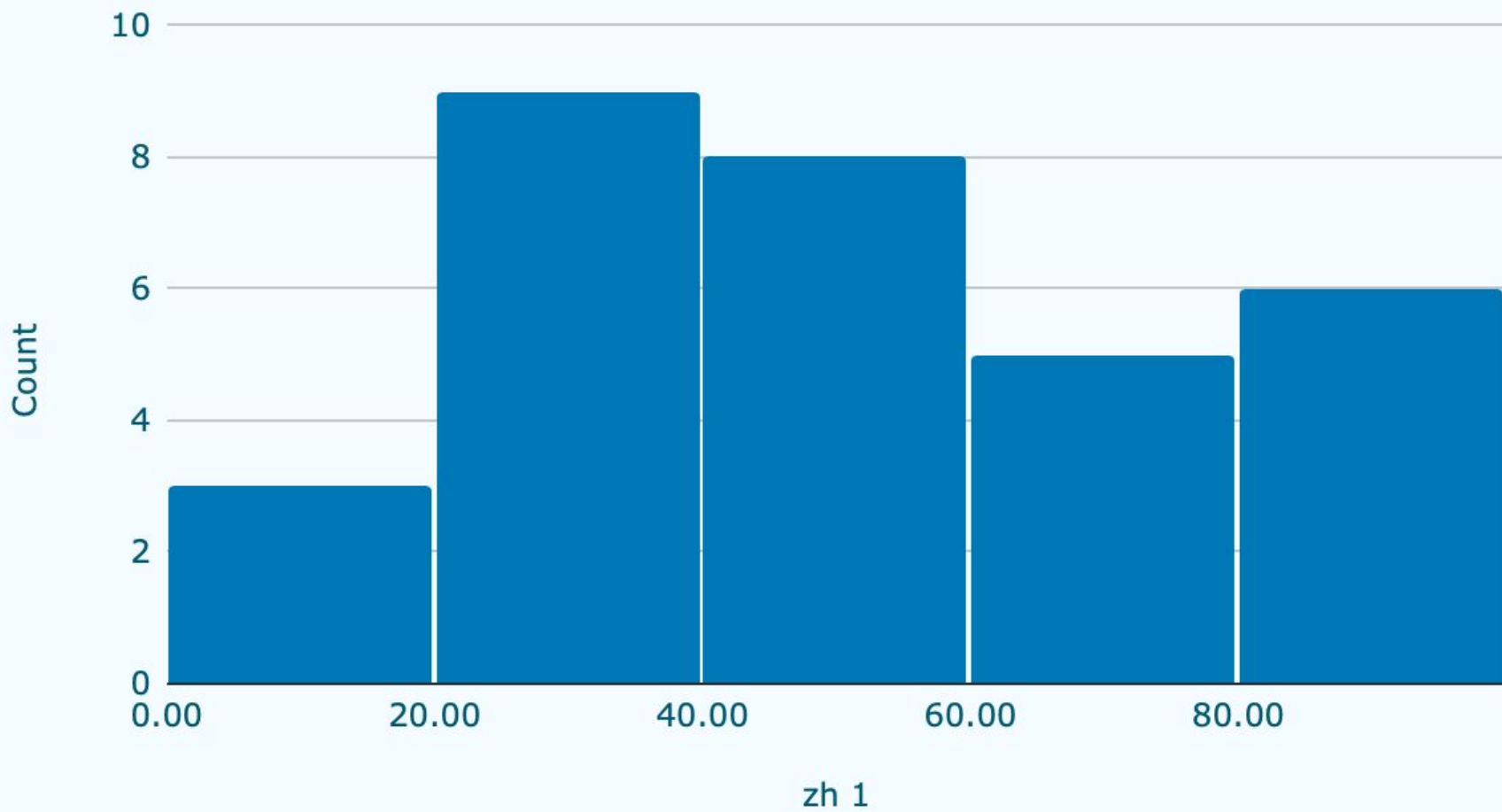


Distribution of zh 1



A t-próba

Adatfájlok

- SATpassage.csv
- selfhelpBook.csv
- Word_recall.csv
- LuckyCharm.csv

Átlagok összehasonlítása: populáció értékeinek becslése konfidencia intervallumokkal

- A kategorikus független változó szintjei által definiált minták átlagai közötti különbségből becsüljük meg a populációk átlagai közötti különbséget
- Módszer:
 - Minták átlagainak konfidencia intervallumainak az egybeesése: ha egybeesnek, nem feltételezhetünk különbséget a populációk között
 - Hatásméret konfidencia intervalluma: ha magában foglalja a nullát, nem feltételezhetünk különbséget a populációk között

Átlagok összehasonlítása: szignifikancia tesztelés

- A kategorikus független változó szintjei által definiált minták átlagai közötti különbség tesztjei
- Kutatási hipotézis: a független változó hatással van a függő változóra, azaz az átlagok között szignifikáns különbség van, azaz a minták különböző populációkból jönnek
- Null hipotézis: a független változó nincs hatással a függő változóra, azaz az átlagok között nincs általánosítható különbség, azaz a minták ugyanabból a populációból jönnek
- A teszt annak a valószínűségét mutatja ki, hogy az adott statisztikákat (átlagokat, szórásokat, stb) kapjuk akkor, ha a null hipotézis a helyes, azaz a null hipotézis valószínűségét

Ismét eloszlás

- Az átlagok összehasonlítására két tesztcsalád van:
 - Paraméteres tesztek: normál eloszlású intervallum vagy arányskála függő változóra
 - Nem-paraméteres tesztek: nem normál eloszlású intervallum vagy arányskála függő változóra és ordinális függő változóra

Paraméteres tesztek családja

- t-próbák:
 - egy független és egy függő változó
 - a független változónak maximum 2 szintje van
- ANOVA:
 - több független változó
 - vagy egy függő változó, aminek több mint 2 szintje van

William Gosset: Student



William Sealy Gosset (1876–1937)



A Guinness-nek dolgozott és Student néven publikált

A t-próbák fajtái

- Egymintás t-próba:
 - egyetlen mintának az átlagát hasonlítjuk egy referencia értékhez
 - pl. a véletlennél jobban teljesítettek-e a diákok egy feleletválasztós teszten
- Páros t-próba:
 - ismételt méréses dizájn: a kísérleti személyeket két különböző kondícióban mérjük le
 - a két kondíción elért pontszámok átlagait hasonlítjuk össze
- Független mintás t-próba:
 - független mintás dizájn: két különböző csoportot mérünk le a két kondícióban
 - a két csoport átlagát hasonlítjuk össze

A t-próbák logikája

Általános számítási elv:

$$t = \frac{\text{két összehasonlítandó érték különbsége}}{\text{Standard Hiba}}$$

$$SE = \frac{SD}{\sqrt{N}}$$

A szignifikancia

- A t értékeknek van egy eloszlása (a Student t -disztribúció)
 - az egyes t -értékek valószínűségét mutatja random módon kiválasztott mintákra (vagyis arra az esetre, ha a null hipotézis a helyes)
 - <http://www.distributome.org/V3/calc/index.html>
- Az eloszlást befolyásolja az elemszám (illetve a **szabadságfok**)
- A p értéket az eloszlásról olvassuk le

Szabadságfok (degrees of freedom, df):

A változtatható értékek száma abban az esetben, ha az átlag adott
 $df = N - 1$

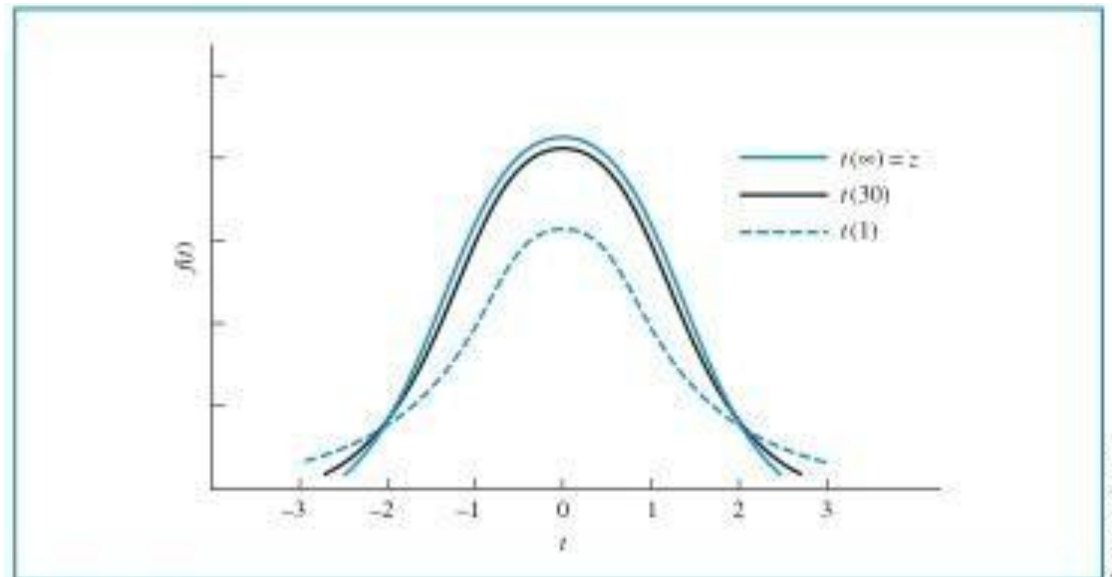


Figure 12.5

t distribution for 1, 30, and ∞ degrees of freedom.

Egymintás t-próba

Egymintás t-próba képlete

$$t = \frac{M - X}{SE}$$

M: a minta átlaga

X: az érték, amihez az átlagot hasonlítjuk

SE: a minta Standard Hibája

Példa: Katz et al 1990

- SAT (szövegértési teszt) úgy, hogy a szöveget nem lehet elolvasni (SATpassage.csv, 1: volt szöveg, 0: nem volt szöveg)
- 100 kérdés van, mindegyikre 5 válaszlehetőség közül kell választani
 - pontszám: helyes válaszok száma
- Hipotézis: a véletlennél jobban teljesítenek a diákok
- Null hipotézis: a véletlen szintjén teljesítenek a diákok

Jasp

Szűrjük ki azokat, akik olvasták a szöveget

T-Tests > One Sample T-Test

Tests: Student

Test value: 20

az átlag és a referenciavérték közötti különbség (M - X)

szabadságfok = N-1

One Sample T-Test ▼

	t	df	p	Mean Difference
Score	20.59	27	< .001	26.57

Note. Student's t-test.

Note. For the Student t-test, location parameter is given by mean difference d .

Note. For all tests, the alternative hypothesis specifies that the population mean is different from 20.

$$t = (M - X) / SE$$

Az érték, amihez az átlagot hasonlítjuk

Eredmények jelentése

Katz et al. (1990) were interested in whether students could guess the answers in tests of reading comprehension without reading the passage. They presented students with exam questions similar to those on the SAT, which required them to answer 100 five-choice multiple-choice questions about a passage. One of the groups ($N = 28$) was asked to answer the questions without reading the passage. A second group was allowed to read the passage, but they are not of interest to us here.

If participants perform purely at random, those in the NoPassage condition would be expected to get 20 items correct just by chance. A one-sample t test revealed that participants in the NoPassage condition scored significantly above chance ($M = 46.57$, $t(27) = 20.59$, $p < .001$). That is, it is possible to answer some of the questions correctly without even a single glance at the text to which they refer.

Eredmények jelentése

Katz et al. (1990) hipotézise szerint szövegértési teszteken a diákok a szöveg elolvasása nélkül is ki tudják találni a helyes válaszokat. Egy 100 kérdésből álló, feleletválasztós szövegértési tesztet végeztek diákokkal, ahol 5 lehetséges válasz közül kellett választani. A diákok egyik csoportjának ($N = 28$) a szöveg elolvasása nélkül kellett válaszolni a kérdésekre. Jelen jelentésben csak ezzel a csoporttal foglalkozunk.

Ha a diákok a szöveg elolvasása nélkül random módon válaszolnak a kérdésekre, a százból várhatóan 20 választ találnak el. Egy egymintás t-próba eredménye szerint a diákok szignifikánsan a véletlen szintje felett teljesítettek a szöveg elolvasása nélkül ($M = 46,57$, $t(27) = 20,59$, $p < .001$). Ez arra utal, hogy a szövegértési tesztek nem csupán szövegértést mérnek, hiszen bizonyos mértékben a szöveg elolvasása nélkül is ki lehet következtetni a helyes választ.

Független mintás t-próba

Független mintás t-próba képlete

$$t = \frac{(M_1 - M_2) - X}{SE_{DM}}$$

M_1 és M_2 : a két csoport átlagainak a különbsége

X : 0 (az érték, amihez hasonlítjuk az átlagok különbségét)

SE_{DM} : a két csoport átlagai közötti különbségnek a Standard hibája

Az átlagok különbségének Standard Hibája

- A két kondíciót két különböző csoport teljesíti. Mindkettőnek van egy átlaga.
- A két átlag között van egy különbség.
- Ennek a Standard Hibáját a két csoport szórásából (varianciájából) és elemszámából becsüljük meg

$$SE_{DM} = \sqrt{\left(\frac{SD_1^2}{N_1} + \frac{SD_2^2}{N_2} \right)}$$

A varianciák egyenlősége

- Mivel a t képletében a Standard Hibát a két csoport varianciájából becsüljük meg, igazítani kell a számításon, ha a két variancia nagyon különbözik egymástól.
- Ennek a tesztje (equality of variances): Levene teszt
 - Ha a teszt szignifikáns ($p < ,05$), Welch korrekciót kell használni

Szignifikancia

- A t-eloszlásból: annak az esélye, hogy a kapott t értéket kapjuk, ha a null hipotézis a helyes
- A t eloszlása függ a szabadságfoktól

$$df = N_1 - 1 + N_2 - 1$$

Hatásméret

Cohen's d

$$d = \frac{M_2 - M_1}{SD_{pooled}}$$

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

SD: a két csoport összevont szórása
Az elemszámtól független
hatásméret.

Bármilyen értéket felvehet. $d = 0,8$
többnyire már erős hatásnak
számít.

r^2 (determinációs együttható)

$$r^2 = \frac{t^2}{t^2 + df}$$

Az elemszám befolyásolja.

Könnyebb értelmezni, mert 0 és 1
között van az értéke. Azt mutatja,
hogy a független változó a függő
változó varianciájának
hányadrészét jósolja be.

Példa: ismét SATpassage.csv

Most a két csoportot hasonlítjuk össze: azoknak a diákoknak a pontjait, akik olvasták a szöveget, azoknak a pontjaival, akik nem olvasták a szöveget.

Hipotézis: Bár a véletlennél jobban teljesítenek a diákok a szöveg elolvasása nélkül, de azért rosszabbul, mint a szöveg elolvasásával.

Jasp

T-Tests > Independent Samples T-Test
Tests: Student vagy Welch
Assumption Checks: Equality of variances

Levene nem szignifikáns,
tehát nem kell Welch
korrekció

Test of Equality of Variances (Levene's) ▼

	F	df	p
Score	2.29	1	0.14

Independent Samples T-Test

	t	df	p	Mean Difference	SE Difference	Cohen's d
Score	-9.21	43.00	< .001	-24.49	2.66	-2.83

Note. Student's t-test.

$$t = (M_1 - M_2) / SE_{DM}$$

A két átlag
különbsége
($M_1 - M_2$) (Location
parameter)

Az átlagok különbségének
Standard Hibája (SE_{DM})

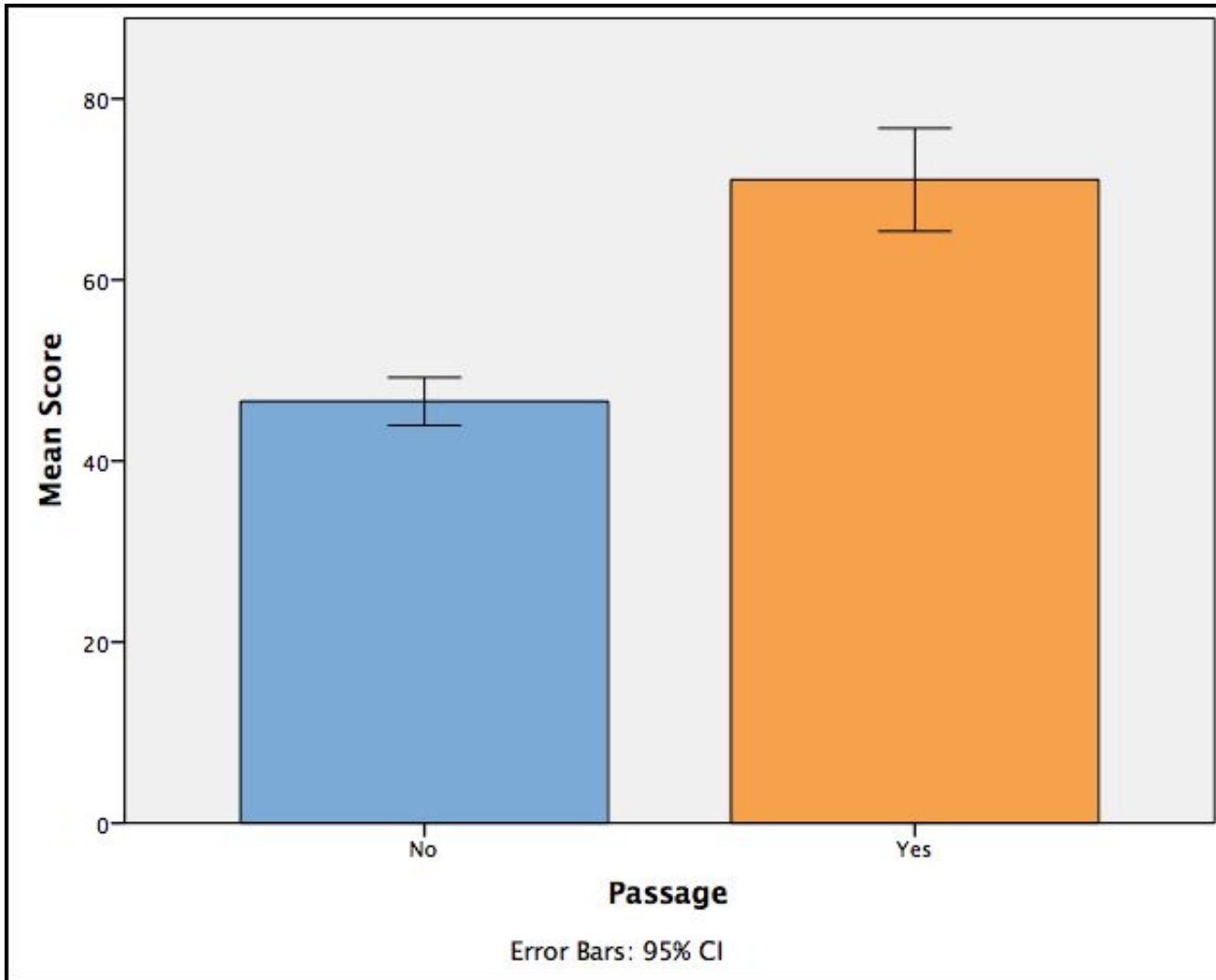


Figure 1. Mean score of the group who did not read the passage and the group who did read the passage.

Az eredmények jelentése

[...] The SAT performance of the group of students who did not read the passage ($N = 28$) was compared to the performance of a group of students who did read the passage ($N = 17$). The results are displayed in Figure 1. An independent samples t test revealed that the students who read the passage had significantly higher scores ($M = 71.06$, $SD = 11.06$) than the No-passage group ($M = 46.57$, $SD = 6.83$), ($t(43) = -9.21$, $p < .001$). While people may show better than chance performance without reading the passage, their scores will not approximate the scores of those who read the passage. This effect was very large (Cohen's $d = 2.83$).

Az eredmények jelentése

[...] A tesztet a szövegeket nélkül kitöltő csoport (N = 28) teljesítményét összehasonlítottuk a szövegeket elolvasó csoport (N = 17) teljesítményével. Az eredményeket az 1. ábra szemlélteti. A független mintás t próba szerint that a szöveges csoport szignifikánsan magasabb pontszámot (M = 71,06; SD = 11,06) ért el, mint a szöveg nélküli csoport (M = 46,57; SD = 6,83), ($t(43) = -9,21; p < .001$). Tehát bár a véletlennél jobban teljesítenek a diákok a szöveg elolvasása nélkül, de lényegesen rosszabbul, mint a szöveg elolvasásával. Ez a hatás nagyon erősnek bizonyult (Cohen's $d = 2,83$).

Páros t-próba

Páros t-próba képlete

$$t = \frac{M_D - X}{SE_D}$$

M_D : a két kondícióon elért pontok különbségeinek átlaga (D: difference)

X: 0, az érték, amihez hasonlítjuk a különbségek átlagát

SE_D : a két kondícióon elért pontok különbségeinek a Standard hibája

A különbségek Standard Hibája

- Minden kísérleti személynek két pontszáma van: az egyik kondícióban és a másik kondícióban elért pontszám
- A különbségek: minden egyénre, a két pontszáma közötti különbség
- Ezeknek a különbségeknek van átlaga, szórása és Standard Hibája
- N: a párok száma

$$SD_D = \sqrt{\frac{\Sigma(D - M_D)^2}{N - 1}} = \sqrt{\frac{SS_D}{N - 1}}$$

$$SE_D = \frac{SD_D}{\sqrt{N}}$$

Hatásméret

Cohen's d

$$d = \frac{M_2 - M_1}{SD_{\text{of differencescores}}} = \frac{t}{\sqrt{N}}$$

SD: a különbségek szórása

Az elemszámtól független hatásméret.

Bármilyen értéket felvehet. $d = 0,8$ többnyire már erős hatásnak számít.

r^2 (determinációs együttható)

$$r^2 = \frac{t^2}{t^2 + df}$$

Az elemszám befolyásolja.

Könnyebb értelmezni, mert 0 és 1 között van az értéke. Azt mutatja, hogy a független változó a függő változó variáciájának hányadrészét jósolja be.

Példa: selfhelpBook.csv

Önsegítő könyvek hatékonyak-e a párkapcsolatok rendezésében:

Lemérik párok párkapcsolati boldogságát két könyv elolvasása után: Men are from Mars, Women are from Venus és statisztika tankönyv. 6 hónap telik el a két könyv elolvasása között, a ksz-ek fele az egyik könyvvel kezdi, a másik fele a másik könyvvel kezdi.

A pontszám a könyvek elolvasása utáni párkapcsolati boldogságot mutatja.

Jasp

T-Tests > Paired Samples T-Test
Tests: Student

A különbségek átlaga (M_D)
(Location Parameter in JASP)

$$df = N - 1$$

Paired Samples T-Test

	t	df	p	Mean Difference	SE Difference	Cohen's d
women - statbook	2.71	499	0.01	1.53	0.56	0.12

Note. Student's t-test.

$$t = (M_D - 0) / SE_D$$

A különbségek Standard
Hibája (SE_D)

Eredmények jelentése

To test the hypothesis that self-help books are useful, the relationship happiness of 500 couples was measured after reading *Men are from Mars, Women are from Venus* and after reading a statistics book. The two books were read six months apart. Half of participants read the self-help book first while the other half read the statistics book first. On average, reported relationship happiness after reading the self-help book ($M = 20.2$, $SD = 9.98$) was significantly higher than in the control condition ($M = 18.49$, $SD = 8.99$), $t(499) = 2.71$, $p = .007$. However, the small effect size estimate ($d = 0.12$) indicates that this difference was not of practical significance.

Eredmények jelentése

A kutatás az önsegítő könyvek hatását vizsgálja. Egy kérdőívvel 500 pár kapcsolatának a minőségét mérték a *Men are from Mars, Women are from Venus* önsegítő könyv elolvasása után, és egy statisztika könyv elolvasása után. A két könyv elolvasása között 6 hónap telt el. A párok egyik fele az önsegítő könyvet olvasta először, míg a másik fele a statisztika könyvet. Az eredmények szerint a kapcsolatok minősége szignifikánsan jobb volt ($t(499) = 2,71$; $p = ,007$) az önsegítő könyv elolvasása után ($M = 20,2$; $SD = 9,98$), mint a kontroll kondícióban ($M = 18,49$; $SD = 8,99$). A gyenge hatásméret ($d = 0,12$) azonban arra utal, hogy a különbségnek nincs gyakorlati jelentősége.

Normalitás problémák

- Ha az eloszlás egy kicsit eltér a normáltól, mert van egy-két kiugró értékünk:
- Konfidencia Intervallumokat érdemes nézni
 - Location parameter konfidencia intervalluma: a tesztnek megfelelő értéket mutatja (a különbségek átlagát illetve az átlagok különbségét) - ne legyen a 0 érték az intervallumban
 - Cohen d hatásméret konfidencia intervallumát: ebben se legyen 0