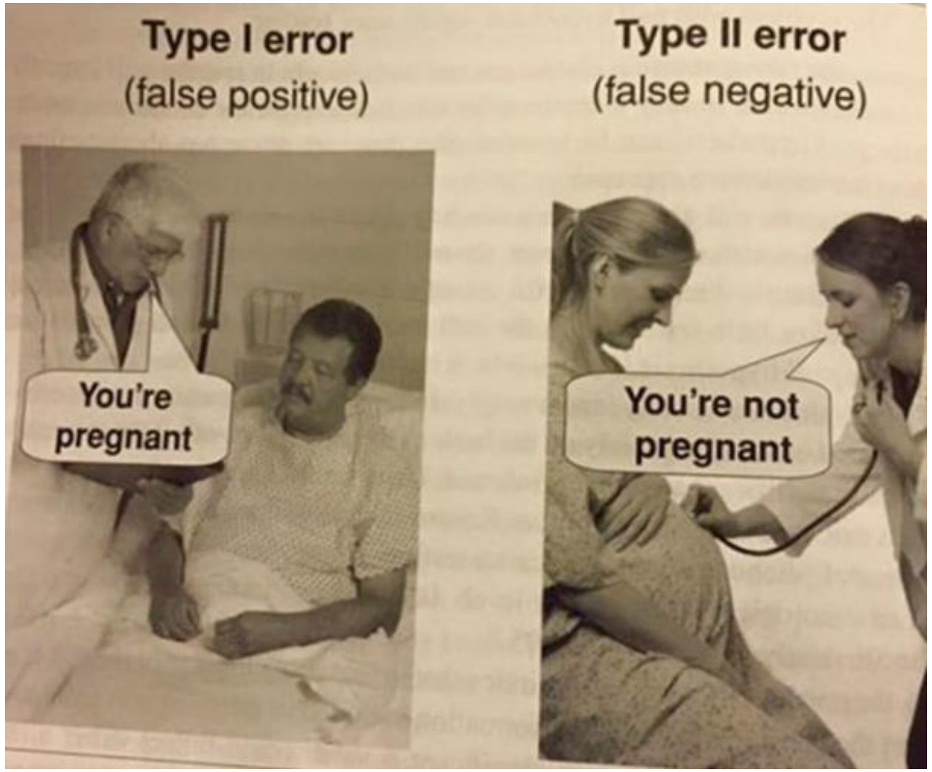
A statisztikában valószínűségekkel dolgozunk, ezért bárhogy döntünk, mindig megvan a valószínűsége, hogy tévedtünk. **Kétféleképpen tévedhetünk**:

Elsőfajú hiba (**α - szint**)

Amikor valamilyen hatásról azt hisszük, hogy létezik, pedig nem.

Amikor a null hipotézist elvetjük, pedig igaz.

A tévedés maximálisan elfogadható valószínűségét a szignifikancia szinttel határozzuk meg, melynek értéke Fisher kritériuma alapján 5% alatt van.



Első- és másodfajú hiba

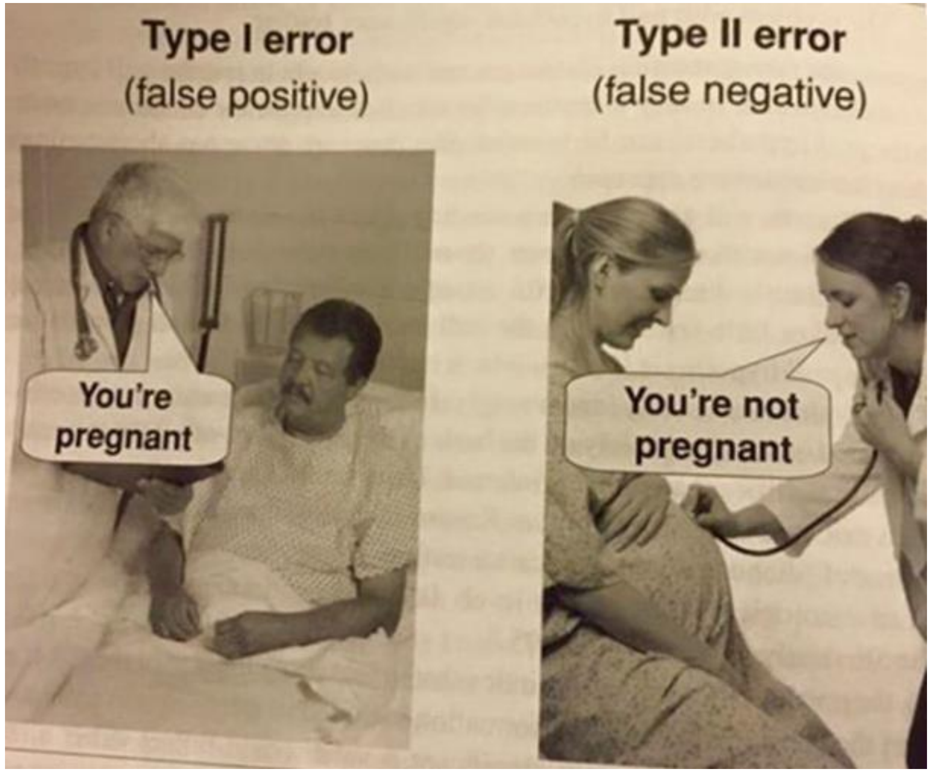
A statisztikában valószínűségekkel dolgozunk, ezért bárhogy döntünk, mindig megvan a valószínűsége, hogy tévedtünk. **Kétféleképpen tévedhetünk**:

Másodfajú hiba (**ß - szint**)

Amikor valamilyen hatásról azt hisszük, hogy nem létezik, pedig csak nem vettük észre.

Amikor a null hipotézist megtartjuk, pedig nem igaz.

Cohen alapján elvárható érték 0,2 alatt van, tehát maximum 20% esélye lehet annak, hogy nem veszünk észre egy meglévő hatást



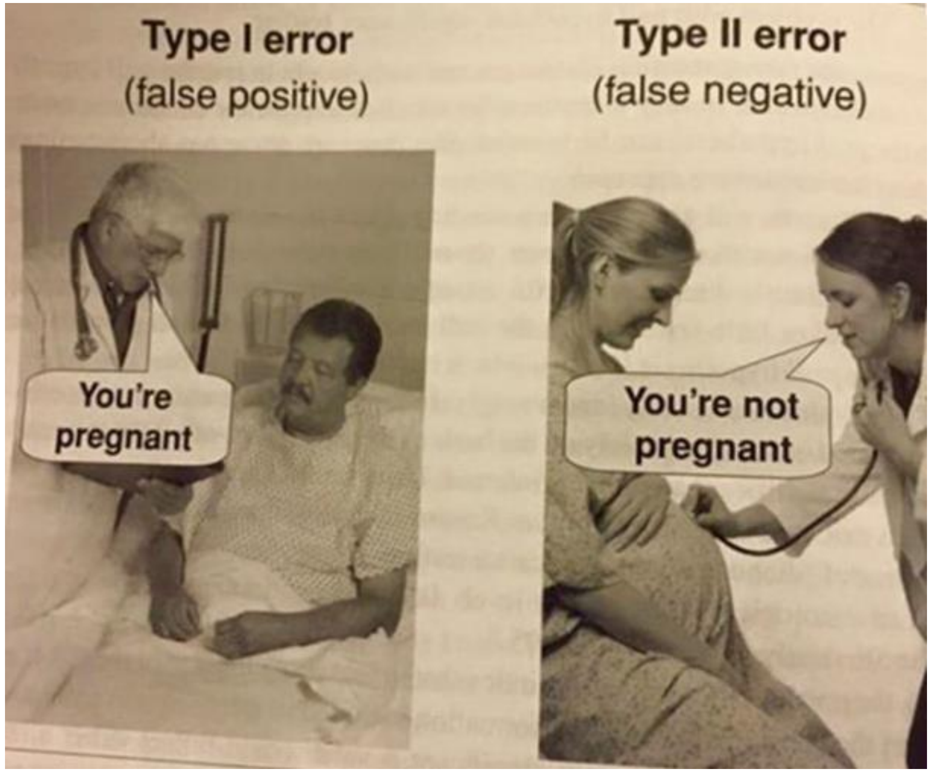
Első- és másodfajú hiba

Az első- és másodfajú hiba másra vonatkozik:

Elsőfajú hiba
Amikor a populációban nincsen jelen a hatás, mi mégis szignifikáns eredményre jutunk a mintáink alapján.

Másodfajú hiba
Amikor a populációban jelen van a hatás, mi mégsem tudtuk azt kimutatni.

A kettő között nem arányos a trade-off.



Másodfajú hiba és statisztikai erő

Statisztikai erő

Arra való képességet méri, hogy

- ha a változónknak van hatása, azt észrevesszük
- a null hipotézist helyesen elutasítsuk, ha nem igaz.

A másodfajú hiba ellentéte

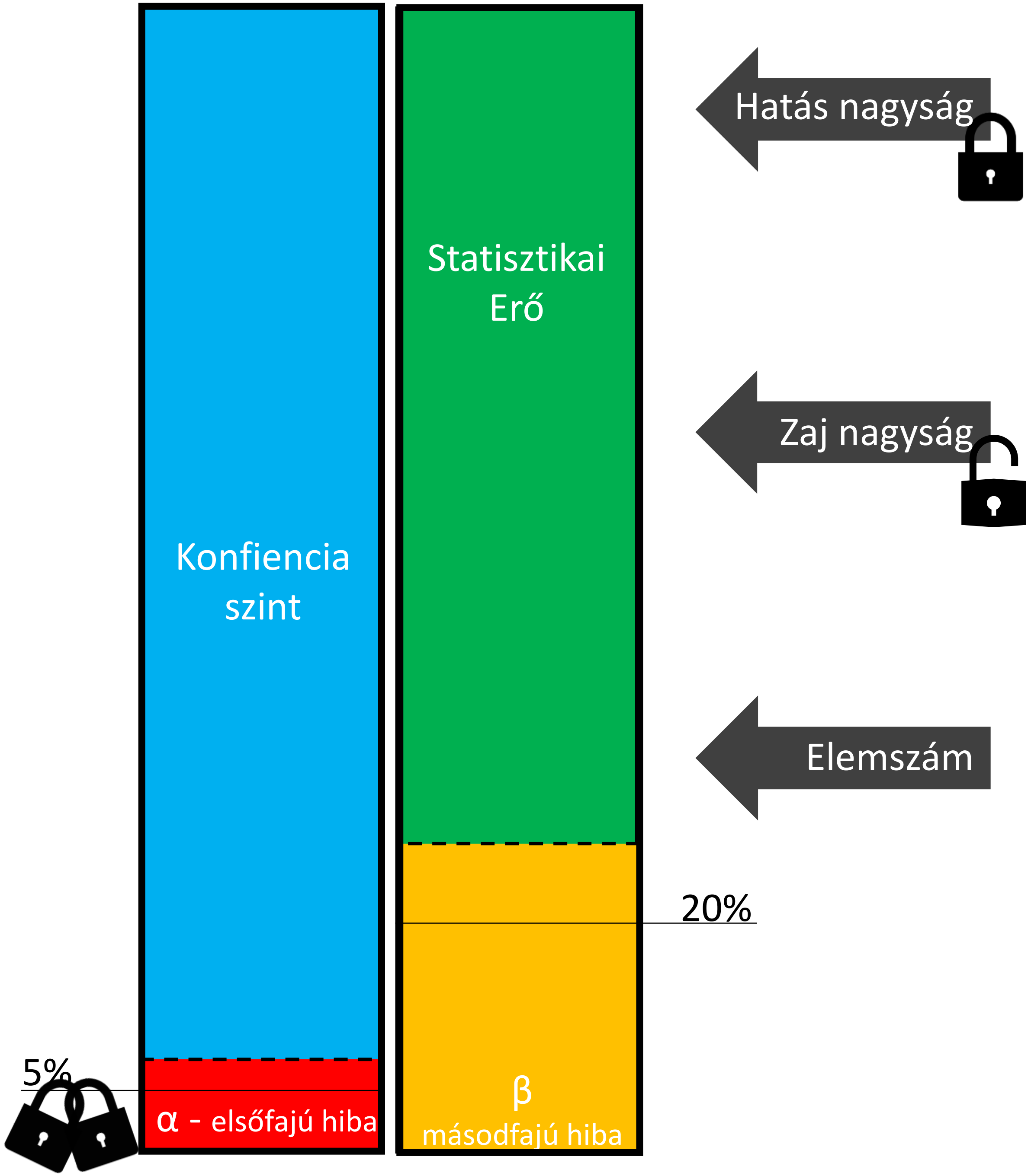
Másodfajú hiba: a populációban a hatás létezik, de mi nem tudtuk kimutatni.
Statisztikai érték: az a képesség, hogy ha a populációban a hatás létezik, mi ki tudjuk mutatni azt.

Elfogadható minimális értéke 1-ß = 0,8; **tehát, akkor elfogadható egy vizsgálat, ha a populációban jelen van a hatás, akkor azt 80% valószínűséggel észre fogjuk venni.**
Akkor jó egy vizsgálat, hogy ha a populációban tényleg ott van a hatás, és mi elvégezzük a vizsgálatot rengetegszer, akkor legalább a vizsgálatok 80%-ban szignifikáns eredményt kapunk.

A próba eredménye

Populáció		Szignifikáns	Nem szignifikáns
	A hatás létezik	A 80%	B 20%
	A hatás nem létezik	C 5%	D 95%

A/(A+B) = statisztikai erő
B/(A+B) = másodfajú hiba valószínűsége
C/(C+D) = elsőfajú hiba valószínűsége, azaz szignifikancia érték

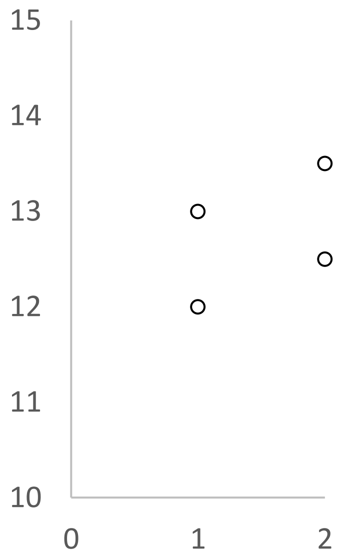


- El kellene érni, hogy **α<5%** és **β<20%** legyen
- **Az első és másodfajú hiba között negatív, nem lineáris kapcsolat van**:
  - minél „szigorúbb” vagyok, annál kisebb lesz az elsőfajú hiba, de annál kevésbé valószínű, hogy egy létező hatást ki tudok majd mutatni.
  - Minél „megengedőbb” vagyok, annál könnyebben mutatok ki hatásokat, de annál inkább fogom tévesen is azt hinni, hogy van hatás.
- A mérés érzékenységét 3 további tényező befolyásolja:
  - **Hatásnagyság**: Minél nagyobb a hatás, annál könnyebb kimutatni (különbség indonéz és norvég emberek magassága között). (Persze szeretnénk minél kisebb hatás kimutatására képesek lenni.)
  - **Zaj**: Minél nagyobb a szórás, annál nehezebb egy hatást kimutatni.
  - **Elemszám**: minél nagyobb az elemszám, annál könnyebb egy hatást kimutatni.

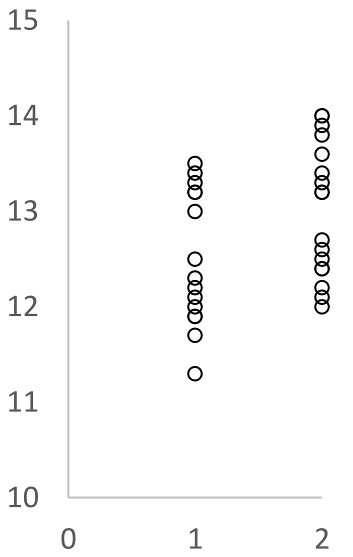
Elemszám és szignifikancia

Minél nagyobb az elemszám, annál kisebb hatás is szignifikáns lesz.

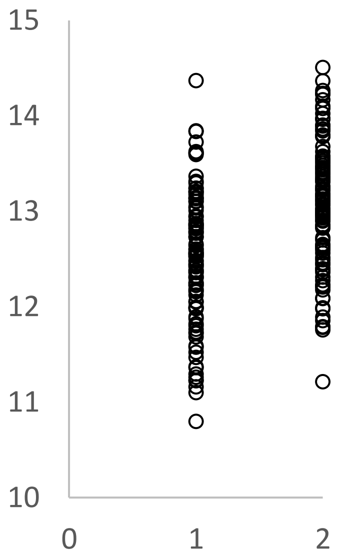
Melyik esetben hiszed el inkább, hogy különbség van a két csoport között?



Áltag: 12,5 és 13
Szórás: 0,7 és 0,7
Elemszám: 2 és 2



Áltag: 12,5 és 13
Szórás: 0,7 és 0,7
Elemszám: 15 és 15



Áltag: 12,5 és 13
Szórás: 0,7 és 0,7
Elemszám: 100 és 100

Ha tényleg ott a hatás, minél nagyobb az elemszám, annál szignifikánsabb lesz az eredmény, ha nincsen ott a hatás, az elemszám növelése nem segít

Elemszám becslés

Az előző összefüggést felhasználhatjuk a vizsgálatunk megtervezésekor.

1. Becsüljük meg, vajon az adott vizsgálatban mekkora hatásra és hibára számíthatunk! Ennek két módja van:
   - Elméleti megfontolások alapján (pl. előzetes tudás)
   - Pilot vizsgálattal
2. A becsült hatás és zajnagyság alapján becsüljük meg, mekkora minimális elemszámra lesz szükségünk, hogy az 5%-os szignifikancia szint mellett 80%-os statisztikai erővel rendelkezzünk.

Elemszám becslés

A már felvett pilotminta értékei alapján megtudhatjuk, hogy körülbelül mennyi fővel kell még felvennünk a tesztet, hogy a hatás kimutatható legyen.

- http://epitools.ausvet.com.au/content.php?page=SampleSize
- Ingyen letölthető sokoldalú program: GPower

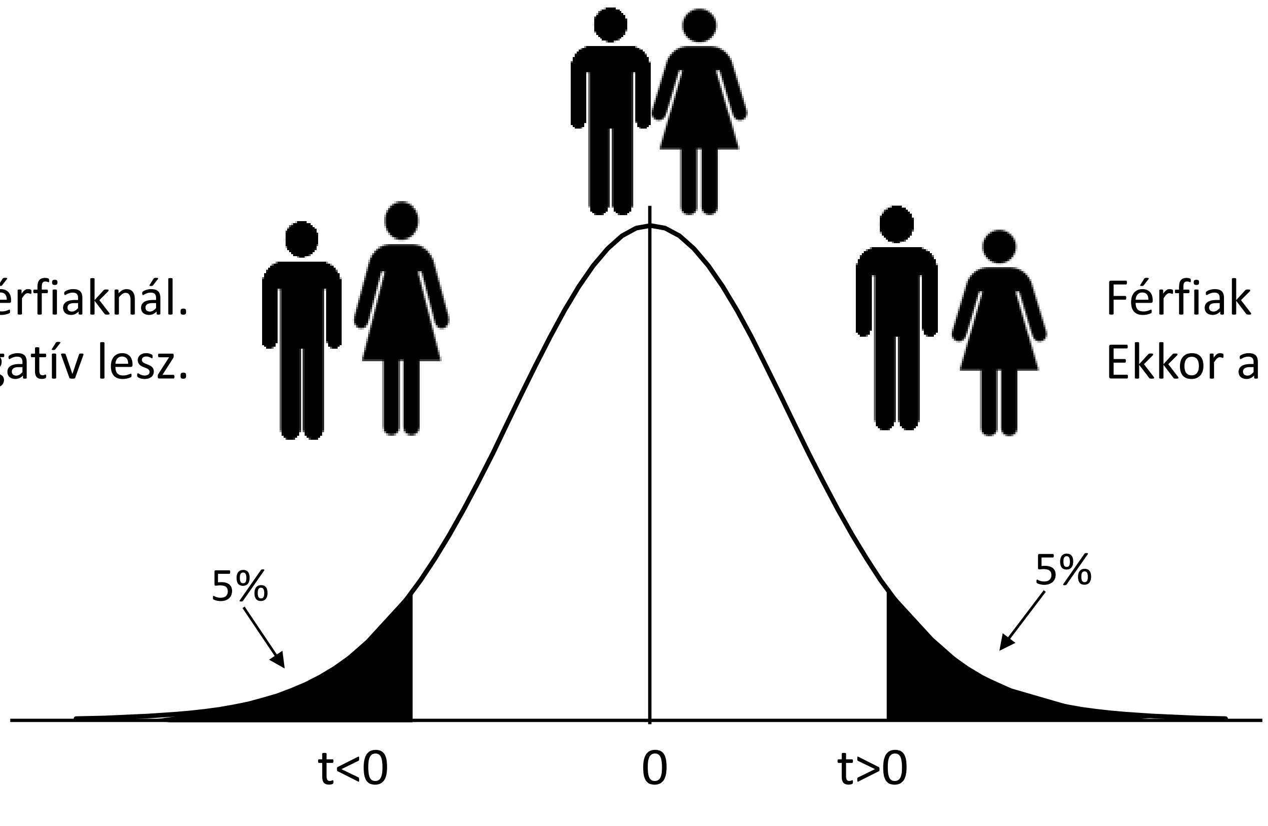
Egy- és kétvégű statisztikai tesztek

Számoljuk a t-értéket úgy, hogy a férfiak átlagmagasságából vonjuk ki a nőkét! Ekkor 3 lehetőség van:

Nincs különbség a férfiak és nők magasságában (null hipotézis).
t-érték nulla körül lesz.

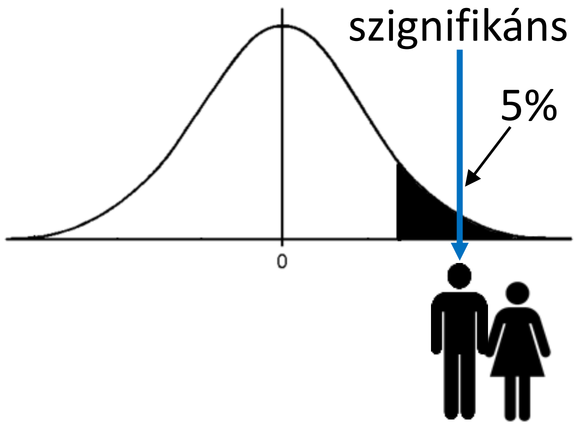
Nők magasabbak a férfiaknál.
Ekkor a t-érték negatív lesz.

Férfiak magasabbak a nőknél.
Ekkor a t-érték pozitív lesz.

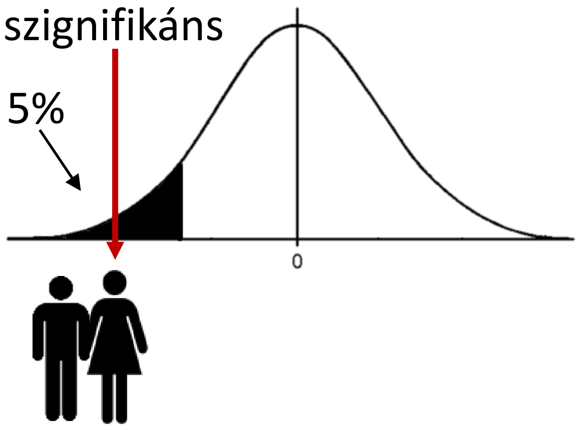


Két külön hipotézis-teszteléssel ellenőrzöm azt a hipotézist, hogy a

Férfiak magasabbak a nőknél.



Nők magasabbak a férfiaknál.

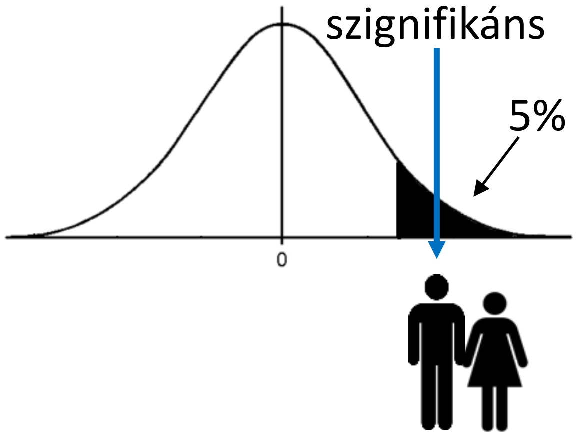


Ekkor viszont az elsőfajú hiba 5%-os valószínűségét kétszer engedtük meg.

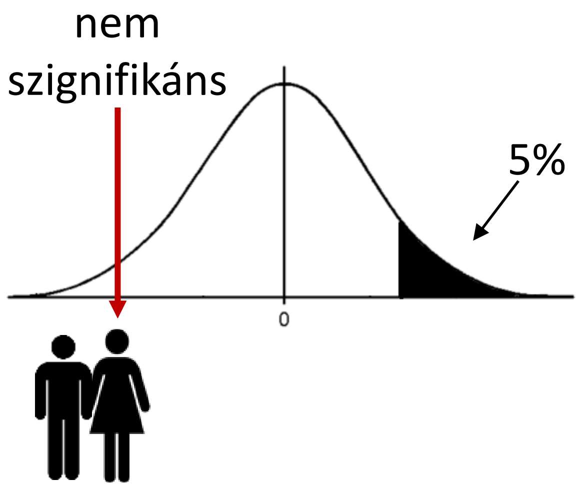
Egyoldalú hipotézistesztelésnél az 5%-nak megfelelő szignifikancia-szintet az eloszlásgörbe egyik oldalán helyezzük el, és a lehetséges különbségek csak egyik oldalát teszteljük.

Például ellenőrizzük azt a hipotézist, hogy „A férfiak magasabbak a nőknél."

Ekkor, ha a különbség elég nagy,
és a férfi minta átlaga a nagyobb,
akkor szignifikáns különbséget fogok kapni a két minta között.

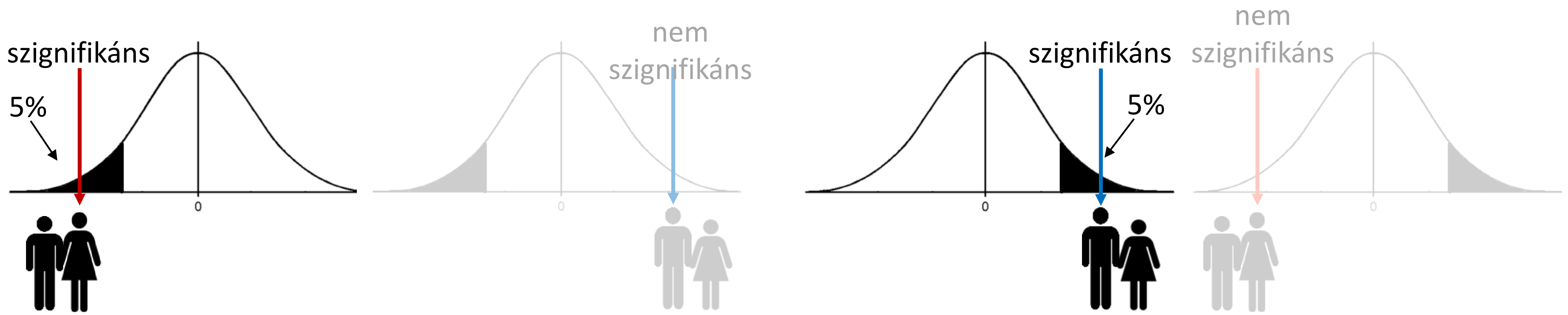


Viszont ugyanakkora különbségnél,
ha mégis a női minta átlaga a nagyobb,
akkor a különbség nem lesz szignifikáns.

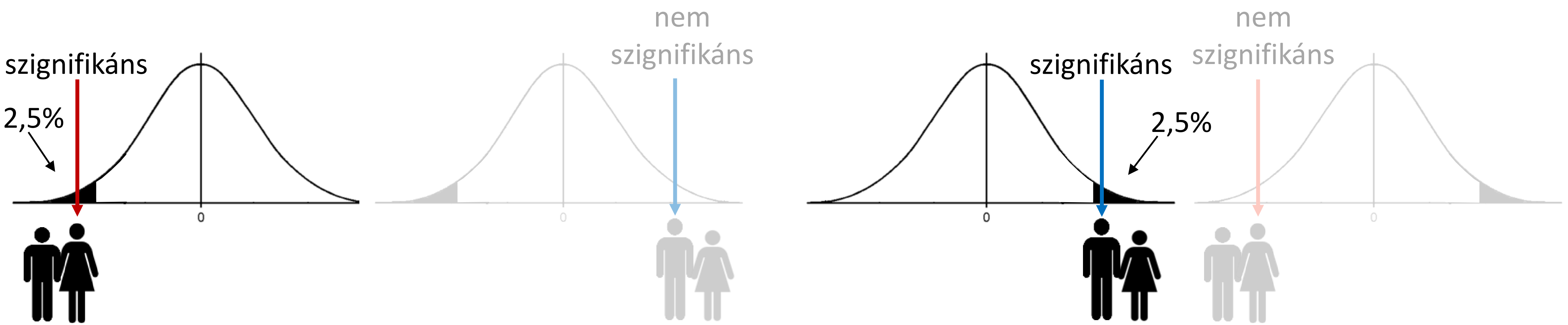


A probléma megoldható, ha két külön hipotézis-teszteléssel ellenőrzöm azt a hipotézist, hogy a

Nők több csokit esznek, mint a férfiaknál. Férfiak több csokit esznek, mint a nőknél.



Ekkor viszont az elsőfajú hiba 5%-os valószínűségét kétszer engedtük meg. Ezt korrigálni kell, ezért mindkét tesztelés során szigorúbb, 2,5%-os alfa szintet használunk.



Egy- és kétvégű statisztikai tesztek

- **Példa egyoldalú hipotézisekre:**
  - A férfiak szorongása magasabb a nőkénél.
  - A magasság és súly között pozitív összefüggés van.
  - A feladatban mért gyorsaság és pontosság negatívan függ össze.
  - A tréninget követően magasabb a személyek éntudatossága, mint a tréninget megelőzően.
  - Az ötven év feletti társadalomban a nemi arány eltér az 50-50%-tól. Ötven év felett több a nő.
  - Különbség van anorexiás és egészséges személyek BMI értékében. Az anorexiások BMI értéke alacsonyabb.

- **Példa kétoldalú hipotézisekre:**
  - Férfiak és nők szorongásában különbség van.
  - A túlórázás mértéke és a munkahellyel való elégedettség közt összefüggés van.

Effect-size

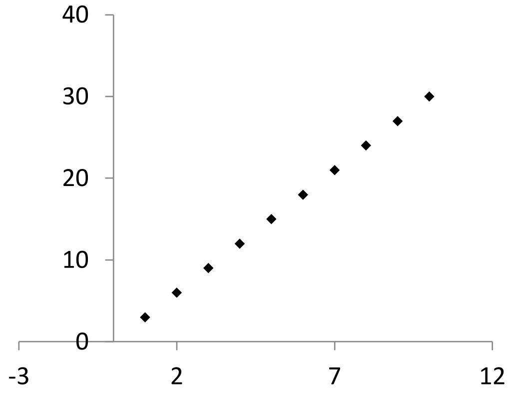
Effect-size (**hatásnagyság**)

A hatás nagyságát adja meg a mintában.
A megmagyarázott és teljes variancia aránya (a modell mennyit magyaráz a függő változó változatosságából)
Több effect-size mutató létezik.

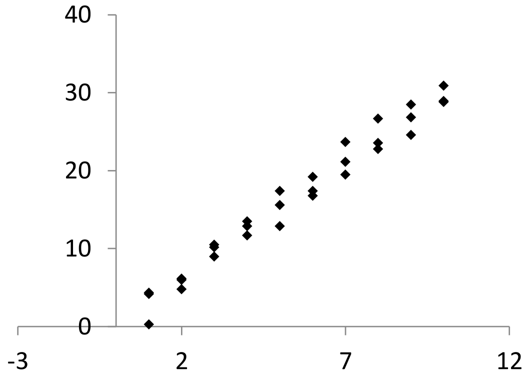
Pearson-féle korrelációs együttható (*r*) és annak négyzete (R^2)

Az *r* 0-tól 1-ig adja meg a megmagyarázott variancia arányát

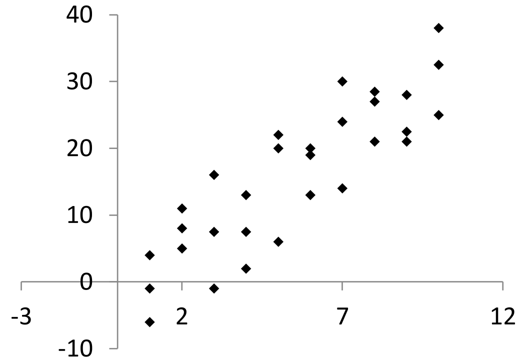
- $r < .10$ – elhanyagolhatóan kicsi hatás
- $.10 \leq r < .30$ – kicsi hatás, a variancia 1%-át tudjuk magyarázni
- $.30 \leq r = .50$ – közepes hatás, a variancia 9%-át tudjuk magyarázni
- $.50 \leq r$ – erős hatás, a variancia 25%-át tudjuk magyarázni



r = 1 az Y varianciáját teljesen megmagyarázza az X



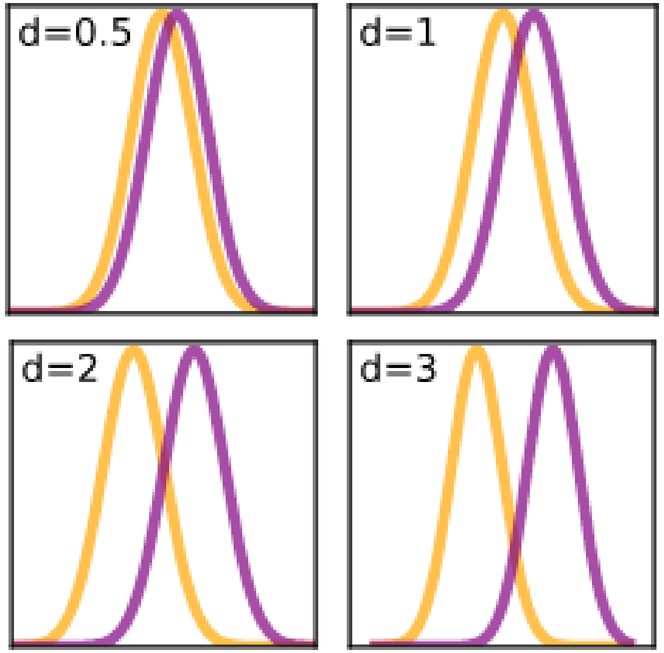
r < 1, de az Y varianciájának nagy részét meg tudjuk X-szel magyarázni

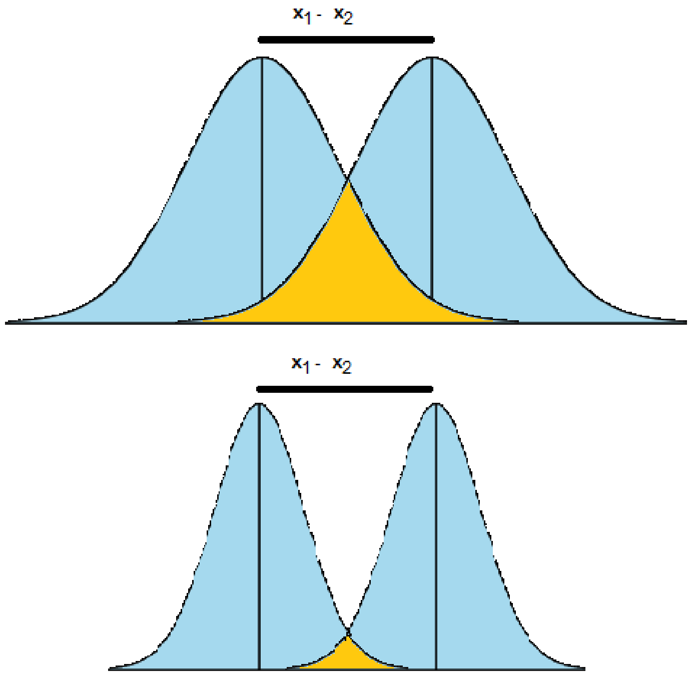


Az effect size még kisebb, az Y varianciájából még kevesebb tudunk X-szel megmagyarázni

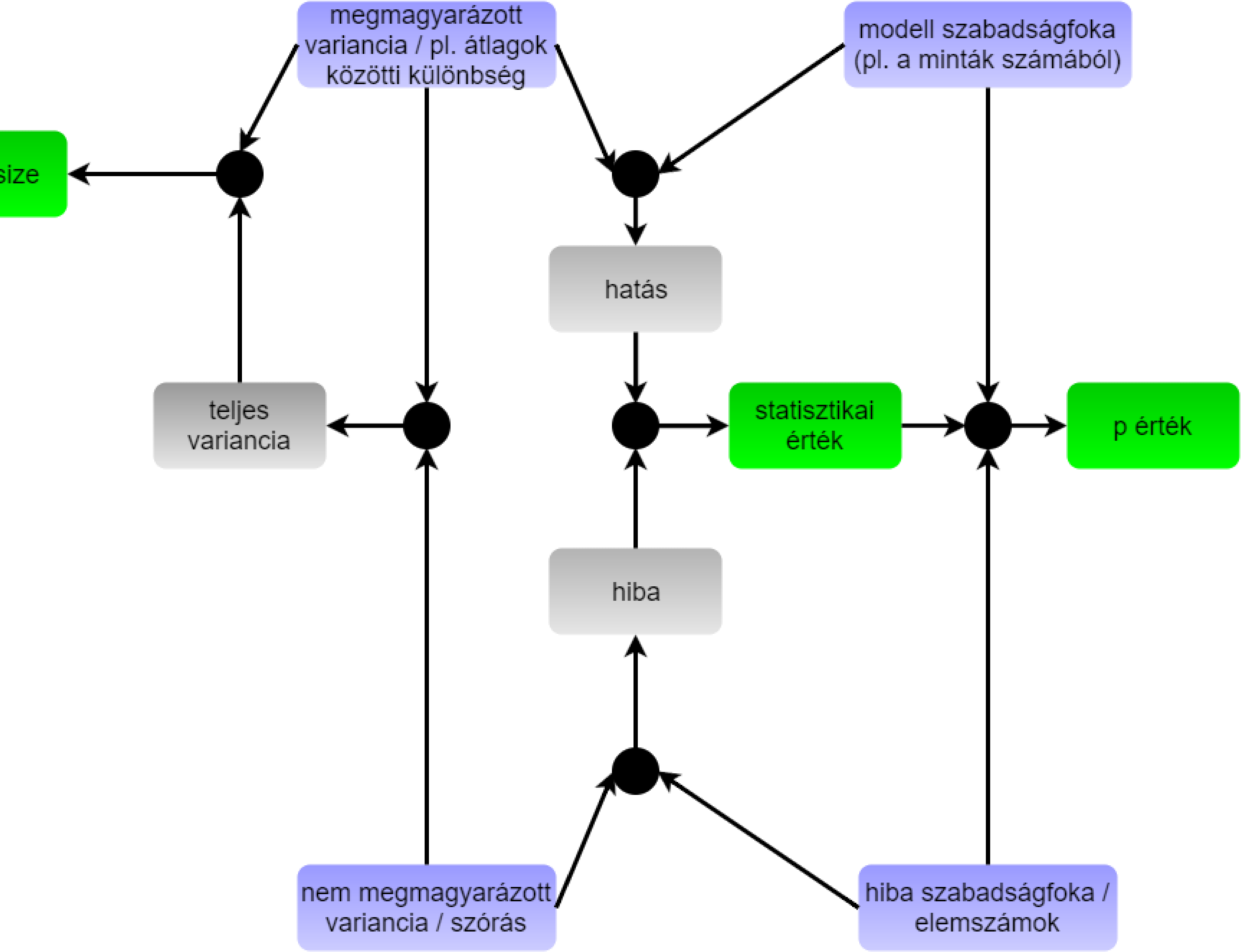
Effect-size

- **Cohen-féle d (delta) érték**
  - A két csoport átlagának különbsége a szórás függvényében
  - Megadja, két minta mennyire van átfedésben egymással
  - $d = \frac{\bar{x}_1 - \bar{x}_2}{s}$





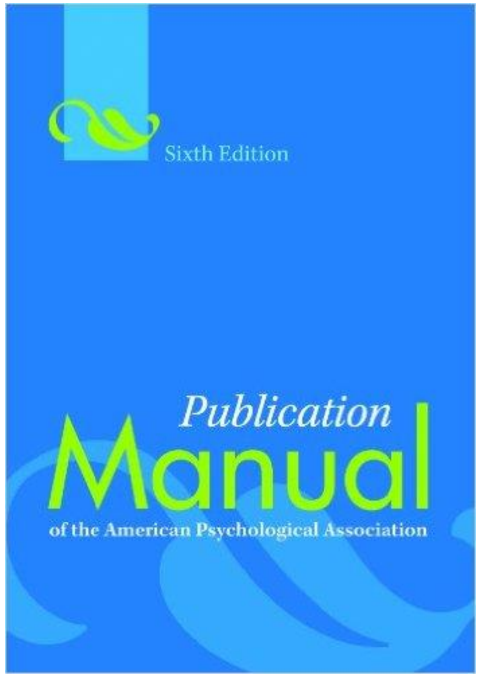
Miből mi következik?



Az ábra koncepciójában igaz, az aktuális próbák számításai eltérhetnek ettől.

- A statisztikai érték, szabadságfok(ok) és p-érték egymásba egyértelműen átszámíthatóak
- Az effect-size a varianciák arányaiból egyértelműen meghatározgató.
- Egyes effect-size mutatók és statisztikai értékek az elemszám ismeretében átszámíthatóak egymásba (pl. ilyen párt alkot az r és t érték)
- **A hatás nagysága (effect size) és valószínűsége (p-érték) közötti összefüggés mindig a szabadságfokok függvénye!** Nem igaz, hogy egy kisebb p-érték mindenképp nagyobb hatással járna.

APA style


Sixth Edition
Publication
Manual
of the American Psychological Association

Publikáció szabályai (APA formátum)

- Minden tesztnek van publikációs formája, melyet követni kell
  - Pl. $t(23) = 1.49$ $p = .023$ (1-tailed) $r = .24$

- **Dőlt betűk**:
  - A statisztikai jelöléseket dőlt betűvel kell írni
    - Pl. átlag: *M* , szórás: *SD* , t-próba: *t*
  - Ez alól kivétel a konfidencia intervallum, amit nem szabad dönteni (95%-os CI vagy CI_{95})

- **Kezdő nulla**:
  - Azoknál a mutatóknál, melyek értéke nem haladhatja meg az 1-et (pl. szignifikanciaszint, effect-size), a tizedesvessző/pont előtti nullát el kell hagyni
    - Pl. $p = .023$ vagy $r = .52$
    - Az eddigi diákon az egyszerűség kedvéért nem így szerepeltek az értékek, mostantól azonban eszerint fogom publikálni a próbák eredményét
    - Figyelj! Nem mindegy, hogy $p = 0.05$ illetve $p = .05$ vagy $p = 0.5$ illetve $p = .5$
  - Ahol viszont meghaladhatja az 1-et, kötelező a kezdő nulla kiírása
      - Pl. $t(26)=0.123$

Publikáció szabályai (APA formátum)

- **p-érték**
  - A p értékét ki kell írni 2 vagy 3 tizedes jegy pontossággal.
    - Ez alól kivétel, ha olyan kicsi/nagy az értéke, hogy nem fér bele a három tizedes jegybe vagy nem állapítható meg pontosan

Eset	SPSS-ben	Helyes jelölés
p nagyon kicsi	.000	$p < .001$
p a megjeleníthető tartományban van	.034	$p = .034$
p túl nagy	.200* ill. 1.000	$p > .200$ ill. $p > .999$

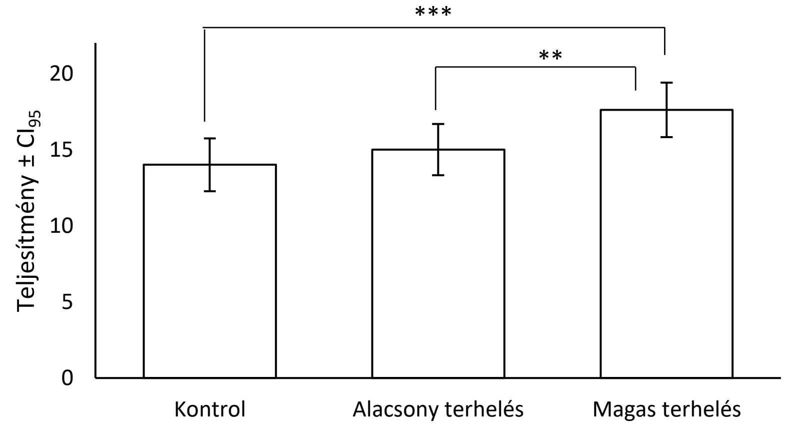
  - Egyes táblázatokban vagy grafikonokon néha csak csillaggal jelöljük a szignifikanciát. Ilyenkor a táblázat alá kell a csillagok jelentését írni.

Tábla 12. Tukey HSD utóvizsgálat a különböző mértékben terhelt csoportok közötti különbség vizsgálatára

		Különbség	*SE*
Kontrol	Alacsony terhelés	-1	0.886
Kontrol	Magas terhelés	-2,6 ***	0.859
Alacsony terhelés	Magas terhelés	-1,6 **	0.915

** szignifikáns $p < .01$ szinten; *** szingifikáns p < .001 szinten

Tukey HSD utóvizsgálat a különböző mértékben terhelt csoportok közötti különbség vizsgálatára



** szignifikáns $p < .01$ szinten; *** szingifikáns $p < .001$ szinten

Publikáció szabályai (APA formátum)

- A statisztikai próbáknál a következő mutatókat kell publikálni:
  - statisztikai érték (pl. t-érték, F-érték)
  - szignifikanciaérték (p-érték) és a próba oldalisága (1-tailed vagy 2-tailed)
  - Effect size (pl. *r, d, η, φ, Ω)* – nagyon ritkán, pl. feltételek tesztelésénél elhagyható
  - Szabadságfok (*df*) – ez nem minden esetben van, de vagy hogy kettő is van
- Minta bemutatásánál minden esetben kell publikálni:
  - Középérték (pl. átlag vagy medián)
  - Szórás
  - Elemszám
- APA tizedespontot ír elő, magyar helyesírás tizedes vesszőt. Választhatsz, de légy konzisztens!
- Felsorolásoknál (pl. több df esetén) tegyél szóközt, hogy elkülönítsd az értékeket pl. *F*(2, 33) a helyes az *F*(2,33) helyett, mely tört számként is érthető
- Egyenlőségjelek köré tégy szóközöket a könnyebb átláthatóság érdekében

Rövidítések gyűjteménye

	Statisztikában gyakran használt jelölés	**Mintára vonatkozó jelölés**	**Populációra vonatkozó jelölés**
Elemszám	*n* vagy *N*	n	N
Átlag (mean)	*M*	$\bar{x}$	μ („mű”)
Szórás (standard deviation)	*SD* vagy *s*	s	σ („szigma”)
Variancia	*Var* vagy s^2	s^2	σ^2
Standard error	*SE* vagy *S.E.*	$\sigma_{\bar{x}}$	
Konfidencia intervallum (confidence interval)	95% CI vagy CI_{95}		
p-érték	*p* vagy *Sig.*		
Pearson-féle korrelációs együttható	*r*		
Cohen-féle delta	*d*		
Standardizált érték (z-érték)	*Z* vagy *z*		
Szabadságfok	*df*		
Elsőfajú hiba	α		
Másodfajú hiba	β		
Null hipotézis	H_0		