Hatásméret, konfidencia intervallumok, statisztikai erő

Mai adatfájl

Saját adataink

p értékre épülő inferenciális statisztika

- A hipotézis összefüggést feltételez két vagy több változó között, ami lefordítható úgy, hogy
  - különbség lesz két vagy több átlag (esetleg medián) között
  - vagy két változón kapott eredmény együtt fog járni (korrelláció)
- De a különbség vagy a korreláció lehet a véletlen műve - ez a null hipotézis
- Mennyi annak az esélye, hogy a null hipotézis áll fenn?

- Ha a populáció adatai normál eloszlásúak, ismerni fogjuk a populáció mintáit jellemző leíróstatisztikáknak (és az azokból származó statisztikáknak) az eloszlását.
- Ha ismerjük valaminek az eloszlását, meg tudjuk becsülni annak a valószínűségét, hogy egy adott értéket vegyen fel
- Pl. ha a hipotézis átlagok különbségét jósolja meg:
  - ha a populáció pontszámainak az eloszlása normál
  - a populációból származó minták átlagainak az eloszlása is normál lesz
  - tehát meg tudjuk mondani, hogy mennyi egy adott átlag esélye (vagyis mennyi annak az esélye, hogy egy adott átlag az adott populációból származó minta átlaga)

Más megközelítések

- Konfidencia intervallumok (Confidence Intervals, CI)
- Hatásméret
- Bootstrapping
- Bayes-féle statisztika

Most az első háromról lesz szó. A negyedikről majd később.

Az átlag konfidencia intervalluma

- A mintát arra használjuk, hogy megbecsüljük a populáció paramétereit (pl. átlagát)
- A minta átlaga híresen megbízhatatlan
- Az átlag konfidencia intervalluma: egy intervallum, amibe a becslés szerint a populáció átlaga esik.
- Különböző típusok: mekkora bizonyossággal állíthatjuk, hogy a populáció átlaga az adott interallumba esik:
  - 95%-os
  - 99%-os

Konfidencia intervallum kiszámítása

- Csak normál eloszlással használható
- A z értékekhez társított valószínűségekre épül
- 95%-os bizonyossággal, a populáció átlaga -1,96 és 1,96 z érték közé esik

Tehát 95% konfidencia intervallum:

- Alsó határa: -1,96(SE) + M
- Felső határa: 1,96(SE) + M

https://www.mathsisfun.com/data/standard-normal-distribution-table.html

JASP és konfidencia intervallumok

Lerajzolja (error bar), de nem számítja ki!

T-Tests
- Ha két csoportot hasonlítok össze: Independent Samples
- Ha két kondíciót hasonlítok össze: Paired Samples

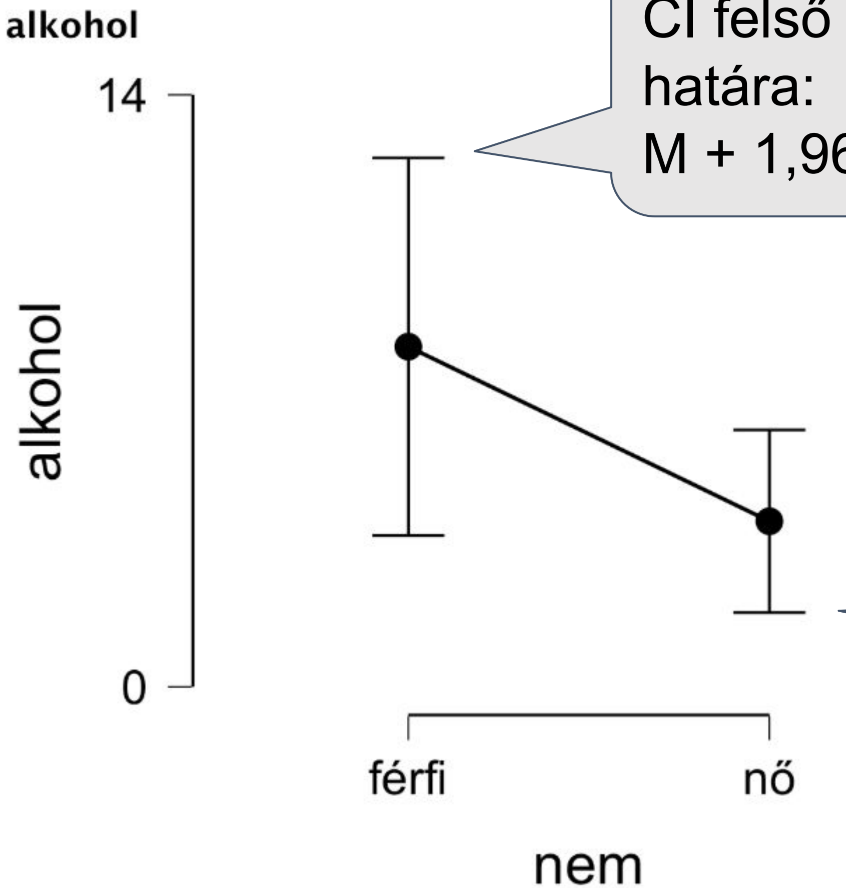
Mindkét esetben, kipipálni:
- Descriptives
- Descriptives plots

Group Descriptives

	Group	N	Mean	SD	SE
alkohol	férfi	46	8.043	15.030	2.216
	nő	47	3.915	7.351	1.072

Standard Hiba (SD osztva N négyzetgyökével)

Descriptives Plot



CI felső határa: M + 1,96SE

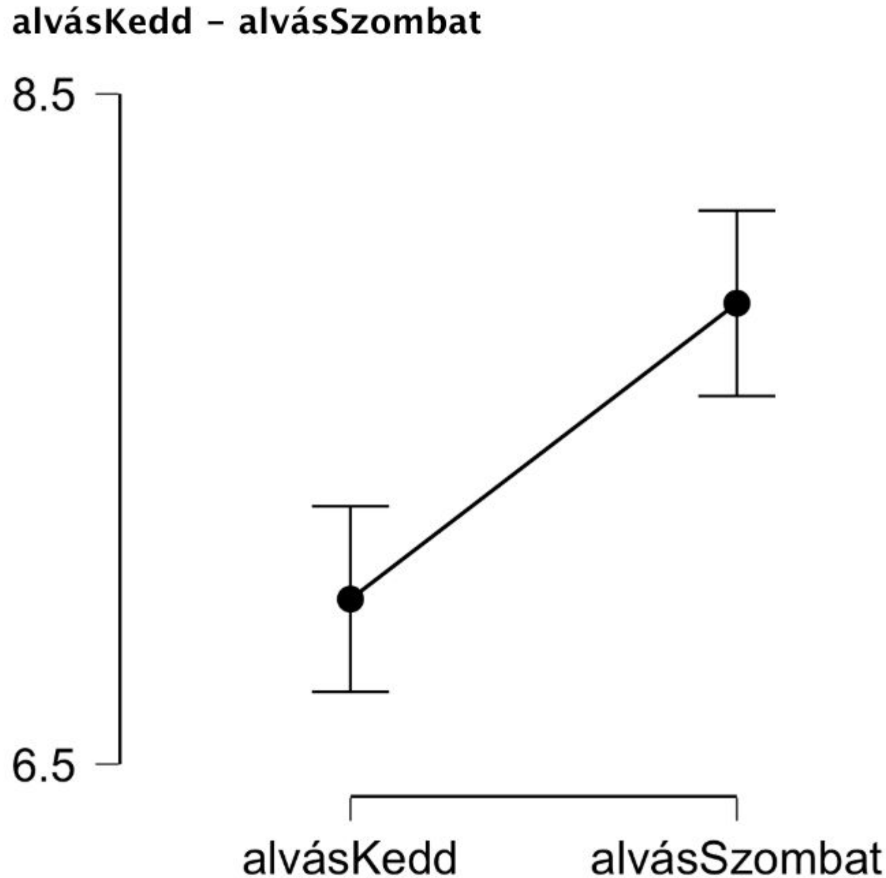
CI alsó határa: M - 1,96SE

Jelentéshez:
M (nő) = 3,92[1,81; 6,02]
M (férfi) = 8,04[3,40; 12,39]

Descriptives

	N	Mean	SD	SE
alvásKedd	93	6.992	1.501	0.156
alvásSzombat	93	7.875	1.481	0.154

Descriptives Plot



Ha a populáció átlaga a konfidencia intervallum legmagasabb értéke kedden és a lagalacsonyabb értéke szombaton, még mindig messze vannak egymástól. Tehát a keddi “minta” és a szombati “minta” nem jöhet ugyanabból a populációból.

Bootstrapping

- Az 1,96SE módszer csak egy matematikai megközelítése a konfidencia intervallumoknak.
- A bootstrapping a precízebb módszer (a JASP-ben egyelőre korlátozottan érhető el):
  - a mintából random módon kivesz egy almintát
  - kiszámolja az alminta átlagát
  - ezt megismétli mondjuk ezerszer
  - a kapott ezer átlagnak levágja az alsó 2,5 százalékát és a felső 2,5 százalékát
  - a megmaradt legalacsonyabb átlag a 95% CI alsó határa és a megmaradt legmagasabb átlag a felső határa

Még mire jó a konfidencia intervallum

- Konfidencia intervalluma nemcsak az átlagnak lehet
- Hanem más statisztikáknak is, mint például a hatásméretnek
- a p értékek alternatívájaként vagy kiegészítőjeként használható

A hatásméret

- A hipotézis prediktálhat átlagok közti különbséget vagy két változó értékeinek együttjárását (korrelációt)
- A p érték megmondja annak a valószínűségét, hogy a lemért különbség vagy együttjárás a véletlen műve,
- De semmit nem mond arról, hogy mekkora a különbség vagy milyen mértékű a korreláció.
- A különbség vagy korreláció mértéke a **hatásméret.**
- **Az APA kötelezően előírja a hatásméret jelentését.**

Két átlag különbségének hatásmérete: Cohen's d

$$d = \frac{(M_1 - M_2)}{SD}$$

Ahol M1 és M2 a két átlag, a szórást pedig mindenki máshogy számolja a két szórásból. Egy jellemző képlet (részleteket ld. később):

$$SD = \sqrt{\frac{(N_1 - 1)SD_1^{\,2} + (N_2 - 1)SD_2^{\,2}}{N_1 + N_2 - 2}}$$

Cohen's d informálisan: a két átlag szórásegységben mért (sztenderdizált) különbsége

Hatásméret, CI és p

Férfiak és nők alkoholfogyasztása

$p = ,047$

CI:

M (nő) = 3,92[1,81; 6,02]
M (férfi) = 8,04[3,40; 12,39]
(Bőven összeérnek, van rá esély, hogy a nők populációja átlagosan ugyanannyi alkoholt fogyaszt, mint a férfiaké)
$d = 0,35$
(a férfiak egyharmad szórással többet ittak, mint a nők)

Keddi és szombati alvás

$p < ,001$

CI:

M(kedd) = 6,99[6,68; 7,29]

M(szomb) = 7,88[7,57; 8,18]

(Nem érnek össze)

$d = 0,47$

(majdnem egy fél szórással aludtak többet szombaton, mint kedden)

Más hatásméretek

- r a korrelációhoz, r vagy r^2 több más dizájnhoz
- η^2 (Eta négyzet) vagy ω^2 (Omega négyzet)
- ϕ (phi), Cramer’s V vagy odds ratio

Többnyire konvertálni lehet őket egymás között

JASP: a dizájntól függően a megfelelő statisztikai teszttel együtt adja a hatásméretet

Hatásméret és konfidencia intervallum

- A hatásméret konfidencia intervallumát is ki lehet számolni:
  - a hatásméret a mintára vonatkozik
  - a hatásméret konfidencia intervalluma az az intervallum, amibe becsléseink szerint a populációra vonatkozó hatásméret esik
- JASP kiszámolja:
  - nők és férfiak alkoholfogyasztása: $d = 0,35[0,005; \infty]$
  - alvás kedden és szombaton: $d = 0,46[0,68; 0,25]]$

Egy metaelemzés

- Intervenciós programok serdülőkori depresszió kezelésére (Horowtiz & Garber, 2006)
- Több kutatás összegzése, mindegyiknél a hatásméret számít a p értékektől függetlenül

Forest plot: hatásméretek és konfidencia intervallumaik

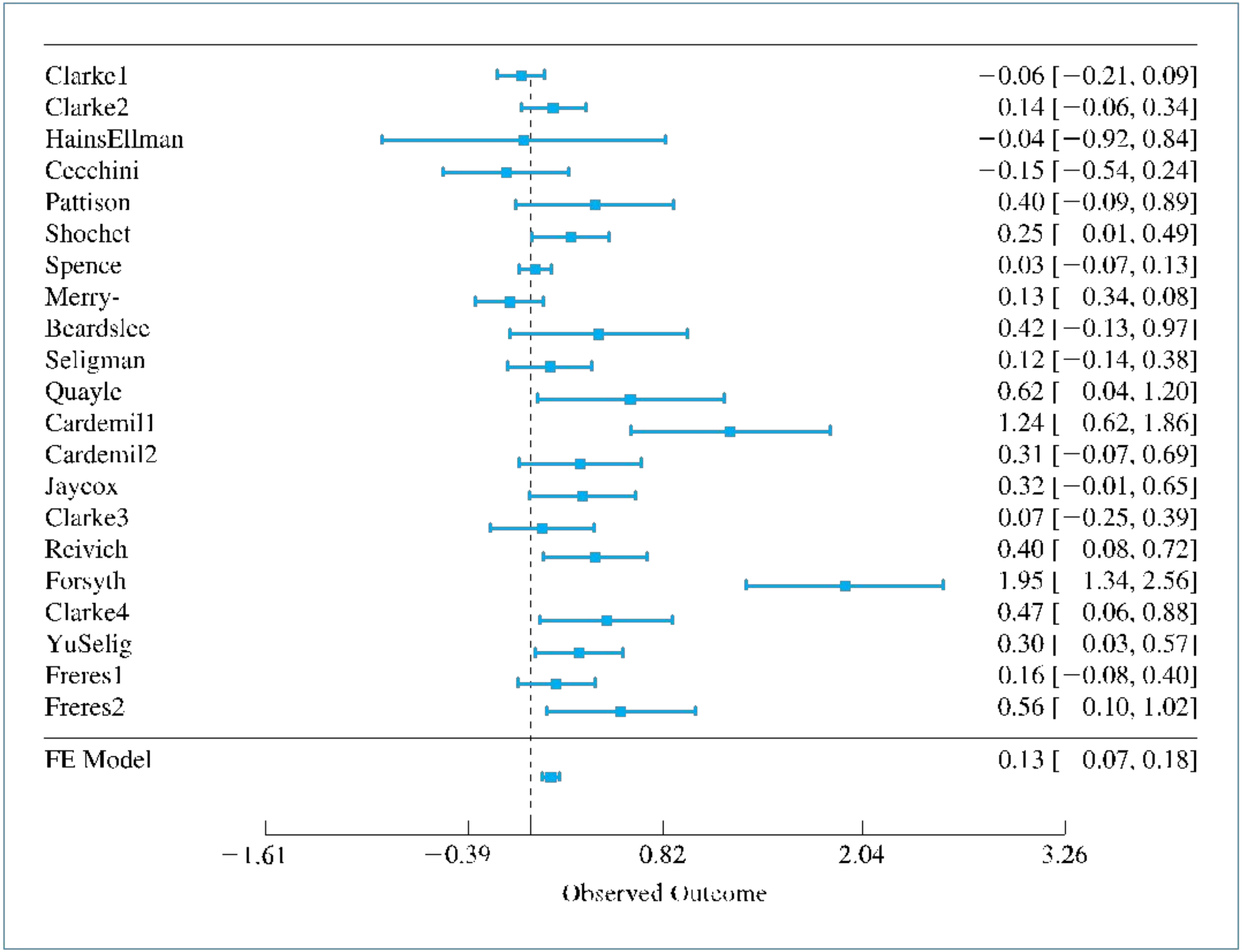


Figure 21.1
Forest plot for Horowitz and Garber's data.

- A hatásméreteket súlyozzák az elemszám és szórás figyelembevételével
- Majd kiszámítják a súlyozott hatásméretek átlagát, és annak a konfidencia intervallumát
  - Ezt mutatja az alsó FE (Fixed Effect) model

Más téma: Statisztikai erő

- Annak a valószínűsége, hogy p értékre támaszkodó statisztikai modell megtalál egy létező hatást a mintában (mennyire érzékeny a statisztikai modell).
- Statisztikai erő = 100 – a téves riasztás esélye
- A statisztikai erőt erősen befolyásoló tényezők:
  - a minta elemszáma
  - a hatásméret
  - Minél kisebb a hatásméret, annál nagyobb minta kell ahhoz, hogy a modell elég kicsinek mutassa a null-hipotézis esélyét
  - Ez azt is jelenti, hogy ha elég pénzem és időm van adatokat gyűjteni, akár mikroszkopikus hatást id ki tudok mutatni.

Mekkora mintára van szükség?

- Csak úgy lehet megbecsülni, ha tudjuk a várható
  - hatásméretet (vagyis az átlagokat és a szórást)
  - egy vagy kétszélű hipotézissel dolgozunk-e
  - mi a p érték, ami alatt elvetjük a null-hipotézist (alpha)

https://www.stat.ubc.ca/~rollin/stats/ssize/n2.html

Jó statisztikai erőnek számít: 0,80.

Elfogadható: 0,70 körül