

Minták eltérésének statisztikai elemzése

BBNPS02000

Készítette: Soltész-Várhelyi Klára

t-próba

Hol tartunk az anyagban?

Elmélet

A statisztika azért van, hogy megállapítsd elhíhated-e a mért mintában talált hatásokat a populációban is létezőnek.

Adatfeldolgozás

Adatok felkészítése a statisztikai elemzésekre, hibák, outlierok szűrése. A próbák végzése feltételekhez kötődik, megbeszéltük a négy leggyakrabban előforduló feltétel ellenőrzésének módjait (függetlenség, skála típusosság, normalitás és szóráshomogenitás)

Statisztikai próbák

Változók közötti kapcsolat vizsgálata

Van összefüggés a szubjektív jólét és az önismeret közt?

Parametrikus korreláció
Pearson korreláció

Nem parametrikus korreláció
Spearman, Kendall-féle tau

Minták közötti különbség vizsgálata

Különbözik-e férfiak és nők szorongásértéke? És a biztonságosan, elkerülően vagy ambivalensen kötődő személyek párkapcsolati elégedettsége? Magasabb-e tréninget követően az önismeret?

1 minta és konstans különbsége

Parametrikus
Egymintás t-próba

2 minta különbsége

Összefüggő minták

Parametrikus Összefüggő mintás
t-próba

Nem parametrikus Wilcoxon
előjeles rang teszt, McNemar teszt

Független minták

Parametrikus
Független mintás t-próba

Nem parametrikus
Mann-Whitney, Kolmogorov
Smirnov Z, Moses extreme
reaction, χ^2

Kettőnél több minta Különbsége

Összefüggő minták

Parametrikus
Repeated Measures ANOVA

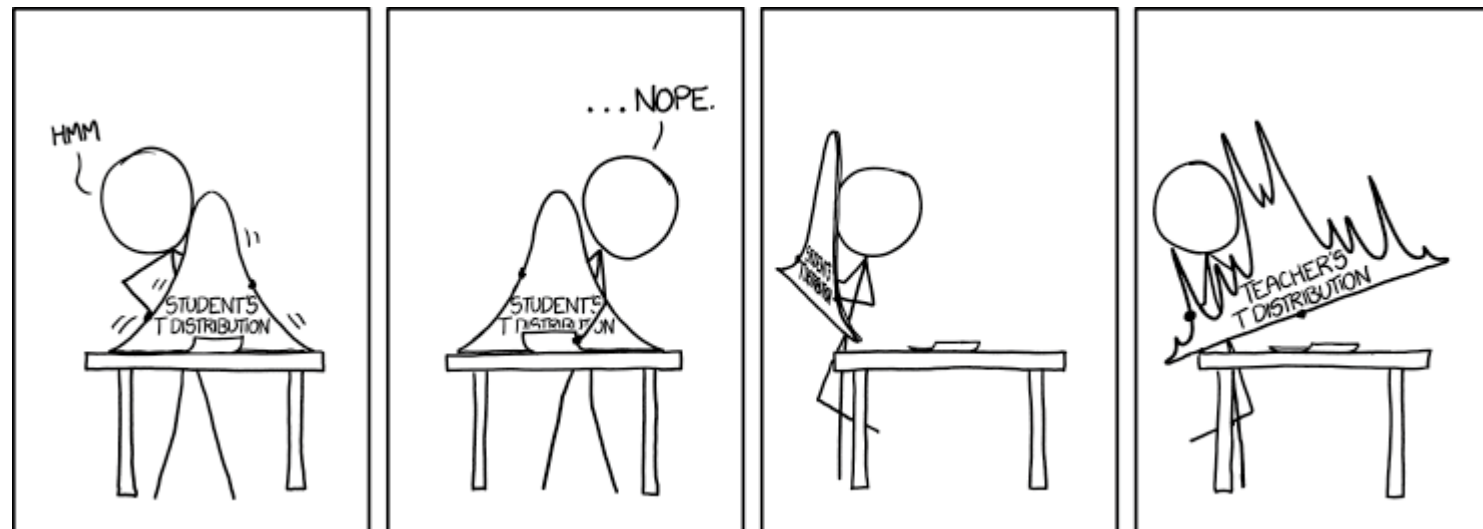
Nem parametrikus
Friedman ANOVA

Független minták

Parametrikus
One-Way ANOVA

Nem parametrikus Kruskal-
Wallis teszt

Elméleti alapok



Különbség vizsgálata?

- Előző félévben megfigyeltünk tulajdonságokat, és próbáltunk együttjárásokra következtetni (ezek voltak a korrelációs elemzések)
- Most **mintákat hozunk létre**, és **azokat hasonlítjuk össze**
 - Például megnézzük, *van-e különbség*
 - *férfiak és nők magasságában.*
 - itt a két minta a férfiak és nők, a csoportosító tényező pedig egy inherens tulajdonság, a nem.
 - *alkoholt vagy vizet fogyasztók logikai készségeiben.*
 - itt a két mintát azok alkotják akik alkoholt, és azok, akik vizet fogyasztottak, a csoportokat pedig mi hozzuk létre azzal, hogy vagy alkoholt vagy vizet itatunk velük.
 - *a terápia előtti és utáni depresszió mértékében.*
 - Itt a két mintában ugyanazok a személyek szerepelnek, de egyszer a terápiát megelőzően, egyszer pedig azt követően mérjük meg őket.

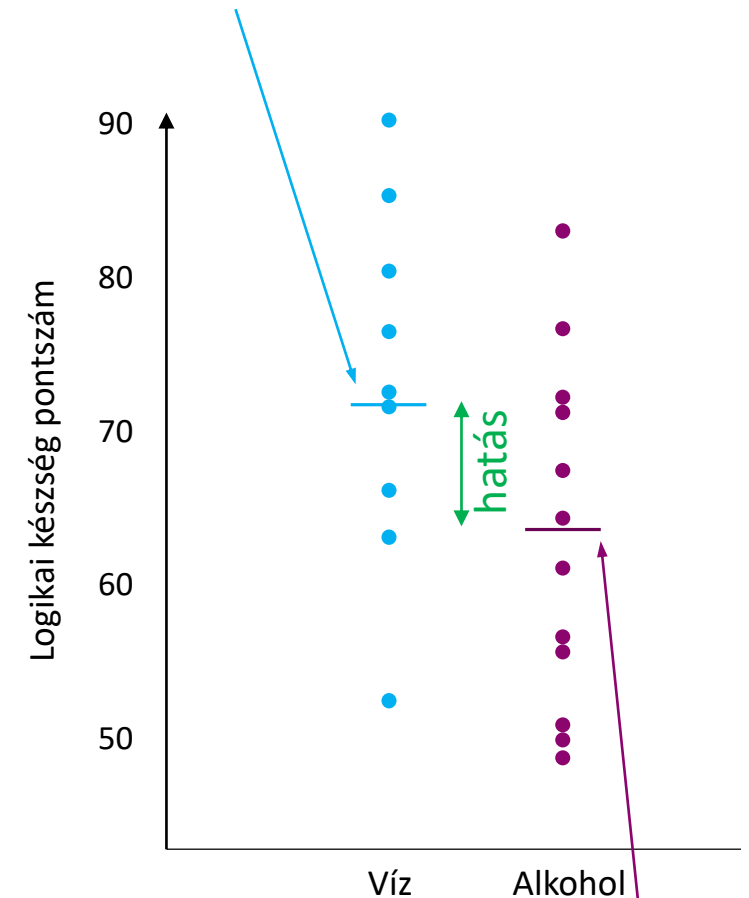
Függő és független változó

- Mivel mi hozzuk létre a mintákat, a két változó nem lesz egyenrangú:
 - Megkülönböztetünk **függő és független változókat**
 - **Független változó**: amit mi hozunk létre, aminek a hatását vizsgáljuk (pl. nem, kísérleti/kontrol csoport)
 - **Függő változó**: amire a független változó hat (pl. a magasság, a logikai készség)
- A legtöbb esetben következtethetünk az ok-okozati kapcsolat irányára
 - Például ha különbséget találunk férfiak és nők magasságában, akkor feltételezhetjük, hogy ha van ok-okozati kapcsolat, akkor a nem befolyásolja a magasságot, és nem fordítva, azaz nem a magassága alapján sorolunk valakit férfinak vagy nőnek.
- De azért itt is óvatosnak kell lenni az ok-okozattal!
 - Attól még, hogy különbséget találunk, nem biztos, hogy van ok-okozati kapcsolat, és nem véletlen egybeesésről van szó
 - Ha feltételezhető az ok-okozati kapcsolat, általában a csoportosító tényező az okozó, de nem mindig, például kosarasok és autóversenyzők magassága közötti különbség

Példa kísérlet

- Vizsgáljuk, hogy az alkohol hat-e a logikai készségekre!
- Hogyan vizsgáljuk?
 - Hívunk egy csomó embert. Véletlenszerűen két csoportba osztjuk őket. Az egyik csoportban lévőekkel alkoholt itatunk. A másik csoportban lévőekkel vizet. Ezt követően kitöltenek egy logikai feladatokat tartalmazó tesztet.
- Mit várunk a **null hipotézis** teljesülése esetén? Azaz mekkora különbséget várunk a két csoport teljesítménye között akkor, ha az alkoholnak NINCS hatása a logikai készségekre?
 - 0-át. Azt várjuk, hogy nem lesz különbség a két csoport teljesítménye között, azaz a két csoporthoz predikált érték (a két csoport átlaga) azonos lesz.
 - Azaz $\mu_{\text{víz}} - \mu_{\text{alk}} = 0$
- Mit jósol az **alternatív hipotézis**? Azaz mit várunk akkor, ha az alkoholnak VAN hatása a logikai készségekre?
 - Azt, hogy az alkoholt fogyasztóknak alacsonyabb lesz a értéke. Azaz a két csoporthoz predikált érték eltér egymástól.
 - Azaz $\overline{M}_{\text{víz}} - \overline{M}_{\text{alk}} > 0$
 - A két predikált érték közötti különbséget nevezzük **hatásnak**.

Vízivókhöz predikált logikai érték



Alkoholt ivókhöz predikált logikai érték

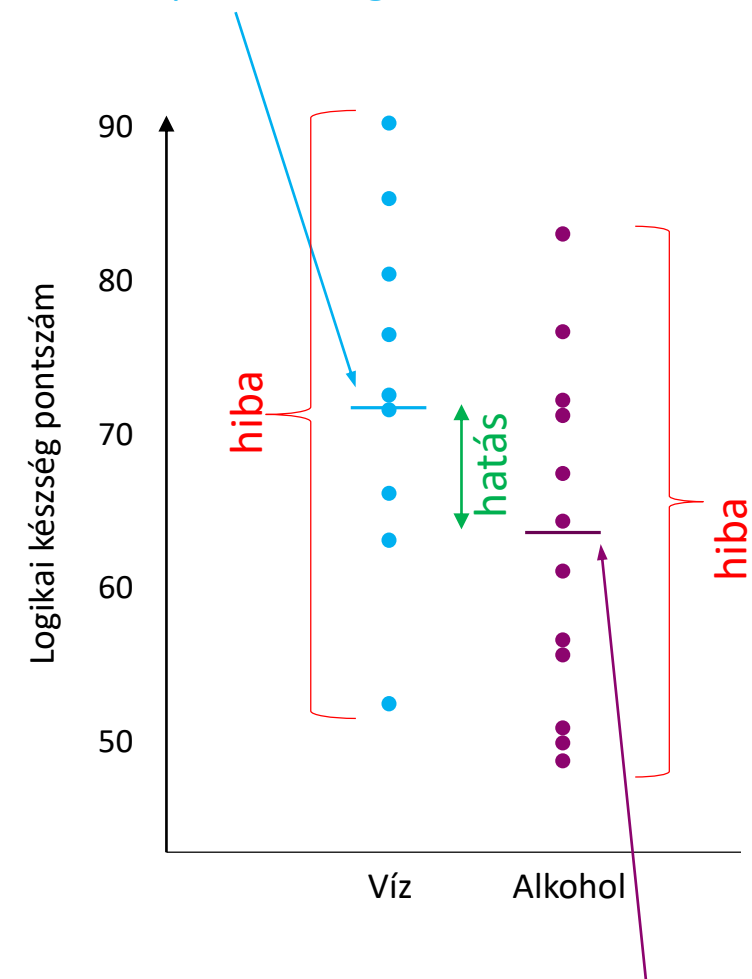
- Azt is látni fogjuk, hogy nem minden vizet ívó teljesít ugyanolyan jól. Azaz csoporton BELÜL is vannak olyanok, akik jobbak és rosszabbak. Ezek a különbségek adódhatnak a genetikából, a személy képzettségéből, fáradtságából, stb. Ezt nem tudjuk a kísérleti manipulációval magyarázni, ezért ezek az eltérések a vizsgálat szempontjából **zaj**nak tekinthetők. Hasonlóképpen az alkoholt ivók között is lesz különbség.
- Különbséget találunk a **csoportok között** és a **csoportokon belül** is.
- Korábban megtanultuk, hogy a *statisztikai érték* = $\frac{\text{hatás}}{\text{zaj}}$
- t-próbánál azt jelenti, hogy $t = \frac{\text{minták közötti eltérés}}{\text{mintákon belüli eltérések}}$

Két csoport átlaga közötti különbség, azaz a hatás

$$t = \frac{(\overline{M}_{\text{víz}} - \overline{M}_{\text{alk}})}{\text{becsült közös SE}}$$

A zaj a két csoport szórásából számolódik az elemszámokat is figyelembe véve. A képlet eltérő lehet attól függően, hogy hasonló-e a két minta mérete, szórása, ebbe nem megyünk részletesen bele.

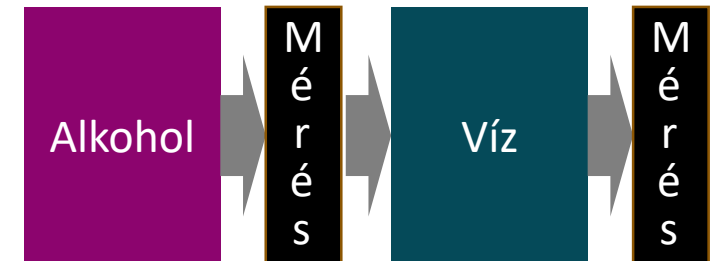
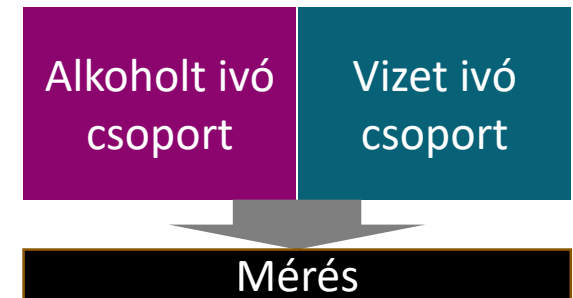
Vízivókhoz predikált logikai érték



Alkoholt ivókhoz predikált logikai érték

Minták kialakítása

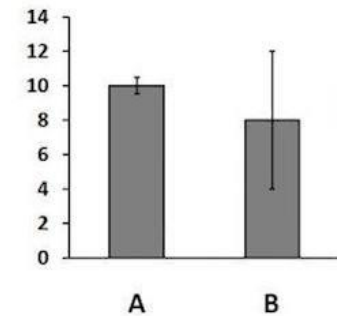
- Az eddig elhangzott elmélet azokra az esetekre igaz, amikor különböző személyeket tartalmazott a két minta. De az előző kísérletet úgy is lefolytathattuk volna, hogy egyetlen csoportot vizsgálunk két különböző KONDÍCIÓBAN: először alkoholt itatunk velük, és megmérjük a logikai készségeik, majd (miután kijózanodtak) vizet itatunk velük, és újra lemérjük őket.
- **Azaz mintákat kétféleképp hozhatunk létre:**
 - **Független minták összehasonlítása** (vagy más néven csoportközi összehasonlítás / between-subject design / between-group design)
 - Függetlenek a minták, mert különböző emberek vannak a két mintában.
 - Ekkor beszélhetünk **kísérleti csoportokról**
 - **Összetartozó minták vizsgálata** (vagy más néven ismételt mérések vizsgálata / Repeated Measures / within-subject design)
 - Összefüggők a minták, mert ugyanazon embereket mérjük le kétszer, két kísérleti kondícióban különböző időben
 - Ekkor NEM beszélhetünk két csoportról (hiszen ugyanazon személyek vannak mind a két mintában), csak **kondíciókról**



Független mintás t-próba

MR T TEST

"THAT AIN'T SIGNIFICANT, FOOL!"



csop	logikai
Alkohol	67
Víz	81
Víz	97
Víz	36
Alkohol	31
Víz	53
Alkohol	39
Víz	66
Alkohol	46
Víz	96
Víz	56
Alkohol	69
Alkohol	81
Alkohol	74
Víz	61
Alkohol	86
Alkohol	97
Alkohol	83
Alkohol	24
Alkohol	77
Alkohol	57
Alkohol	57
Víz	53

Független és függő változó

- Független mintás t-próba
 - Két csoportot hasonlít össze
 - A csoportok függetlenek, azaz különböző személyek vannak a két csoportban
 - Az adatbázisban két változó van, az egyik egy csoportosító változó (itt csop), a másik a mért érték (itt logikai).
 - A csoportosító változót független változónak, a mért értéket függő változónak nevezzük.
 - Többféleképp meg lehet fogalmazni, hogy mit várunk:
 - A két csoport között *különbség lesz*. Az alkoholt ivóknak rosszabb lesz a logikai teljesítménye.
 - A kísérleti manipuláció *hat* a mért értékre. Az, hogy alkoholt vagy vizet kap valaki, befolyásolja a személy logikai teljesítményét.
 - Az alkoholfogyasztás összefüggésben van a logikai teljesítménnyel, az alkoholt fogyasztók gyengébben fognak teljesíteni. (bár ez utolsó megfogalmazás is helyes, a saját dolgod nehezíted meg azzal, hogy úgy fogalmazod meg, mintha korrelációs vizsgálathoz írnál hipotézist.)

- Fontos, hogy a vizsgálat ne tartalmazzon szisztematikus hibát. A szisztematikus hiba egy mintán belül minden eredményt azonos irányban és hasonló mértékben befolyásol, azaz torzít, ezért statisztikai módszerekkel megtalálni igen nehéz – kiküszöböléséhez átgondolt kísérlettervezés szükséges.
- **Szisztematikus hiba** adódhat
 - a mérés vagy mérőműszer pontatlanságából, az eredmények helytelen értékeléséből
 - (pl. valamelyik logikai feladatban nincs helyes megoldás, ezért mindenki kevesebb pontot tud csak szerezni)
 - a kísérleti körülmények illetve kísérleti személyek helytelen megválasztásából
 - (pl. a csoportok kialakításánál a vizes csoportba csupa matekszakos kerül, az alkoholos csoportba pedig szaktársak – a matekosok amúgy is jobbak a logikai feladatokban, így a két csoport teljesítménye közötti különbségről nem tudjuk eldönteni, hogy az alkohol vagy a képzés hatása)
- **Randomizáció** független mintás t-próba esetén azt jelenti, hogy **a két csoportba véletlenszerűen legyenek a kísérleti személyek szétosztva.**

Független mintás t-próba feltételei

- **Feltételek** – a t-próba parametrikus teszt (tananyag ismételhető előző félévi példa alapján)
 - A **kitöltők függetlenek** egymástól
 - Ez a klasszikus függetlenség feltétel
 - A **csoportok függetlenek** egymástól
 - Ez csak annyit jelent, hogy különböző személyek kerültek a két csoportba
 - A függő változó **skála típusú**
 - A függő változó **normál eloszlású** csoportonként ellenőrizve
 - Mind **a két minta külön-külön kell normál eloszlást kövessen**
 - Ellenőrzésére használható mind a három tanult módszer.
 - Shapiro-Wilk teszt
 - Z-tesztek
 - Robosztussági körülmények: (ha **mintánként** megvan a 40 fő, a két minta megközelítőleg azonos méretű (az arány kisebb, mint 1,5, azaz a nagyobb nem éri el a kisebb méretének másfélszeresét), és a szóráshomogenitás teljesül, akkor a normalitás feltételét nem kell ellenőrizni, mivel a t-próba robusztus a normalitás sérülésére.)
 - **Szóráshomogenitás**
 - Levene tesztrel ellenőrizhető, de nem kell külön elvégezni, mert be van építve a próbába

- statgyakGY2_02_t-proba_alkohol_fuggetlen.sav
- Az adatbázis 3 változót tartalmaz:
 - csop: Csoportosító változó, 0 = víz, 1 = alkohol
 - logikai: logikai feladatban mért teljesítmény, értéke 0-100ig terjedhet
 - raghiba: a kísérlet során beszédében hány ragozási hibát vétett (az első száz mondatot figyelembe véve)
- Az adatbázist beolvasva természetesen először hibákat, outliereket keresünk.
- Ha az adatbázist tisztának tekinthetjük, a következő két hipotézist ellenőrizzük:
 - *Alkohol fogyasztása csökkenti a logikai teljesítőképeséget.*
 - *Alkohol fogyasztása növeli a természetes beszédben előforduló ragozási hibák számát.*

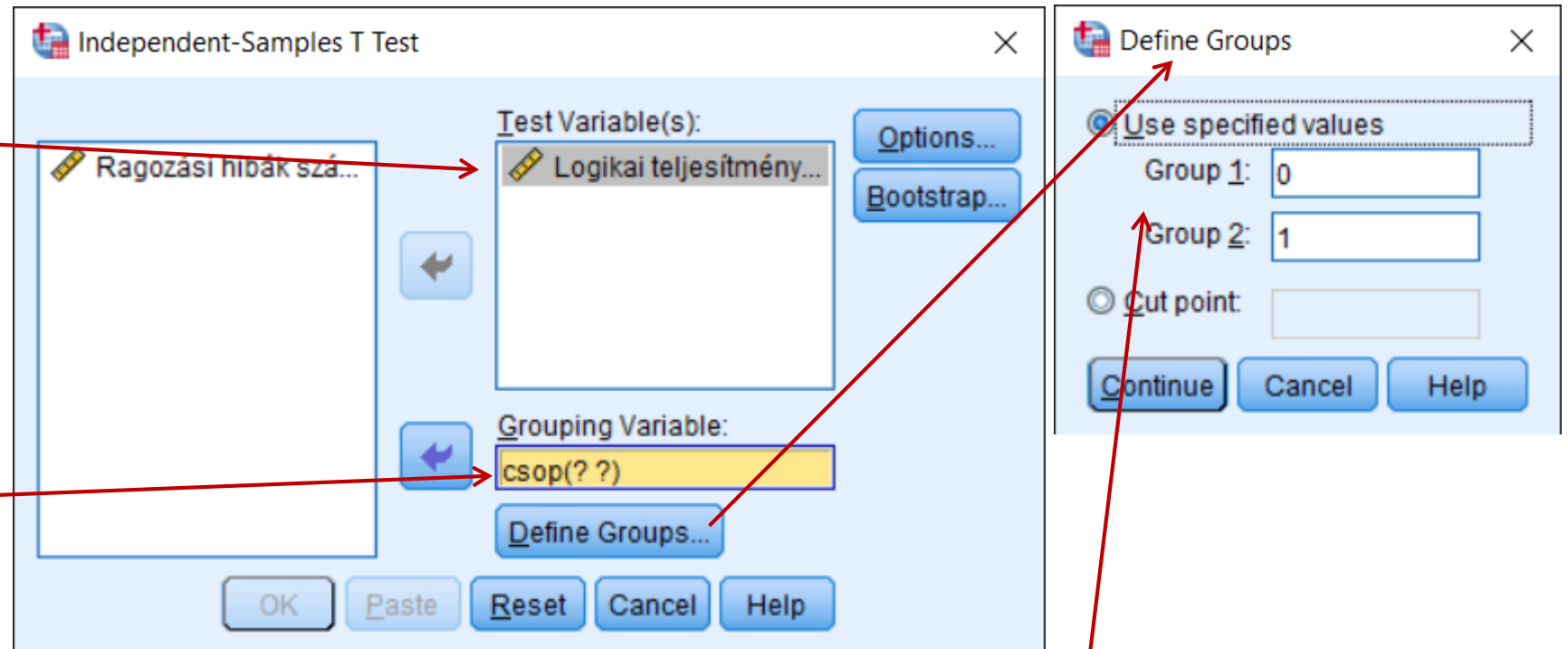
- Kezdjük az első hipotézissel! - *Alkohol fogyasztása csökkenti a logikai teljesítőképeséget.*
- Próbaválasztás
 - Két csoportot (alkoholt és vizet ivók) szeretnék összehasonlítani logikai teljesítmény tekintetében, ezért független mintás t-próbát választok
- Feltételek (előző félévben megtanultuk őket, most csak összefoglalom az eredményeket)
 - A csoportok függetlennek tekinthetők.
 - A kísérleti személyek is függetlenek.
 - A logikai teljesítmény egy 0-tól 100-ig terjedő skálán mérődik.
 - A normalitás teljesül a Shapiro-Wilk tesztre hivatkozva. Alk: $W(48) = .955$ $p = .064$; Víz: $W(43) = .969$ $p = .296$.
 - Szóráshomogenitás Levene-teszttel vizsgálva sérül. $F(1, 89) = 5.636$ $p = .020$
 - Van egy feltétel, ami nem teljesül. Mit tehetünk?
 - Következő órán látni fogjuk, hogy ilyenkor általában valamilyen másik próbát választunk, melynek nem feltétele a kérdéses feltétel teljesülése.
 - Most szerencsés helyzetben vagyunk. A szóráshomogenitás sérülése annyira gyakori, hogy az SPSS-ben beépítették a korrekciót is a t-próba elemzése mellé, így kikérhetjük a próbát, csak majd a korrigált értékeket kell néznünk!

Független mintás t-próba az SPSS-ben

- Analyze / Compare means / Independent sample t-test

Függő változó

Csoportosító változó



Csoportok definiálása – meg kell adni, milyen értékkel van a két összehasonlítandó csoport jelölve. Nálunk most 0 és 1, mert a vizet ivókat 0-val, az alkoholt ivókat 1-gyel jelöltük.

Független mintás t-próba outputja

Group Statistics					
	Csoport	N	Mean	Std. Deviation	Std. Error Mean
Logikai teljesítmény	Víz	43	67,30	19,101	2,913
	Alkohol	48	56,25	25,860	3,733

Leíró statisztikák

- Leolvasható, hogy a vizet ivók átlagos teljesítménye 67.30, míg az alkoholt ivóké csak 56.25. Ha ez a különbség szignifikáns, akkor jól gondoltuk, az alkohol csökkenti a logikai teljesítményt.
- Érdekes azt is megfigyelni, hogy az alkoholt fogyasztók esetén nagyobb a szórás, mint a víz ivók esetén (víz: 19.101, míg alkohol 25.860). Azt, hogy a szórásokban a különbség szignifikáns, már tudjuk a Levene teszt eredményéből.
Ezt lehet úgy értelmezni, hogy az alkohol nem mindenkire hatott hasonló mértékben, van, akinek jobban, van, akinek kevésbé rontotta a teljesítményét, így összességében az alkoholt fogyasztók csoportján belül változatosabb eredményeket láthatunk.

		Independent Samples Test								
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Logikai teljesítmény	Equal variances assumed	5,636	,020	2,297	89	,024	11,052	4,813	1,490	20,615
	Equal variances not assumed			2,334	85,989	,022	11,052	4,735	1,640	20,464

- t-próba táblája
- Először csak foglalkozzunk az első két oszloppal, mert ez tartalmazza a **Levene-tesztet**. Mivel a szóráshomogenitás-vizsgálat be van építve a t-próbába, nem szükséges azt előre kikérni, elegendő a feltételt itt ellenőrizni. A Levene-féle szóráshomogenitás teszt szignifikáns, azaz a szórások feltehetően nem egyeznek meg a két csoportban.
- **Innentől két sort látunk:**
 - **A felső sort** (Equal variances assumed) **akkor kell nézni, ha** a Levene-teszt nem volt szignifikáns, azaz a **szóráshomogenitás teljesül**. Ez a sor tartalmazza a t-próba eredményeit.
 - **Az alsó sort** (Equal variances not assumed) **akkor kell nézni, ha** a Levene-teszt szignifikáns lett, azaz a **szóráshomogenitás sérült**. Ez a sor tartalmazza a korrigált t-próba eredményeit.
- Itt most a szóráshomogenitás sérült, így az alsó sort nézzük!

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Logikai teljesítmény	Equal variances assumed	5,636	,020	2,297	89	,024	11,052	4,813	1,490	20,615
	Equal variances not assumed			2,334	85,989	,022	11,052	4,735	1,640	20,464

- **t**
 - A t-érték a t-próba statisztikai értéke
 - Mikor pozitív/negatív? Ha az első csoportban (víz) nagyobb az átlag, akkor a t értéke pozitív, ha a másodikban (alkohol), akkor negatív
 - A $t = \frac{(\overline{M}_{\text{víz}} - \overline{M}_{\text{alk}})}{\text{becsült közös SE}}$ képlet ellenőrizhető is: a két minta átlaga közötti különbséget a Mean Difference oszlop, a becsült közös standard error, a Std. Error Difference oszlopok tartalmazzák.
- **df**
 - szabadságfok
 - A korrigálatlan t-próbában értéke a teljes elemszám-2 (43+48-2 = 89). A korrigált t-próbában változik az értéke, gyakran tört szám lesz.

		Independent Samples Test								
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Logikai teljesítmény	Equal variances assumed	5,636	,020	2,297	89	,024	11,052	4,813	1,490	20,615
	Equal variances not assumed			2,334	85,989	,022	11,052	4,735	1,640	20,464

- **Sig (2-tailed)**

- A kapott t értéket összehasonlítja azzal, amekkora ilyen adatok esetén véletlenszerűen is adódhat, annak a valószínűsége, hogy ekkora különbség a véletlen műve
- Ez mindig two-tailed szerepel a táblázatban.
 - Ha one-tailed szignifikanciát szeretnénk, akkor ezt a szignifikancia értéket kell 2-vel osztani. Mivel nekünk van hipotézisünk a különbség irányára is, így nyugodtan használhatunk 1-tailed szignifikancia értéket, ami a táblázatban szereplő érték kettővel elosztva, $p = .011$ lesz.
- **Mivel $p < .05$, szignifikáns különbséget találtunk a két csoport átlagos logikai teljesítménye között, és a leíró statisztikából tudjuk, hogy az alkoholt ivók teljesítettek rosszabbul.**



Independent Samples Test

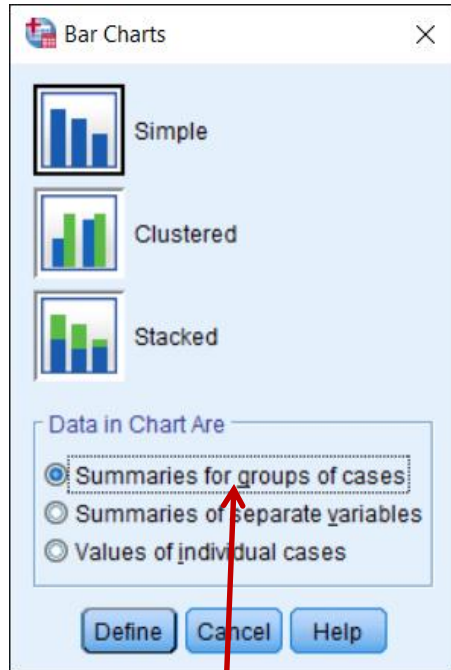
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Logikai teljesítmény	Equal variances assumed	5,636	,020	2,297	89	,024	11,052	4,813	1,490	20,615
	Equal variances not assumed			2,334	85,989	,022	11,052	4,735	1,640	20,464

- **95% Confidence Interval of the Difference**
- A két csoport közötti különbséghez tartozó 95%-os konfidencia intervallum. A különbség populációátlagára ez a két határ között helyezkedik el 95%-os valószínűséggel. Magyarul legalább másfél, legfeljebb húsz pontnyi különbségre számíthatunk a két csoport között.
- Hogy egy különbség szignifikáns-e, a konfidencia intervallumból is leolvasható:
 - Ha az alsó és felső határ is pozitív, akkor elég biztosak lehetünk, hogy van különbség a két csoport között, az első csoport jobban teljesít.
 - Ha az alsó és felső határ is negatív, akkor is elég biztosak lehetünk, hogy van különbség a két csoport között, de a második csoport teljesít jobban.
 - Ha az alsó negatív, a felső pozitív, akkor az is lehet, hogy az egyik csoportban nagyobb a teljesítmény, meg az is, hogy a másikban, így a két csoport közötti különbség akár pozitív, akár negatív is lehet.

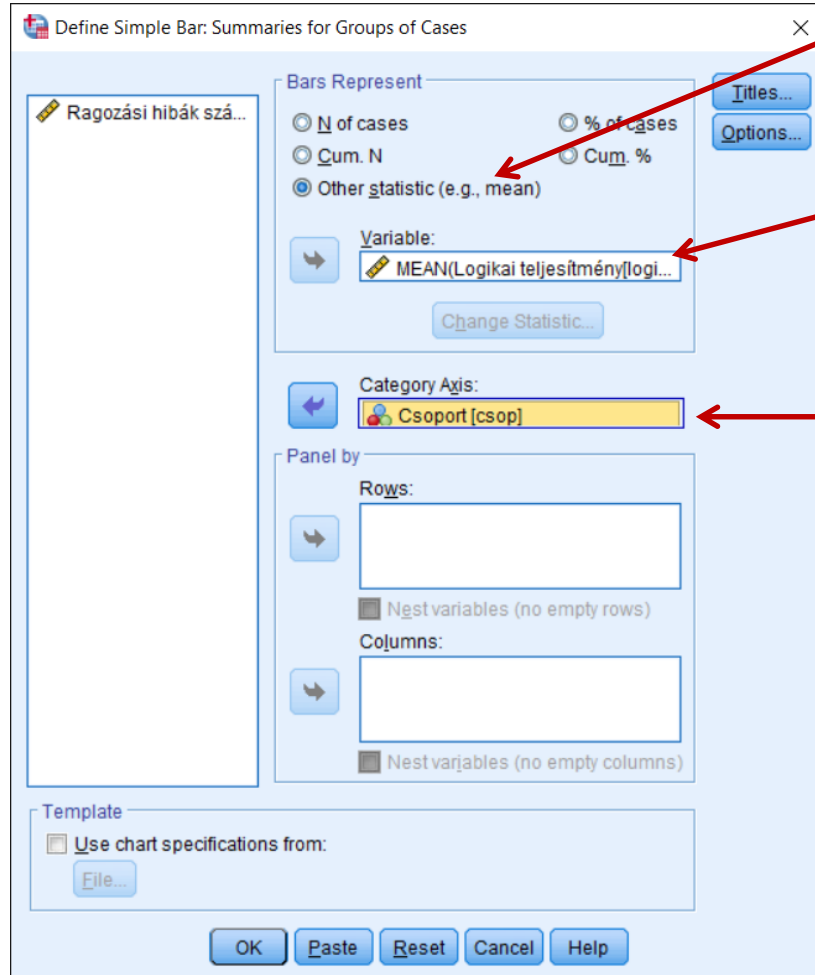
- Előző félévben megtanultuk, hogy egy statisztikai eredmény publikálásakor két dolgot kell minden esetben megadni
 - Statisztikai érték és szignifikancia érték – ebből tudjuk, az eredmény általánosítható-e a populációra
 - Effect-size – ebből tudunk következtetni a hatás nagyságára
 - (előző félévben tanult korreláció kakukktojás ebből a szempontból: az r-érték az effect-size volt, a p-érték természetesen a szignifikancia érték, és korrelációhoz tartozó statisztikai értéket pedig nem vált szokássá publikálni)
- t-próba során két **effect-size mutató** közül választhatunk:
 - **r-érték**
 - Elterjedtebb. Értelmezését előző félévben megtanultuk. SPSS nem számolja ki, ezért a képletbe kell behelyettesíteni.
 - $r = \sqrt{\frac{t^2}{t^2 + df}}$ nálunk ez most $\sqrt{\frac{2.334^2}{2.334^2 + 85.989}} = .224$, azaz gyenge különbség
 - **Cohen-féle delta**
 - $d = \frac{\bar{X}_1 - \bar{X}_2}{s}$, ahol $S = \frac{(n_1 - 1) * s_1 + (n_2 - 1) * s_2}{n_1 + n_2 - 2}$

Grafikus megjelenítés

- Mivel független csoportok átlagait hasonlítjuk össze, így oszlopdiagrammot választunk.
- Graphs / Legacy Dialogs / Bar



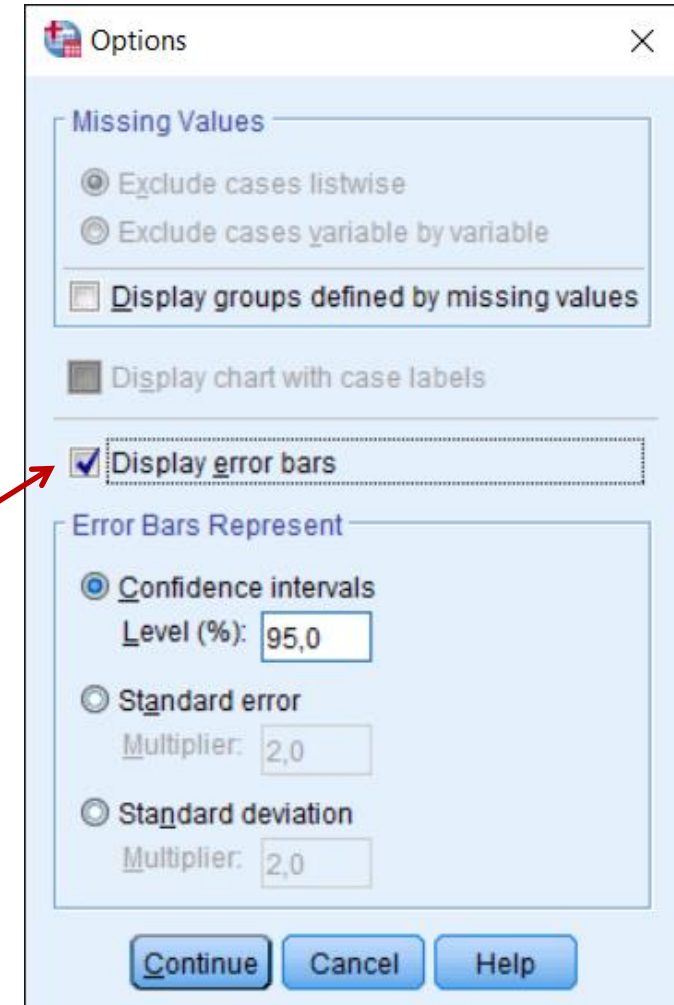
Független csoportok esetén



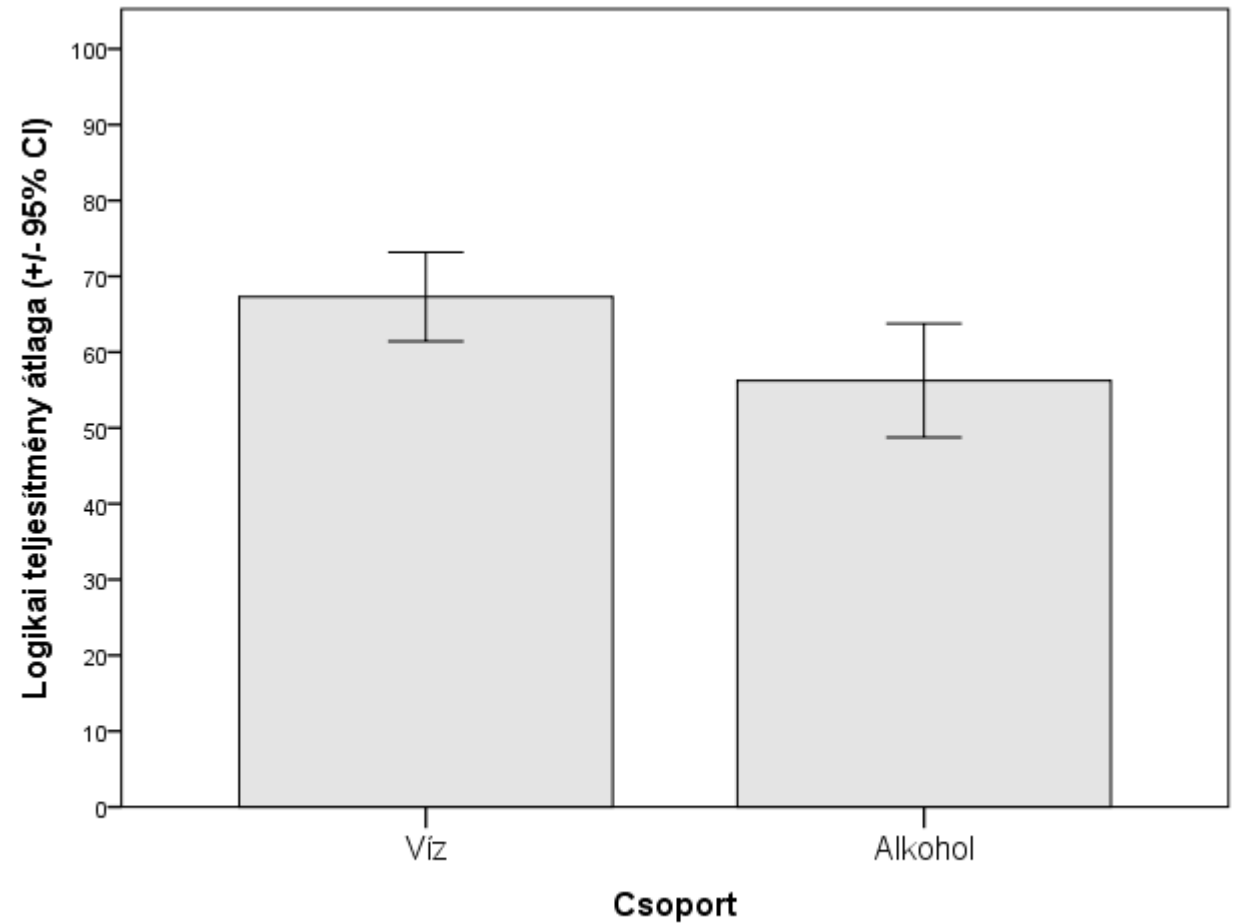
Kiválasztjuk, hogy az átlagokat szeretnénk megjeleníteni. Ezt követően betehető a függő változó

Független változó

Opciókon belül jelöljük, hogy kérünk hibasávot is. Három közül választhatunk, konfidencia intervallum, standard error, szórás



- A grafikon szerkesztését előző félévben tanultuk. Mire figyelj:
 - A hibaszáv adatait az Y-tengelyen zárójelben feltűntettem
 - A feliratokat magyarra írtam át, ahol kellett
 - A betűméret minimum 12pont, különben a dolgozatba átillesztve nem lesz olvasható
 - A skálázást 0-tól 100-igra állítottam, mivel ez a logikai teljesítmény teljes tartománya
 - Esztétika – eltűntettem a ronda háttérét, átszíneztem...

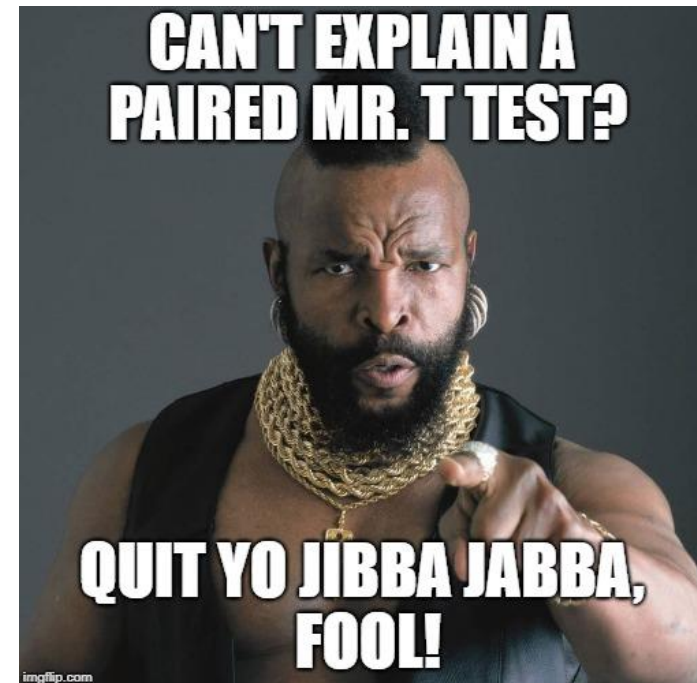


Eredmények publikálása

- Publikálás formája
 - $t(df) = [t \text{ érték}] p = [érték] ([oldaliság]) [effectsize]$
 - Mindig publikáljuk a minták átlagát és szórását is!
 - Fontosabb eredményeket grafikonon is mutassuk meg.
- Szövegesen:
 - A vizet ($M = 67.30$ $SD = 19.101$) és alkoholt ($M = 56.25$ $SD = 25.860$) fogyasztó személyek logikai feladatban nyújtott teljesítménye között szignifikáns, gyenge különbséget találtunk $t(85.989) = 2.334$ $p = .011$ (1-tailed) $r = 0.244$. Az alkoholt fogyasztók teljesítménye rosszabb volt, mint vizet ivó társaiké. Az eredmény megfelel hipotézisünknek, miszerint bla-bla-bla...
 - Érdeemes megjegyezni azt is, hogy az alkoholt fogyasztók teljesítménye nem csak alacsonyabb, de változatosabb is volt, a Levene-teszt eredményei alapján az alkoholt fogyasztó csoport szórása szignifikánsan nagyobb $F(1, 89) = 5.636$ $p = .020$. Ezt azzal lehet magyarázni, hogy bla-bla-bla...

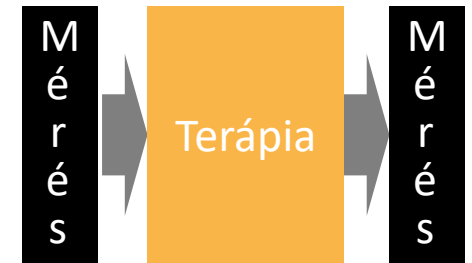
- Nézzük a második hipotézist!
- *Alkohol fogyasztása növeli a természetes beszédben előforduló ragozási hibák számát.*
- Itt is két csoportot, az alkoholt és vizet fogyasztókat hasonlítjuk össze, ezért a független mintás t-próbát választjuk. Ellenőrizd a feltételeket, kérd ki a próbát! Az eredményt megint értelmezhetjük egyoldalúan, így a szignifikancia értéket kettővel osztva $p = .026$, szignifikáns különbséget találunk.
- A dolgozatba ez kerülne:
- Az alkoholt és vizet fogyasztók beszédében előforduló ragozási hibák számát független mintás t-próbával hasonlítottuk össze. A próba feltételei teljesülnek, mindkét minta normáeloszlást követ (Alk: $W(48) = .964$ $p = .146$; Víz: $W(43) = .956$ $p = .098$), a szórások homogénnek tekinthetők ($F(1, 89) = 0.0004$ $p = .985$). A két csoport között egyoldalú tesztelést alkalmazva szignifikáns, gyenge különbséget találtunk $t(89) = -1.970$ $p = .026$ (1-tailed) $r = 0.204$. Az alkoholt fogyasztó csoport ($M = 6.063$ $SD = 3.124$) átlagosan másfél hibával többet ejtett beszédében, mint vizet ivó társaik ($M = 4.767$ $SD = 3.138$).

Összefüggő mintás t-próba



Az összefüggő mintás t-próba mintái

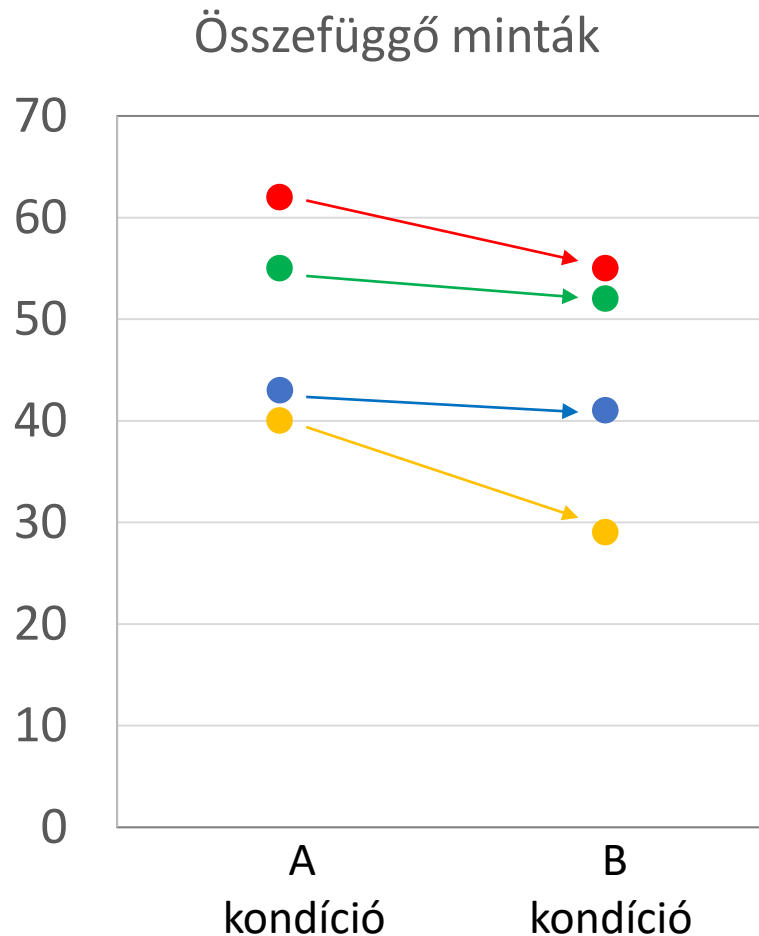
- Az előző kísérletet úgy is lefolytathattuk volna, hogy minden személlyel itatunk alkoholt, majd lemérjük a logikai teljesítményüket. Ezt követően (megvárjuk, amíg kijózanodnak), majd minden személlyel itatunk vizet, és egy az előzővel megegyező nehézségű teszten újra lemérjük őket.
- Ekkor **egyetlen csoport** van, aki **kétféle kondícióban** vesz részt
- Az összefüggő mintás elrendezés előnyei
 - Érzékenyebb, azaz kisebb hatás is kimutatható vele (később megbeszéljük, miért)
 - Kevesebb ember kell, hiszen csak egy csoport van
 - Jól használható, ha valamilyen beavatkozás hatását szeretnénk vizsgálni, tipikus előtte-utána kísérleti elrendezés (pl. terápia előtti és azt követő depressziószint összehasonlítása)



- Hátrányai
 - Az emberektől nagyobb energia-befektetést igényel, hiszen mindenki kétszer is le lesz mérve
 - Nem minden kísérleti elrendezés működik így, hiszen a mérésnek megismételhetőnek kell lenni (ez oké, ha például van két egyforma nehézségű logikai tesztünk, de ha nincs, akkor nem adhatjuk oda ugyanazt kétszer)

Összefüggő mintás t-érték számítása

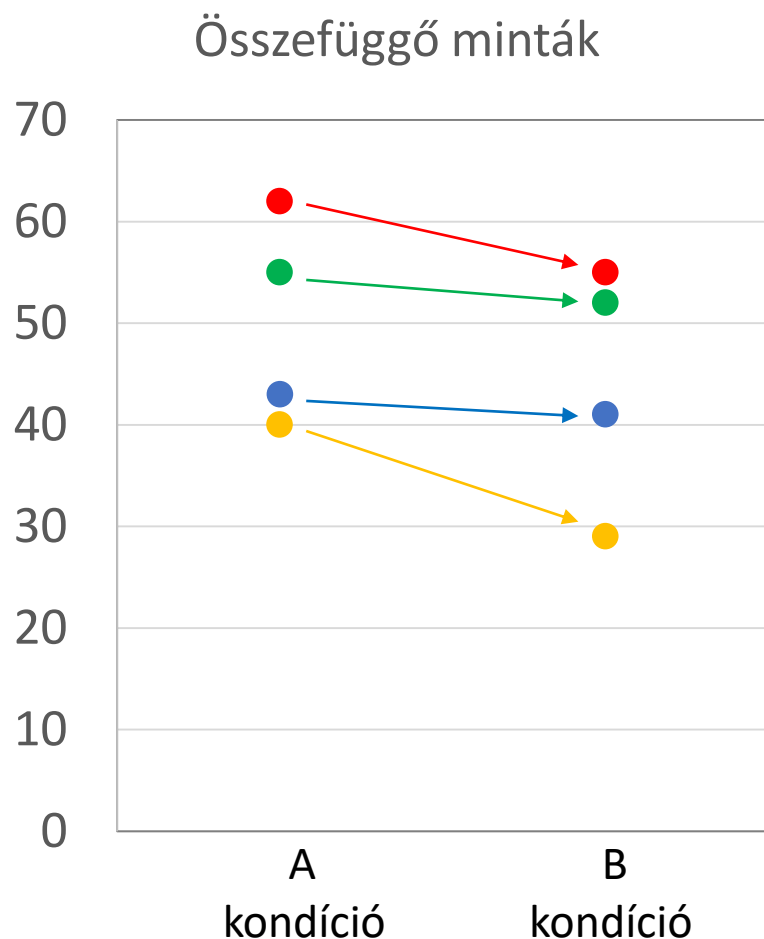
- Szerencsés helyzetben vagyunk az összefüggő mintás elemzés során – mivel minden embernek van adata mind a két kondícióban, így az embereket saját magukhoz tudjuk hasonlítani.



	A kondíció	B kondíció	Különbség
Piros ember	62	55	62-55 = 7
Zöld ember	55	52	3
Kék ember	50	41	9
Sárga ember	40	29	11

- A két kondíció különbségét, azaz a **hatást** számolhatjuk úgy, hogy megnézzük, mekkora a **változás a személyeken belül**.
- A mintában az átlagos különbség $\bar{D} = \frac{7+3+9+11}{4} = 7.5$, vagyis átlagosan 7.5 ponttal érték el a személyek kevesebbet a B kondícióban.

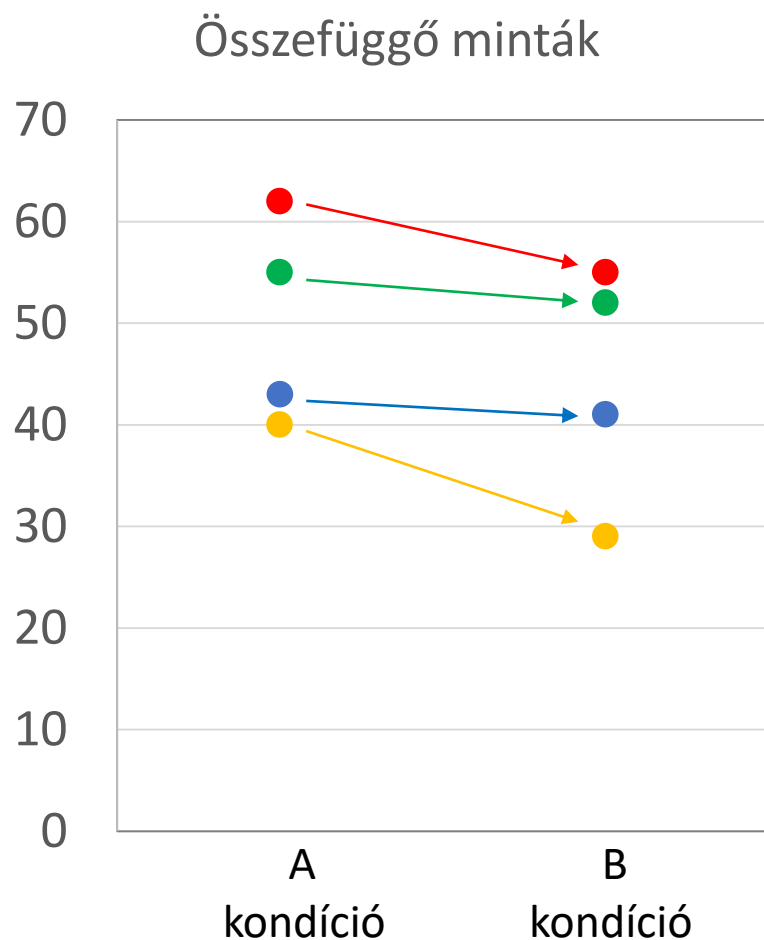
- Szerencsés helyzetben vagyunk az összefüggő mintás elemzés során – mivel minden embernek van adata mind a két kondícióban, így az embereket saját magukhoz tudjuk hasonlítani.



	A kondíció	B kondíció	Különbség
Piros ember	62	55	7
Zöld ember	55	52	3
Kék ember	50	41	9
Sárga ember	40	29	11

- Látjuk, hogy a személy kétféleképp is különböznek:
 - Magukban az értékekben (pl. a piros 62-pontot, a sárga 40 pontot szerzett az A kondícióban).
 - Abban is, hogy mekkora a megváltozás a két kondíció között (pl. a zöld csak 3 ponttal szerzett kevesebbet a B kondícióban, míg a sárga 11-gyel).
- Az összefüggő mintás t-próbában a párosíthatóságnak köszönhetően csak a második esettel kell foglalkozni.

- Szerencsés helyzetben vagyunk az összefüggő mintás elemzés során – mivel minden embernek van adata mind a két kondícióban, így az embereket saját magukhoz tudjuk hasonlítani.

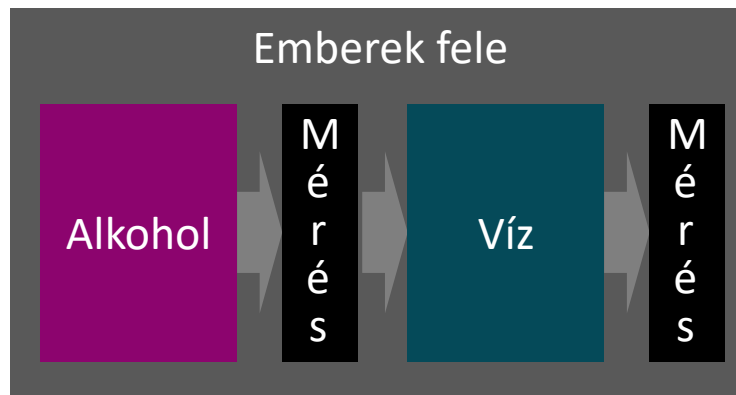


	A kondíció	B kondíció	Különbség
Piros ember	62	55	7
Zöld ember	55	52	3
Kék ember	50	41	9
Sárga ember	40	29	11

Az S_D
ezek
szórása

- A **zaj** a személyeken belüli megváltozásban látható változatosságból adódik. A zajt reprezentáló standard error ebből számolódik. $SE = \frac{S_D}{\sqrt{N}} = \frac{3.41565}{\sqrt{4}} = 1.7078$
- Végül a t-érték képlete a hatás és zaj hányadosa:
- $$t = \frac{\bar{D}}{\frac{S_D}{\sqrt{N}}} = \frac{7.5}{1.7078} = 4.392$$

- Fontos, hogy a vizsgálat ne tartalmazzon szisztematikus hibát. A szisztematikus hiba egy mintán belül minden eredményt azonos irányban és hasonló mértékben befolyásol, azaz torzít, ezért statisztikai módszerekkel megtalálni igen nehéz – kiküszöböléséhez átgondolt kísérlettervezés szükséges.
- Összefüggő minta esetén, mivel a mérés ugyanazon embereken történik kétszer is, az emberek **tanulnak** és **fáradnak**, illetve **megunhatják** a feladatot
 - Fáradnak vagy megunják: a másodiknak felvett mérésben rosszabb lesz a teljesítményük
 - Tanulnak: a másodiknak felvett mérésben jobb lesz a teljesítményük
 - Bár a tanulás és fáradás ellentétesen hat, nem ismerjük ezek mértékét, így nem tudjuk a szisztematikus hibát kiküszöbölni
 - Ellensúlyozhatjuk ezt a két hatást, ha az emberek egy részét először A hatás éri, majd B hatás, a másik felével pedig fordítva



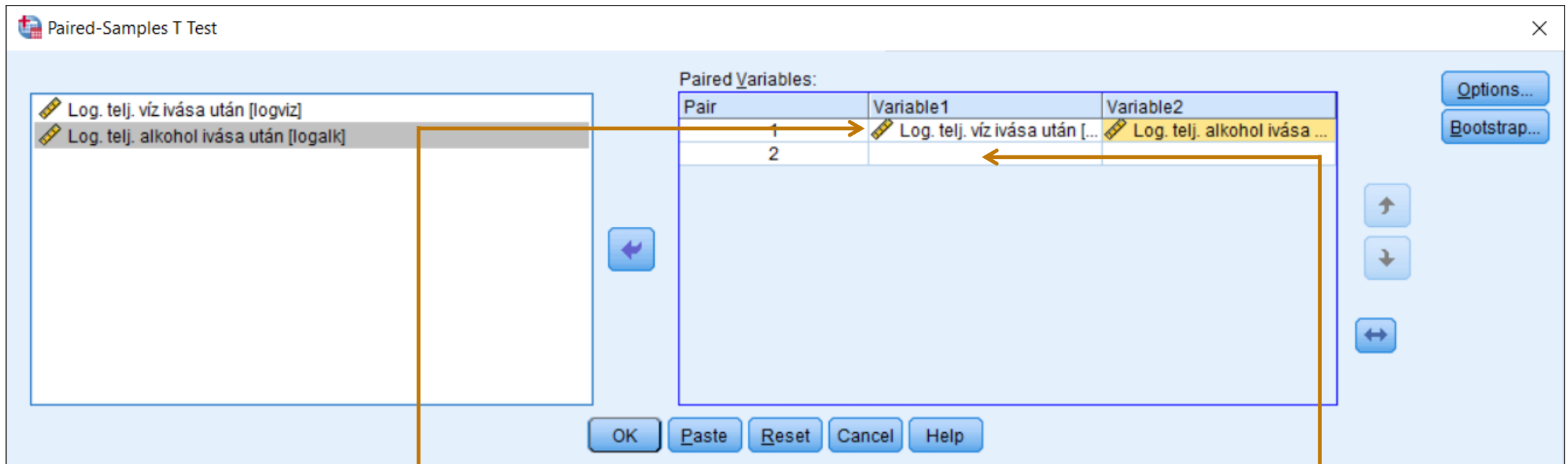
Összefüggő mintás t-próba feltételei

- Feltételek – (a t-próba parametrikus teszt)
 - A kitöltők függetlenek egymástól
 - A változók legalább skála típusúak
 - A minták normál eloszlásúak
 - (A szóráshomogenitást összefüggő minták esetén nem kell ellenőrizni. Egyrésről, mivel azonos személyeket tartalmaz a minta, ezért nagyobb valószínűséggel teljesül, másrésről a t-érték számolásánál nincs jelentősége, hiszen a személyek nyers pontértékében látható változatossággal a próba nem számol.)

 logviz	 logalk
54	41
57	39
97	100
82	66
45	34
36	16
87	81
79	73
57	29
75	65
52	21
62	64
65	62
55	25
97	97
46	12
67	62
95	88
28	21
66	56
49	24

- statgyakGY2_02_t-proba_alkohol_osszefuggo
- Az adatbázis két változót tartalmaz:
 - logviz: logikai feladatban nyújtott teljesítmény víz ivását követően
 - logalk: logikai feladatban nyújtott teljesítmény alkohol ivását követően.
- Randomizáció:
 - A vizsgálat tervezése során figyeltünk a randomizációra. A kísérleti személyek egyik fele először vizet ivott, majd másnap alkoholt, a másik fele pont fordítva, először alkohol, majd másnap vizet.
 - Az ismételt mérésből adódik, hogy a kísérleti személyeknek kétszer is ki kellett tölteni a logikai feladatsort. Ugyanazt a feladatsort természetesen nem kaphatták meg kétszer (tanulás hatása), ezért két különböző, de egyforma nehézségű feladatsort használtunk. Hogy melyik feladatsort kapta meg a kísérleti személy először, véletlenszerű volt.
- Az adatbázist beolvasva természetesen először hibákat, outliereket keresünk.
- Ha az adatbázist tisztának tekinthetjük, a következő két hipotézist ellenőrizzük:
 - *Alkohol fogyasztása csökkenti a logikai teljesítőképeséget.*

- A feltételek ellenőrzésének módját előző félévben tanultuk. Mindkét kondíció esetén feltételezhető a normál eloszlás (víz: $W(43) = .964$ $p = .0198$; alkohol: $W(43) = .954$ $p = .086$).
- Analyze – Compare means – Paired sample t test



Az összetartozó mintás T-próba mindig változópárokkal dolgozik. Ide kerül a változópár, amit vizsgálunk

Ha lenne még változónk, ide kerülhetne a következő változópár a következő t-próbához

Összefüggő mintás t-próba outputja

- Leíró statisztikák

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Log. telj. víz ivása után	68,16	43	19,218	2,931
	Log. telj. alkohol ivása után	57,58	43	25,541	3,895

- Korrelációs tábla

- mivel ugyanattól a személytől származik a két adat a két változó között gyakori korreláció, de ez nem követelmény

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Log. telj. víz ivása után & Log. telj. alkohol ivása után	43	,930	,000

- T-test tábla
 - Mean
 - Átlagos különbség 2 kondíció között. Kiszámoljuk minden ember eredményének különbségét a két kondícióban, majd ezeket a különbségeket átlagoljuk
 - Std deviation & Std Error Mean
 - Különbségek szórása és standard errorja. Van akinél kicsit nagyobb különbség lesz, van akinél kisebb, a különbség szóródásáról ad infót
 - Konfidencia interevallum
 - A két kondíció közötti különbség populációátlagja ez a két határ között helyezkedik el
 - Ha mind a két szám pozitív vagy mind a két szám negatív, akkor elég biztosak lehetünk, hogy az egyik kondícióban a legtöbbször nagyobb az érték. Ha egyik pozitív, másik negatív, akkor néha egyik kondícióban, néha másokban nagyobb az érték

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Log. telj. víz ivása után - Log. telj. alkohol ivása után	10,581	10,420	1,589	7,375	13,788	6,659	42	,000

- t-érték
 - A próba statisztikai értéke
 - Korábbi képletnek megfelelően le is ellenőrizhető: a különbségek átlaga / különbségek SE-je (itt: $t = 10.581 / 1.589$)
 - Mikor pozitív/negatív? Ha az első kondícióban nagyobb az átlag, akkor a t értéke pozitív, ha a másodikban, akkor negatív
- df
 - Szabadságfok (összefüggőmintás t-próba esetén mindig elemszám-1)
- Sig (2-tailed)
 - Szignifikanca-érték
 - A táblázatban mindig 2-tailed szerepel. Ha a hipotézis egyoldalúan van megfogalmazva, értékét 2-vel osztva megkapjuk a 1-tailed szignifikanca-értéket

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Log. telj. víz ivása után - Log. telj. alkohol ivása után	10,581	10,420	1,589	7,375	13,788	6,659	42	,000

- Az összefüggő mintás t-próba során is két **effect-size mutató** közül választhatunk:

- **r-érték**

- Elterjedtebb. Értelmezését előző félévben megtanultuk. SPSS nem számolja ki, ezért a képletbe kell behelyettesíteni.

- $r = \sqrt{\frac{t^2}{t^2 + df}}$ nálunk ez most $\sqrt{\frac{6.659^2}{6.659^2 + 42}} = .717$, azaz erős különbség

- **Cohen-féle delta**

- $d = \frac{\bar{X}_1 - \bar{X}_2}{s}$, ahol $s = \frac{s_1 + s_2}{2}$

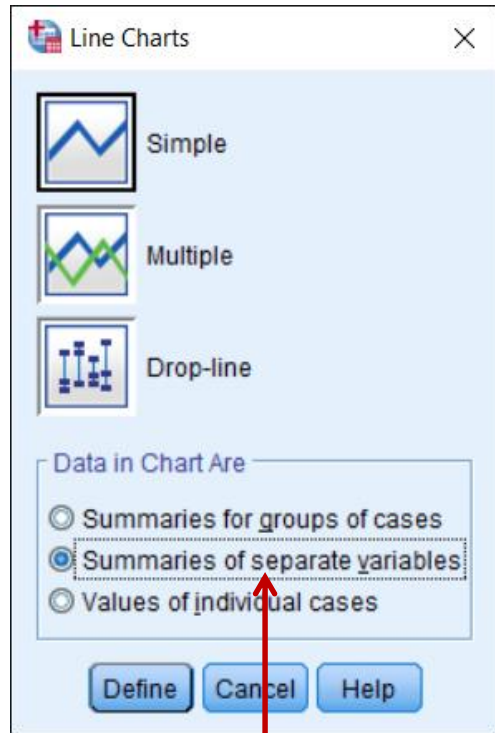
- Érdekességgé rámutatok, hogy a delta képlete a független mintás t-próbában mennyivel bonyolultabb volt. Ez azért van, mert az összefüggő mintás t-próbában mindig azonos az elemszám a két minta elemszáma, így az elemszámokkal le lehet egyszerűsíteni.



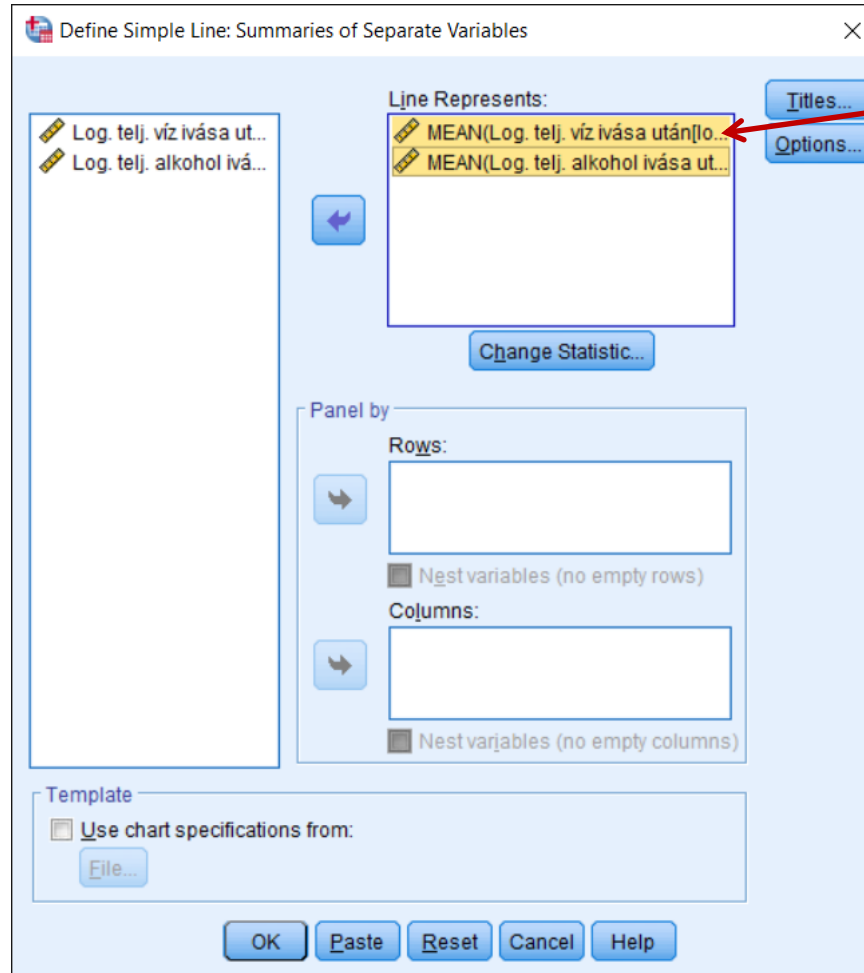
$$s = \frac{(n_1 - 1) * s_1 + (n_2 - 1) * s_2}{n_1 + n_2 - 2} \longrightarrow s = \frac{(n - 1) * s_1 + (n - 1) * s_2}{n + n - 2} \longrightarrow s = \frac{(n - 1) * (s_1 + s_2)}{2n - 2} \longrightarrow s = \frac{(n - 1) * (s_1 + s_2)}{2 * (n - 1)} \longrightarrow s = \frac{(s_1 + s_2)}{2}$$

Grafikus megjelenítés

- Mivel összefüggő mintákat hasonlítottunk össze, így vonaldiagrammot választunk.
- Graphs / Legacy Dialogs / Line

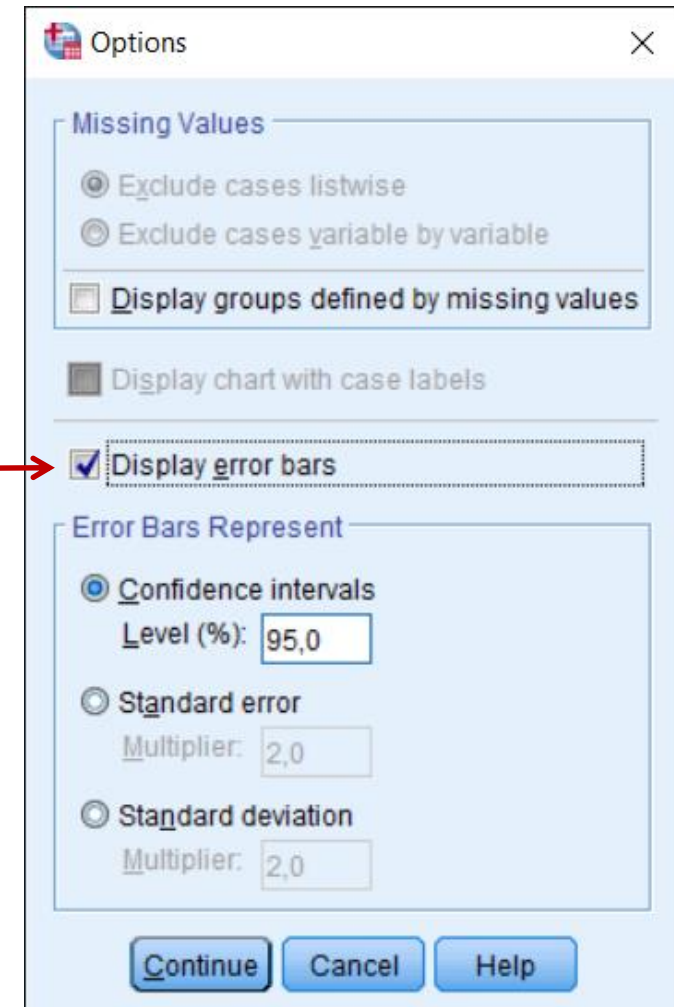


Összefüggő minták
esetén

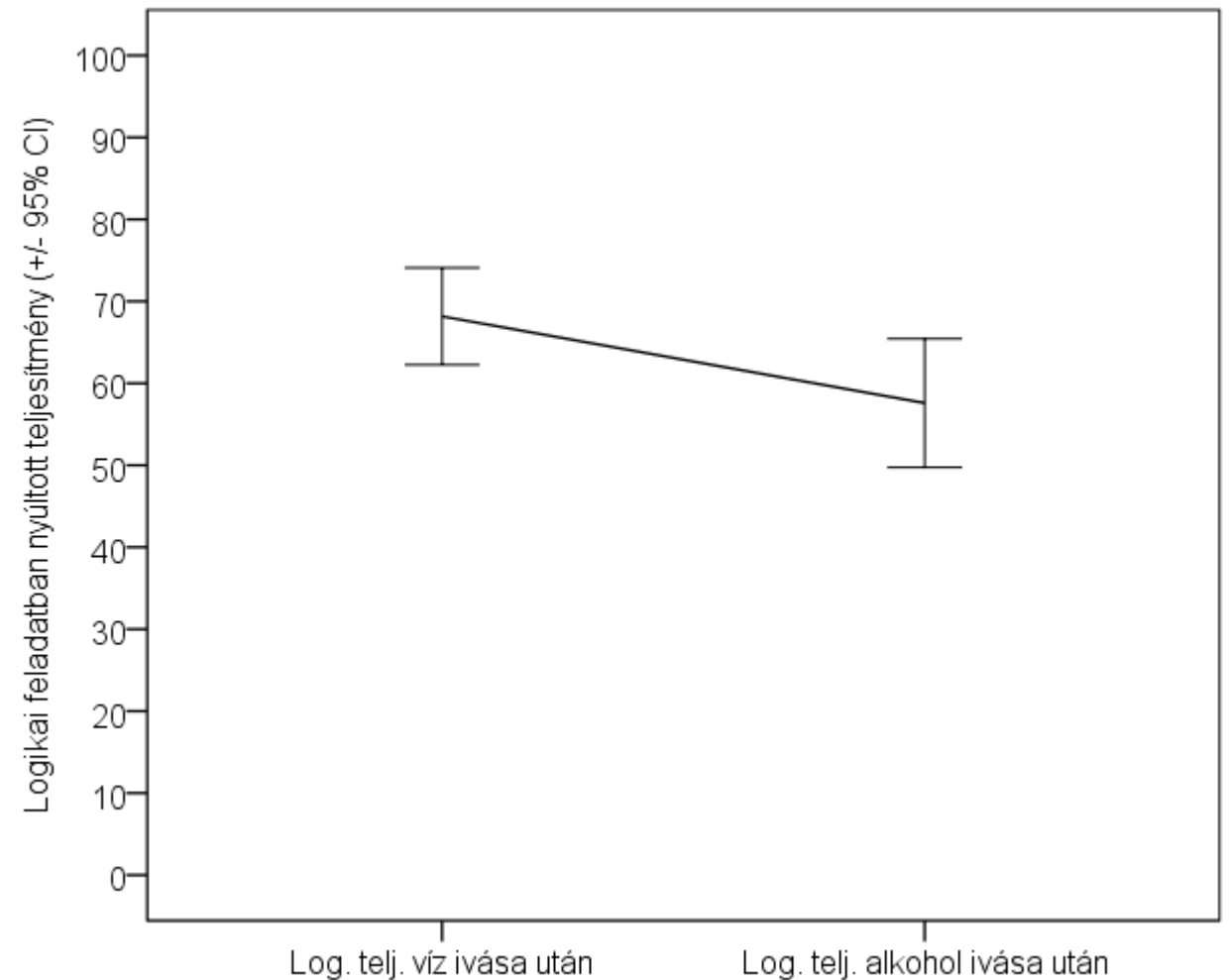


Megjelenítendő
változók

Opciókon belül
jelöljük, hogy
kérünk hibasávot
is. Három közül
választhatunk,
konfidencia
intervallum,
standard error,
szórás



- A grafikon szerkesztését előző félévben tanultuk. Mire figyelj:
 - A hibaszáv adatait az Y-tengelyen zárójelben feltűntettem
 - A feliratokat magyarra írtam át, ahol kellett
 - A betűméret minimum 12pont, különben a dolgozatba átillesztve nem lesz olvasható
 - A skálázást 0-tól 100-igra állítottam, mivel ez a logikai teljesítmény teljes tartománya
 - Esztétika – eltűntettem a ronda háttérét, átszíneztem...



Eredmények publikálása

- Publikálás formája
 - $t(df) = [t \text{ érték}] \ p = [érték] \ ([oldaliság]) \ [effectsize]$
 - Mindig publikáljuk a minták átlagát és szórását is!
 - Fontosabb eredményeket grafikonon is mutassuk meg.
- Szövegesen:
 - Szignifikáns, nagy különbséget találtunk a személyek logikai feladatban nyújtott teljesítményében akkor, ha vizet ($M = 68.16 \ SD = 19.218$), vagy, ha alkoholt ($M = 57.58 \ SD = 25.541$) fogyasztottak a feladat elvégzése előtt, $t(42) = 6.659 \ p < .001$ (1-tailed) $r = 0.717$. Alkohol fogyasztását követően a személyek teljesítménye rosszabb volt. Az eredmény megfelel hipotézisünknek, miszerint bla-bla-bla...

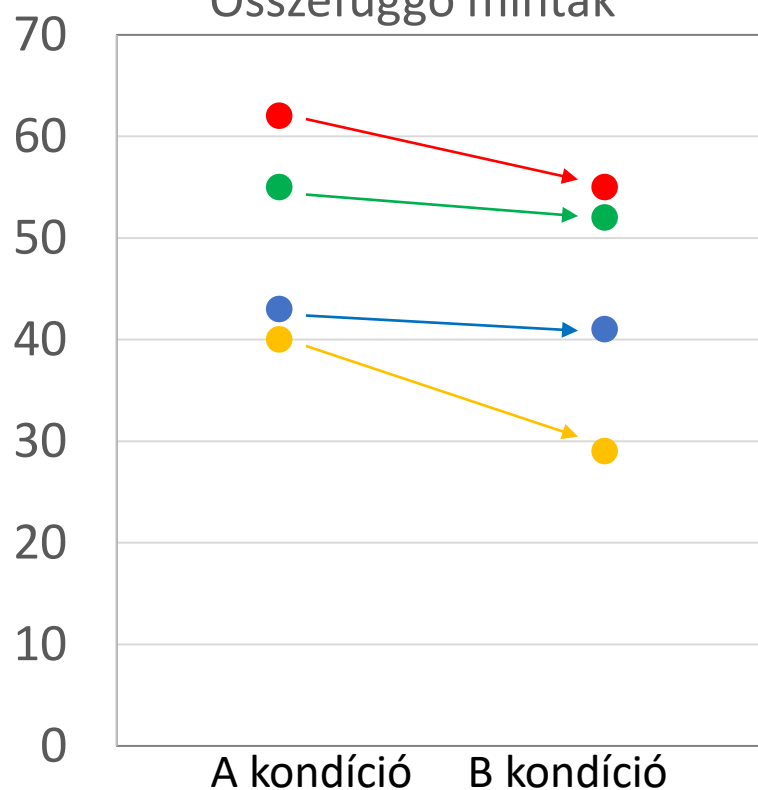


Rövid kitérő az összefüggő és független mintás t-próba érzékenységére

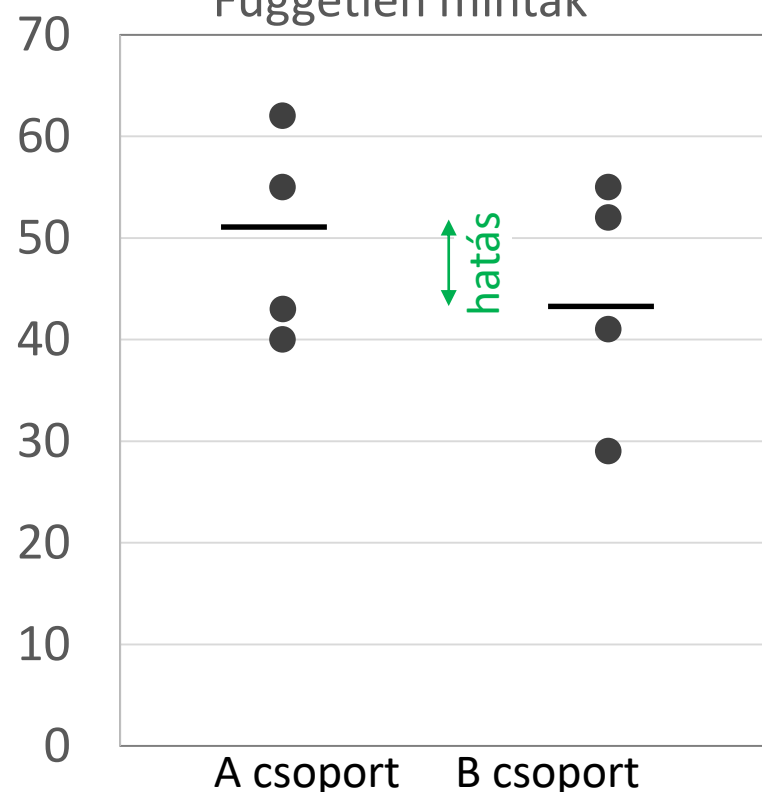
$$t = \frac{\bar{D}}{\frac{S_D}{\sqrt{N}}}$$

$$t = \frac{(\overline{M_{v\acute{ı}z}} - \overline{M_{alk}})}{\text{becsült közös } SE}$$

Összefüggő minták



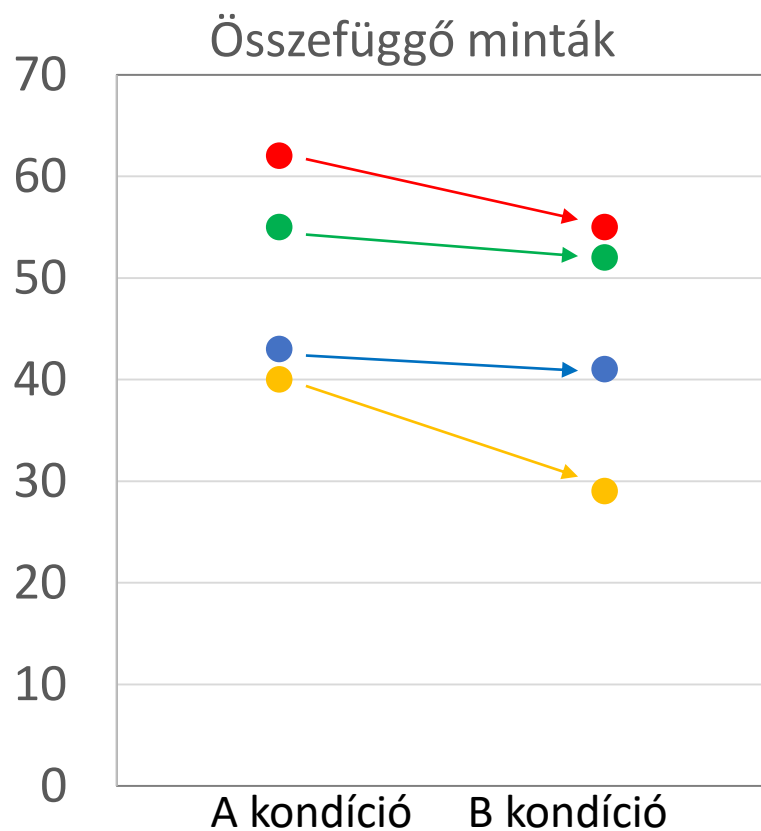
Független minták



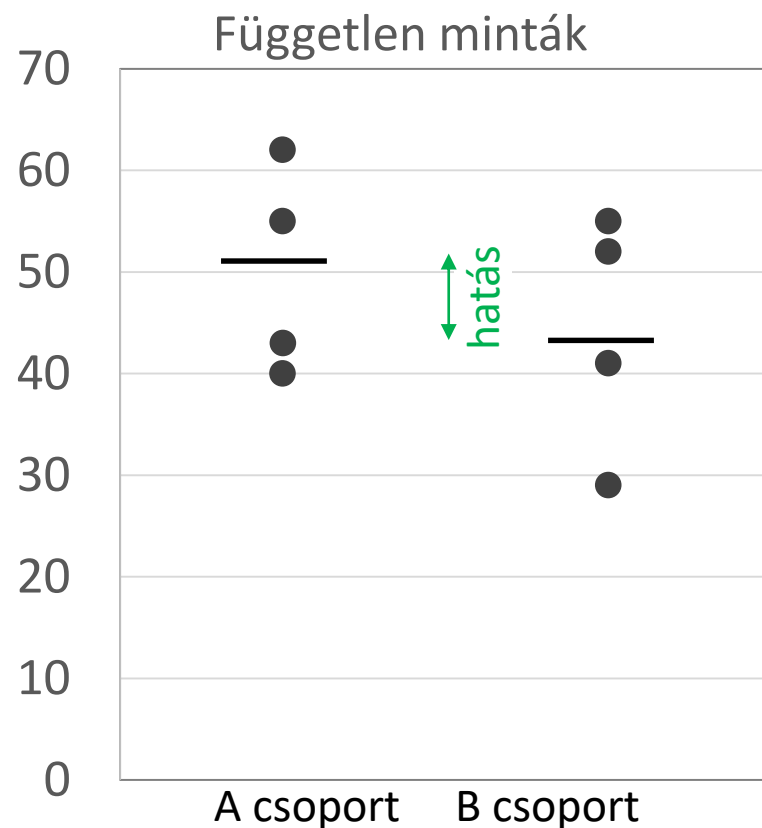
- Látható, hogy a független és összefüggő mintás vizsgálatban ha nem is azonos, de - hasonló mértékű volt a különbség a két minta között (az alkohol kb. 11 ponttal csökkentette a teljesítményt). Sőt a két minta szórása is hasonló volt (a vizes minta szórása 19, az alkoholos mintáé 25 körül volt).
- Az összefüggő próba eredménye mégis sokkal szignifikánsabb lett, nagyobb volt a t-érték, sőt még az effect-size is. Miért van ez?



$$t = \frac{\bar{D}}{\frac{S_D}{\sqrt{N}}}$$



$$t = \frac{(\overline{M_{v\acute{ı}z}} - \overline{M_{alk}})}{\text{becsült közös SE}}$$



- Azért van, mert a hatás azonos mindkét esetben (az mindegy, hogy a különbségeket átlagolod, vagy az átlagok különbségét veszed, az érték ugyanaz lesz)
- A zaj azonban az összefüggő mintás próba esetén kisebb, hiszen ott csak a személyeken belüli megváltozásban lévő változatossággal kell számolni (magyarul azzal, hogy van, akinél kisebb a csökkenés, van, akinél nagyobb). Míg a független mintás elemzésben a személyek közötti különbség is a zaj része (magyarul az, hogy vannak mintán belül is jobban, és kevésbé jól teljesítők).
- **Tehát a hatás mértéke azonos, a zaj viszont nagyobb a független mintás próbában, ezért ez a próba kevésbé érzékeny, nehezebb vele a különbség kimutatása.**



- Elméleti szempontból is érthető ez.

Független mintás	Összetartozó mintás
Hatás Mind a két elrendezésben azonos, adott, hogy az alkohol mennyivel csökkenti a teljesítményt.	
Zaj A kísérleti elrendezéstől függ, a zaj mekkora részével kell számolnunk. Ha nem lenne kísérleti manipuláció (vagy az alkoholnak nem lenne hatása), a két csoport teljesítménye között akkor is lenne valamennyi különbség, mert különböző személyek vannak a két csoportban (különböző IQ, képzettség, fáradtság, stb.)	 Ha nem lenne kísérleti manipuláció (vagy az alkoholnak nem lenne hatása), azt várjuk, hogy az első és második alkalommal nagyon hasonlóan teljesítenek az emberek – hisz ugyanazokról az emberekről van szó. Persze egy kis variancia itt is lenne, hisz más időpontban mérünk a két alkalommal (lehetnek fáradtabbak, más hangulatban, stb.)

Egymintás t-próba

Egy diasor nem bír el három poént Mr. T-vel.
- A hallgatók

Másnéven Student-próba

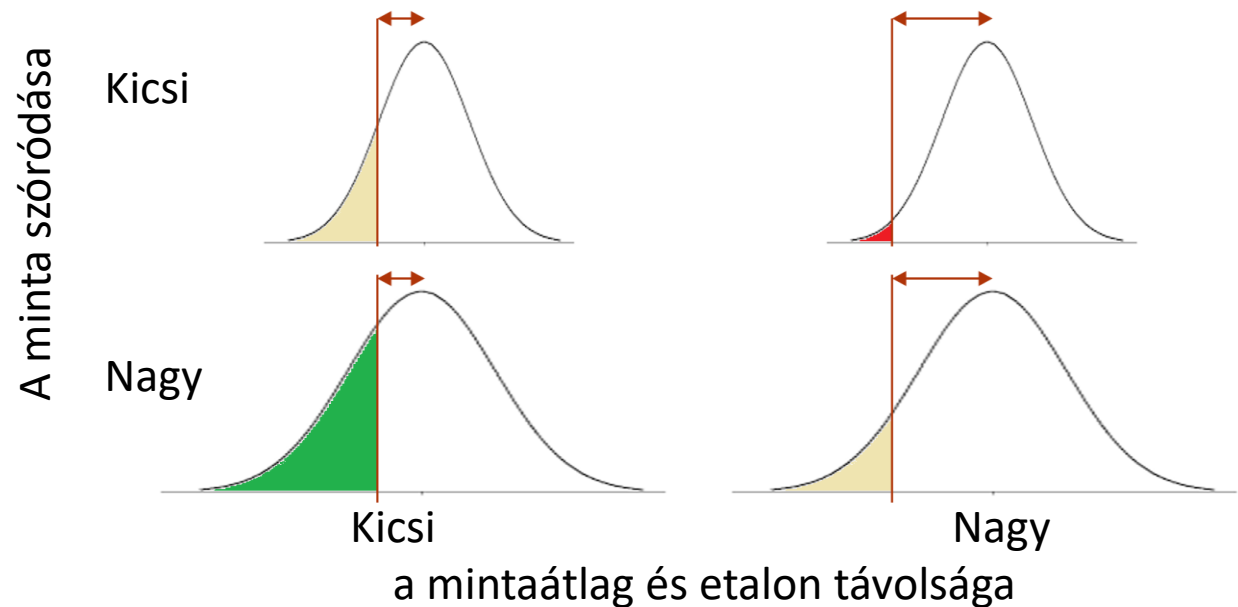
- Egy kis történelem:
 - William Sealy Gosset (Student az álneve) a Guinness sörgyár kémikusa. Az kellett ellenőriznie, hogy a sörök nem térnek el szignifikánsan egy előre meghatározott értéktől (fél liter).
 - Megalkotta a t-próbát, majd Student álnéven publikálta egy matematikai lapban.
- Mire használjuk?
 - Egy szakcikkben azt olvastuk, hogy a magyarok átlagos logikai teljesítménye azon a teszten, amit mi is használtunk, 65 pont. Szeretnénk tudni, hogy a mi mintánk (józanul) a magyar átlagnak megfelelően teljesített-e?
 - Hogyan állapítsuk meg, hogy a mintánk eltér-e egy előre meghatározott értéktől, az úgy nevezett etalontól?
 - Mekkora különbséget várunk a nullhipotézis teljesülése esetén?
 - 0-át, azaz azt, hogy a minta átlaga és az etalon között nem lesz különbség.

- A hatás:
 - A mi mintánk átlaga 68.16 pont, az etalon 65, tehát a mintaátlag és etalon távolsága 3.16 pont. Ez a hatás.
- A zaj:
 - Minél nagyobb a minta szóródása, annál jobban „elnyeli” az etalont is.

- Ebből a próba egyenlete:

$$t = \frac{\overline{M} - \mu}{SE}$$

mintaátlag $\overline{M}$ etalon μ
 SE $\swarrow$
 a minta standard errorja

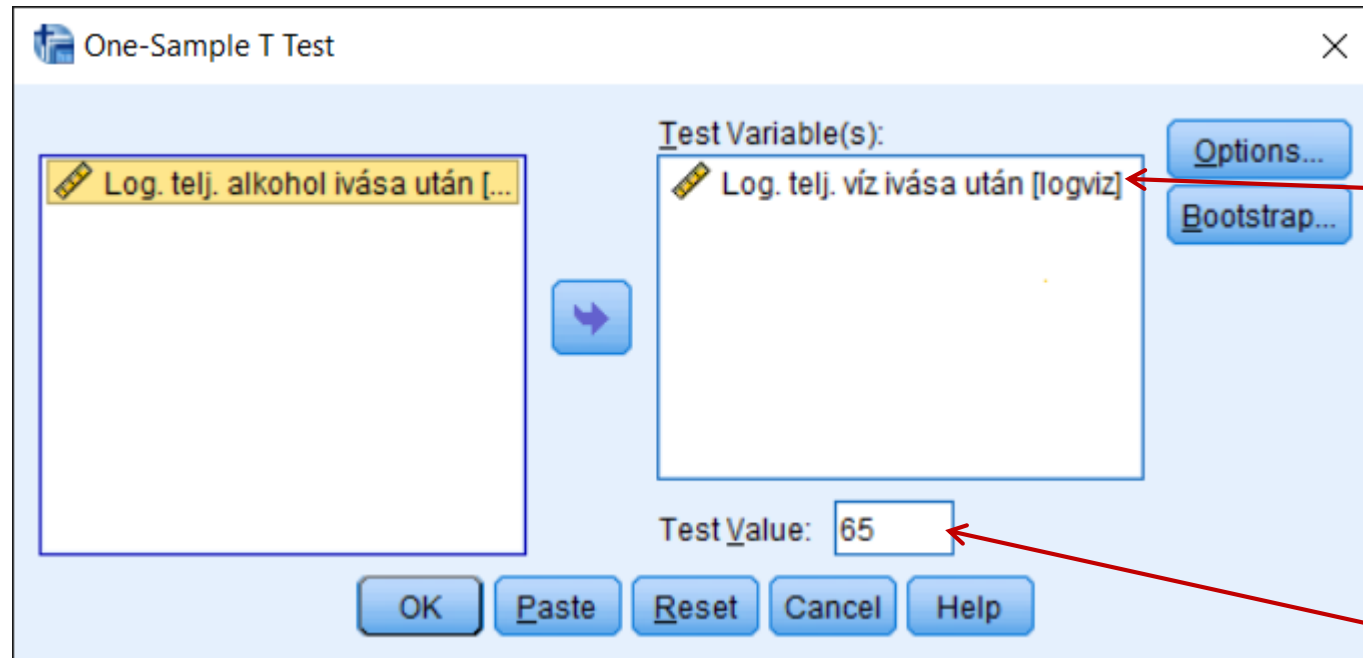


Az egymintás t-próba feltételei

- Feltételek – (a t-próba parametrikus teszt)
 - A kitöltők függetlenek egymástól
 - A minta legalább skála típusú
 - A minta normál eloszlású
 - (Szóráshomogenitás tesztelésre nincs mód, hiszen egyetlen minta van csak)

Egymintás t-próba az SPSS-ben

- Analyze / Compare means / One-sample t-test



Tesztelendő minta

Az etalon, amihez hasonlítunk

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Log. telj. víz ivása után	43	68,16	19,218	2,931

Leíró statisztikák

One-Sample Test

Test Value = 65

Etalon, amihez hasonlítunk

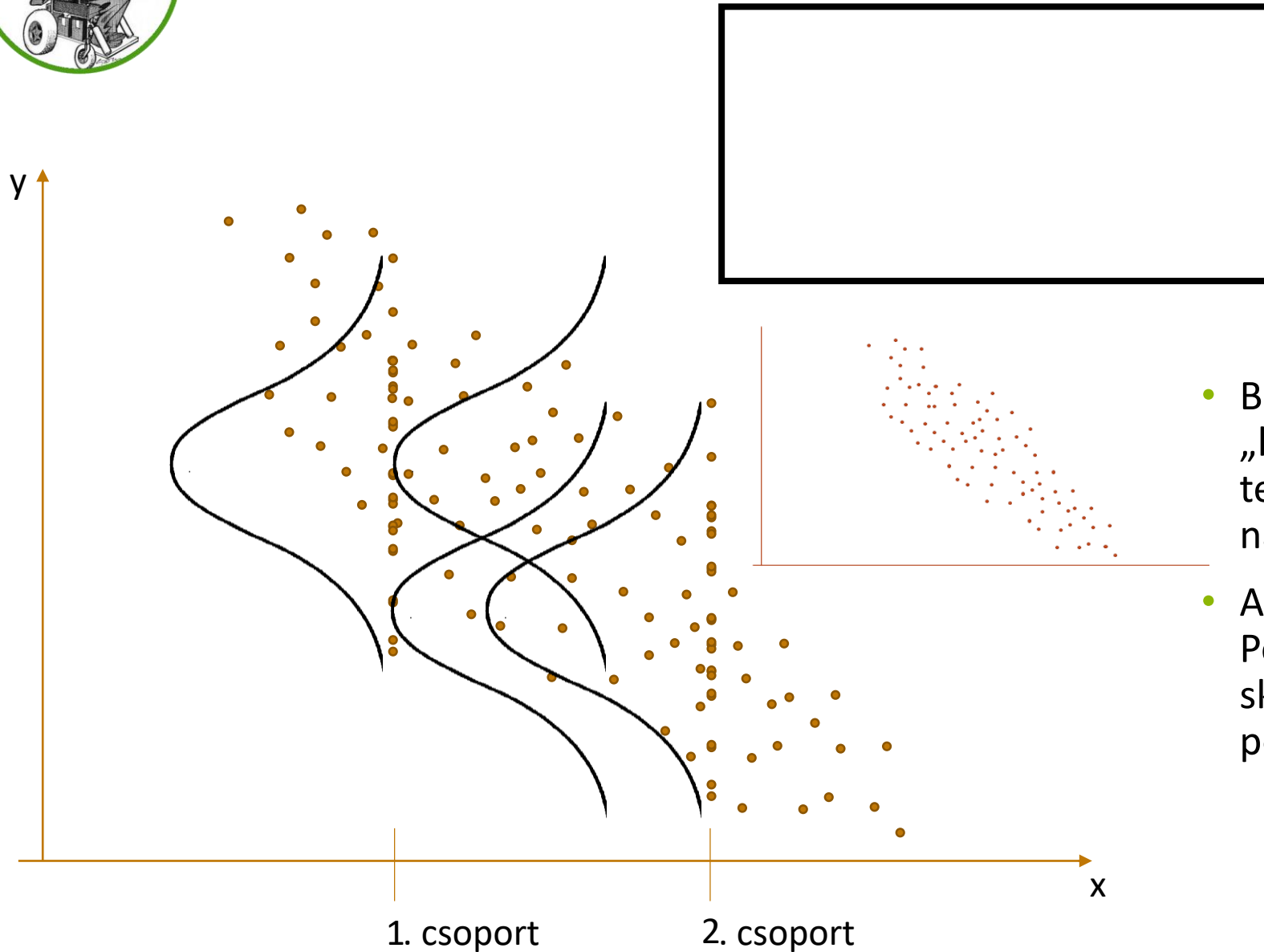
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Log. telj. víz ivása után	1,079	42	,287	3,163	-2,75	9,08

- A vizsgálatban használt minta logikai feladatban nyújtott teljesítménye nem tér el szignifikánsan a szakirodalom alapján magyar populációban elvárt, 65-ös értéktől. $t(42) = 1.079$ $p = .287$ $r = .164$

(effect-size, ha nem ismerős, előző diákon ott van, hogyan kell számolni)



Rövid kitérő a kapcsolat és különbség vizsgálatokra



- Bár a „kapcsolat vizsgálatok” és a „különbség vizsgálatok” a tananyagban teljesen elkülönülnek, gyakorlatban nagyon is közel állnak egymáshoz.
- A „kapcsolat vizsgálat” mintapéldája a Pearson korreláció, mely végezhető egy skála és egy dichotóm változón is, ekkor point biserialnak nevezzük.