

STATISZTIKA DATA SCIENCE-HEZ

SZIGNIFIKANCIA ALAPISMERETEK

Szignifikancia fogalma és kapcsolata a hipotézis teszteléssel

Hipotézis tesztelés lépései – elméletben és gyakorlatban

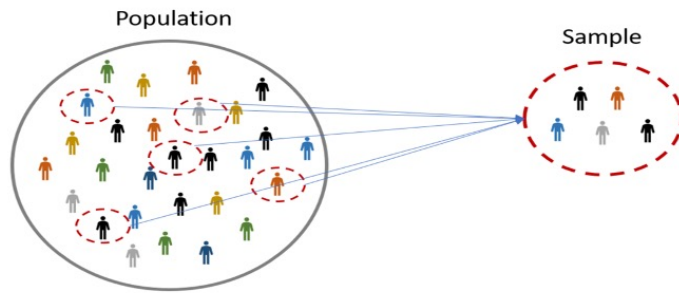
Hipotézis teszteléshez kapcsolódó egyéb statisztikai fogalmak

Összefoglalás



Mennyire a véletlen műve a kapott eredmény?

Mennyire általánosítható a kapott eredmény?



1) Null hipotézis (H_0) = nincs lineáris kapcsolat

2) korrelációs együttható (r) meghatározása
+ következtetések levonása

3) p -érték meghatározása

= annak a valószínűsége, hogy a kapott r érték csupán a véletlen műve (és mégis a H_0 az igaz)

t-teszt és szabadsági fokok meghatározása után egy eloszlás táblázatból

4) ha $p < 0,05^*$

= elutasítjuk a H_0 -t \Rightarrow van lineáris kapcsolat

ha $p > 0,05$

= (most) nem tudjuk elutasítani a H_0 -t \Rightarrow nincs lineáris kapcsolat

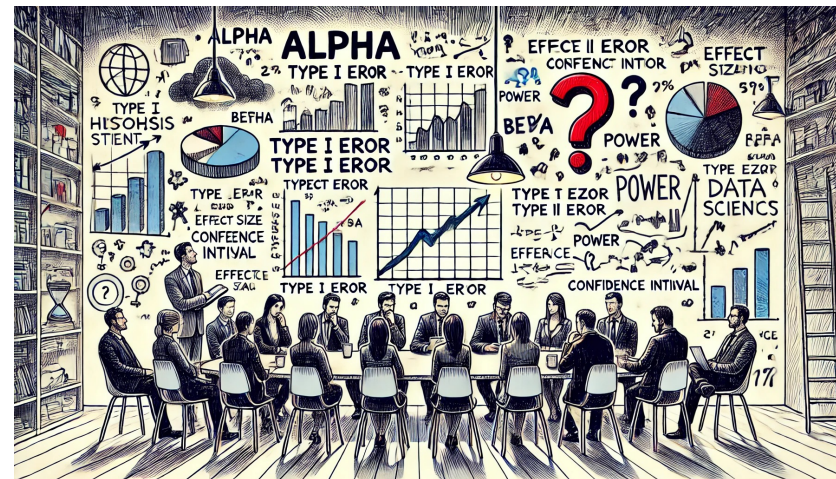
*: 5% kockázatot vállalunk rá, hogy úgy vetjük el a H_0 -t, hogy az mégis igaz

MI AZ A SZIGNIFIKANCIA?

A szignifikancia a statisztikai tesztek egyik kulcsfogalma,

amely azt jelzi, hogy a megfigyelt eredmény mennyire valószínű, hogy a véletlen műve,

és segít eldönteni, hogy egy hipotézis általánosan érvényes-e vagy sem.





**HIPOTÉZIS
TESZTELÉS
FOLYAMATA
ELMÉLETBEN**

- 1) Megfelelő statisztikai teszt kiválasztása
- 2) Nullhipotézis és alternatív hipotézis megfogalmazása
- 3) Teszt statisztika kiszámítása
- 4) p-érték meghatározása
- 5) Döntés - Elutasítjuk /
Nem utasítjuk el a
nullhipotézist

1) MEGFELELŐ STATISZTIKAI TESZT KIVÁLASZTÁSA

Teszt	Mikor alkalmazzuk?	Hétköznapi példa az alkalmazásra	Fajtái
Z-teszt	Normál eloszlású adat, nagy minta ($n > 30$), ismert szórás	Gyárban termékek súlyának ellenőrzése standard értékhez	–
T-teszt	Normál eloszlású adat, kicsi minta ($n \leq 30$), ismeretlen szórás	Két diétás program hatékonyságának összehasonlítása	Egy mintás, független kétmintás, páros t-teszt
ANOVA	Több mint két csoport átlagának összehasonlítása	Három különböző műtrágya típus hatása a növények növekedésére	Egyszerű, kétfaktoros, ismételt mérések ANOVA
F-teszt	Két minta varianciájának összehasonlítása	Két gyártási folyamat varianciájának összehasonlítása	–
Mann-Whitney U teszt	Két független csoport összehasonlítása, nem normál eloszlású adatokkal	Két terápiás módszer hatékonyságának összehasonlítása	–
Kruskal-Wallis teszt	Több mint két független csoport összehasonlítása, nem normál eloszlású adatokkal	Három különböző tanítási módszer hatékonysága	–
Wilcoxon signed-rank teszt	Párosított minták összehasonlítása, nem normál eloszlású adatokkal	Ugyanazon páciensek vérnyomásának vizsgálata két különböző időpontban	–
Friedman-teszt	Többszörös ismételt mérések összehasonlítása, nem normál eloszlású adatokkal	Több diétás program hatékonyságának összehasonlítása ugyanazon résztvevőknél	–
Chi-négyzet teszt	Kategorikus adatok függetlenségének vagy illeszkedésének vizsgálata	Gyógyszer hatékonysága függ-e a páciensek nemétől	Függetlenségvizsgálat, jósági illeszkedés teszt
McNemar-teszt	Párosított minták kategorikus adatainak összehasonlítása	Egy új gyógyszer hatására javul-e a betegek állapota	–



1) MEGFELELŐ STATISZTIKAI TESZT KIVÁLASZTÁSA

Teszt	Adatok típusa	Adatok eloszlása	Vizsgált statisztikai jellemző
Z-teszt	Folytonos	Normál	Átlag
T-teszt	Folytonos	Normál	Átlag
ANOVA	Folytonos	Normál	Átlag
F-teszt	Folytonos	Normál	Variancia
Mann-Whitney U teszt	Folytonos vagy ordinális	Nincs specifikus eloszlásfeltétel	Rangok
Kruskal-Wallis teszt	Folytonos vagy ordinális	Nincs specifikus eloszlásfeltétel	Rangok
Wilcoxon signed-rank teszt	Folytonos vagy ordinális	Nincs specifikus eloszlásfeltétel	Rangok
Friedman-teszt	Folytonos vagy ordinális	Nincs specifikus eloszlásfeltétel	Rangok
Chi-négyzet teszt	Kategorikus	Nincs specifikus eloszlásfeltétel	Gyakoriságok
McNemar-teszt	Kategorikus	Nincs specifikus eloszlásfeltétel	Gyakoriságok

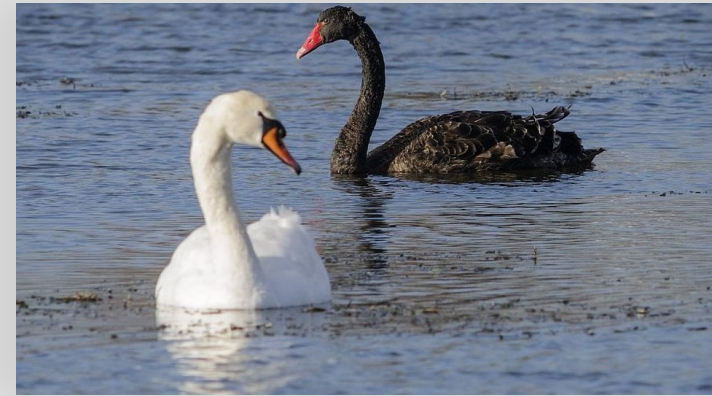


2) NULLHIPOTÉZIS ÉS ALTERNATÍV HIPOTÉZIS MEGFOGALMAZÁSA

H_0 általában azt állítja, hogy **nincs hatás vagy különbség**

- fontos, hogy az állítás cáfolható legyen, mert teljes bizonyossággal igazolni valamit nehezebb

H_1 azt állítja, hogy **van hatás vagy különbség**



2) NULLHIPOTÉZIS ÉS ALTERNATÍV HIPOTÉZIS MEGFOGALMAZÁSA

Teszt	Nullhipotézis	Alternatív hipotézis
Z-teszt	Az átlag megegyezik egy adott értékkel	Az átlag különbözik az adott értéktől
T-teszt	Az átlag megegyezik egy adott értékkel (egy mintás), vagy két csoport átlaga megegyezik (kétmintás)	Az átlag különbözik az adott értéktől (egy mintás), vagy két csoport átlaga különbözik (kétmintás)
ANOVA	Az összes csoport átlaga megegyezik	Legalább egy csoport átlaga különbözik
F-teszt	Két minta varianciája megegyezik	Két minta varianciája különbözik
Mann-Whitney U teszt	Két csoport eloszlása azonos	Két csoport eloszlása különbözik
Kruskal-Wallis teszt	Az összes csoport eloszlása azonos	Legalább egy csoport eloszlása különbözik
Wilcoxon signed-rank teszt	A párosított minták mediánjai megegyeznek	A párosított minták mediánjai különböznek
Friedman-teszt	Az ismételt mérések eloszlása azonos	Legalább egy mérés eloszlása különbözik
Chi-négyzet teszt	A változók függetlenek, vagy a minta eloszlása megegyezik a várt eloszlással	A változók nem függetlenek, vagy a minta eloszlása eltér a várt eloszlástól
McNemar-teszt	A párosított mérések gyakorisági eloszlása azonos	A párosított mérések gyakorisági eloszlása eltér



3) TESZT STATISZTIKA ÉS 4) P-ÉRTÉK MEGHATÁROZÁSA

teszt statisztika = egy adott statisztikai értéket határoz meg, mely megadja az összefüggést a vizsgált változók/csoportok között

p-érték = annak a valószínűsége, hogy a kapott összefüggés csupán a véletlen műve (és mégis a H_0 az igaz)

Teszt	Statisztikai érték kiszámításának módja	P érték meghatározásának módja
Z-teszt	$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$	A standard normális eloszlás alapján (táblázatból vagy szoftverrel)
T-teszt	$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \text{ (egy mintás)}$	A t-eloszlás alapján, szabadsági fokok figyelembevételével
	$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ (kétmintás)}$	A t-eloszlás alapján, szabadsági fokok figyelembevételével
ANOVA	$F = \frac{SS_{between}/df_{between}}{SS_{within}/df_{within}}$	F-eloszlás alapján, szabadsági fokok figyelembevételével
F-teszt	$F = \frac{s_1^2}{s_2^2}$	F-eloszlás alapján, szabadsági fokok figyelembevételével
Mann-Whitney U teszt	$U = n_1 \times n_2 + \frac{n_1 \times (n_1 + 1)}{2} - R_1$	Számított U statisztika alapján (táblázatból vagy numerikus módszerrel)

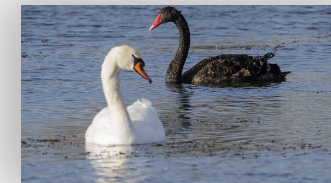
Teszt	Statisztikai érték kiszámításának módja	P érték meghatározásának módja
Kruskal-Wallis teszt	$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1)$	H statisztika alapján, gyakran táblázatból vagy numerikus módszerrel
Wilcoxon signed-rank teszt	Ranked sum of signed differences	Numerikus módszer vagy táblázat alapján
Friedman-teszt	$Q = \frac{12}{nk(k+1)} \sum R_j^2 - 3n(k+1)$	Chi-négyzet eloszlás alapján vagy numerikus módszerrel
Chi-négyzet teszt	$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$	Chi-négyzet eloszlás alapján, szabadsági fokok figyelembevételével
McNemar-teszt	$\chi^2 = \frac{(b-c)^2}{b+c}$	Chi-négyzet eloszlás alapján



5) DÖNTÉS

Nullhipotézis (H_0) és alternatív hipotézis (H_1) állításai:

- H_0 általában azt állítja, hogy **nincs hatás vagy különbség**
- H_1 azt állítja, hogy **van hatás vagy különbség**



$$p < \alpha^*$$

akkor elutasítjuk a H_0 -t
 \Rightarrow **van hatás vagy különbség**

$$p \geq \alpha$$

akkor (most) nem utasítjuk el a H_0 -t
 \Rightarrow **nincs hatás vagy különbség**

*: előre meghatározott szignifikancia szint (általában 0,05), ami megadja, hogy hány % kockázatot vállalunk rá, hogy úgy vetjük el a H_0 -t, hogy az mégis igaz





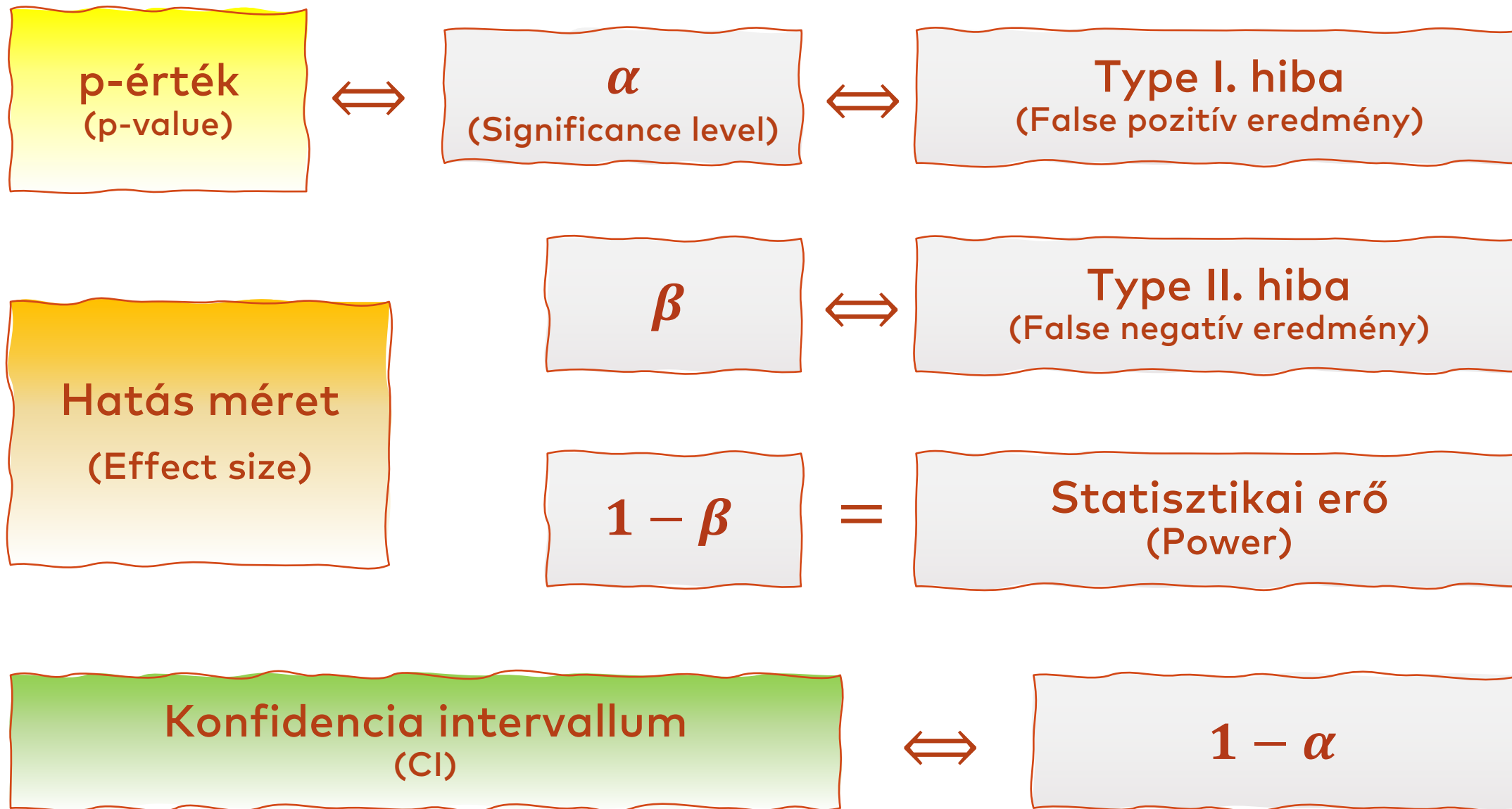
**HIPOTÉZIS
TESZTELÉS
FOLYAMATA
GYAKORLATBAN**

DataKlub_hipotezis_
teszteles.ipynb

**...ÉS ITT MÉG
NINCS VÉGE...**



További
kapcsolódó
fogalmak





HIPOTÉZIS TESZTELÉS FOLYAMATA

- 1) Megfelelő statisztikai teszt kiválasztása
- 2) Nullhipotézis és alternatív hipotézis megfogalmazása
- 3) α és β értékek meghatározás a Type I. és Type II. hibák súlyosságának figyelembe vételével
- 4) Teszt statisztika kiszámítása
- 5) p-érték meghatározása
- 6) Döntés - Elutasítjuk /
Nem utasítjuk el a nullhipotézist
- 7) Effect size meghatározás – hogy megmutassuk a valódi hatást
- 8) Konfidencia intervallum meghatározása, hogy pontosabb becslést adjunk

KAPCSOLÓDÓ FOGALMAK MAGYARÁZATA

Fogalom	Magyarázat	Hogyan kapjuk meg?
p-érték	Annak a valószínűsége, hogy a kapott eredmény csupán a véletlen műve, és a nullhipotézis mégis igaz.	Adathalmazhoz és a vizsgált problémához legjobban illő statisztikai teszthez tartozó képlettel.
α	A szignifikancia szint, ami annak a valószínűsége, hogy tévesen utasítjuk el a nullhipotézist (azaz Type I. hibát vétünk = false pozitív eredményt kapunk).	Mi határozzuk meg a vizsgált problémához illően.
Type I. hiba	Az a hiba, amikor tévesen elutasítjuk a nullhipotézist, azaz false pozitív eredményt kapunk.	-
Konfidencia intervallum (CI)	Egy olyan tartomány, amely adott valószínűséggel ($1-\alpha$) tartalmazza a vizsgált statisztikai paramétert (pl. egy átlagot vagy arányt). A CI megmutatja a becslés bizonytalanságát, és megadja azt az intervallumot, amelyről azt feltételezzük, hogy tartalmazza a valódi értéket.	Adathalmazhoz és a vizsgált problémához legjobban illő képlettel.

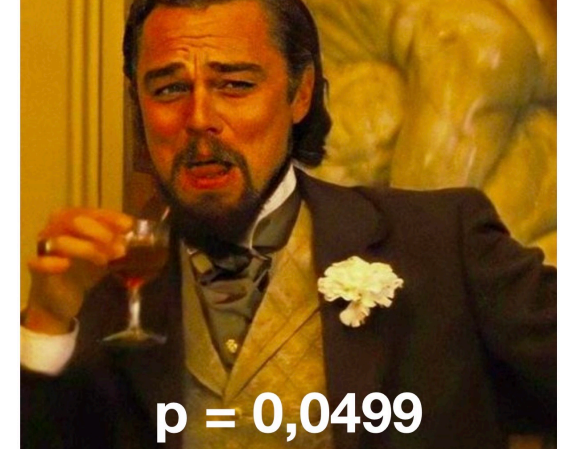


KAPCSOLÓDÓ FOGALMAK MAGYARÁZATA

Fogalom	Magyarázat	Hogyan kapjuk meg?
β	Annak a valószínűsége, hogy helytelenül nem utasítjuk el a nullhipotézist (azaz Type II. Hibát vétünk = false negatív eredményt kapunk).	Mi határozzuk meg a vizsgált problémához illően.
Type II. hiba	Az a hiba, amikor helytelenül nem utasítjuk el a nullhipotézist, azaz false negatív eredményt kapunk.	-
$1 - \beta$ = Statisztikai erő (Statistical Power)	Annak a valószínűsége, hogy helyesen elutasítjuk a nullhipotézist, ha az hamis. A statisztikai teszt érzékenysége a valódi hatás kimutatására. a vizsgálat elég érzékeny a valós különbségek kimutatására. Alacsony power esetén nagyobb az esélye, hogy egy valós hatást nem sikerül kimutatni (azaz Type II. hibát vétünk és false negatív eredményt kapunk).	β határozza meg.
Hatásméret (Effect size)	Míg a szignifikancia teszt azt mondja meg, hogy van-e statisztikailag szignifikáns különbség (azaz a különbség valószínűleg nem véletlen), az effect size érték megmutatja, hogy ez a különbség mekkora a gyakorlatban.	Adathalmazhoz legjobban illő statisztikai teszthez tartozó képlettel.



ÖSSZEFOGLALÁS



THANKS FOR LISTENING



ANY QUESTIONS?

