

Minták eltérésének statisztikai elemzése

BBNPS02000

Készítette: Soltész-Várhelyi Klára

Nemparametrikus próbák
kettőnél több minta összehasonlítására

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

Több független mintát összehasonlító nemparametrikus próbák

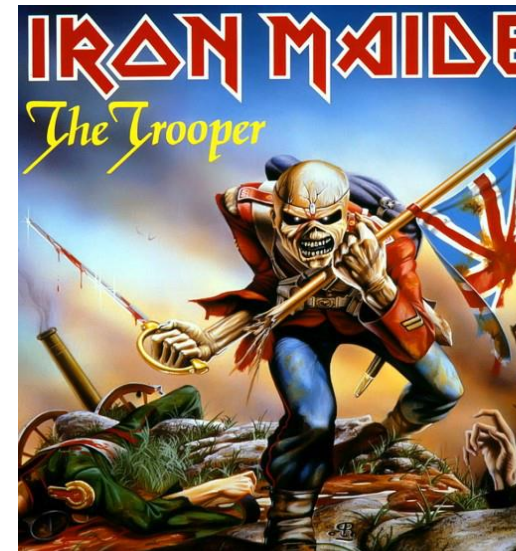
Független mintás ANOVA helyett

- ANOVA helyett használandó nem-parametrikus próba, ha
 - sérül a **szóráshomogenitás**:
 - ha csak szóráshomogenitás sérül, akkor használható a **Welch**, ami be van építve az ANOVÁ-ba
 - sérül a **normalitás**, vagy **ordinális a függő változó**:
 - **Kruskal-Wallis** – a függő változót rangsorolja, majd a rangokon végez egy ANOVÁ-hoz hasonló elemzést
 - **Median** – kis elemszám esetén – azt ellenőrzi hogy a mintából vett kis minták ugyanakkora része esik-e a minta mediánja alá és fölé minden csoportban
 - **dichotóm a függő változó**:
 - **Logisztikus regressziót** lehet végezni, de nem ez nem része a BA-s anyagnak

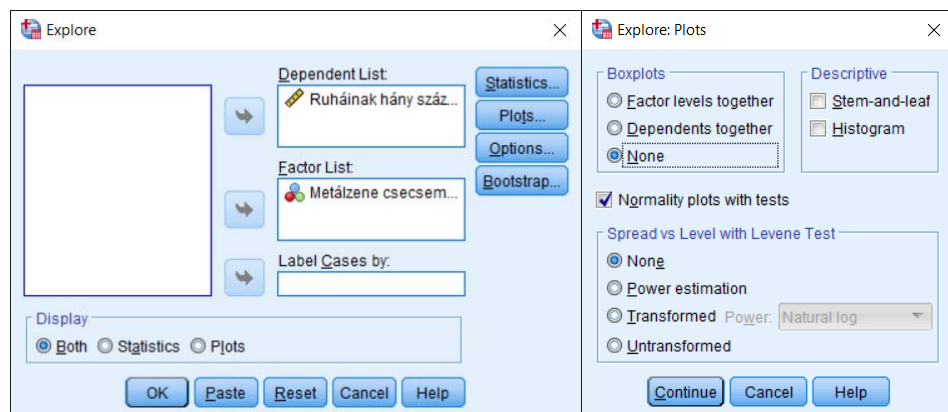
- Welch teszt
 - Lásd a One-Way ANOVA ppt-ben
- Kruskal-Wallis és Median teszt
 - Független kitöltők
 - Független minták
 - Legalább ordinális adatok

- `statgyakGY2_05_nemparaXminta_metal.sav`
- Azt vizsgáltuk, hogy hat-e a csecsemőkori metálzene hallgatás a tinédzserkori ruhaválasztási szokásokra.
- Ehhez öt csoportot alakítottunk ki:
 - Az első csoport egy éves koráig semmilyen metálzenét nem hallgatott
 - A másik négy csoport egy éves koráig naponta 10-15 perc metál zenét hallgatott a következő előadóktól: második csoport Black Sabbath, harmadik csoport Iron Maiden, negyedik csoport Slayer, ötödik csoport Six feet under.
- 15 éves korukban megvizsgáltuk, hogy a ruhatáruk mekkora hányada fekete.
- Azt feltételezzük, hogy a csoportok között különbség lesz.
 - A legkevesebb fekete ruhája azoknak lesz, akik nem hallgattak csecsemőként metál zenét.
 - A metált hallgató csoportoknál pedig azt várjuk, hogy minél durvább előadótól hallgattak zenét, annál nagyobb lesz a fekete ruhák aránya.

- Infóként azoknak, akik nem ismerik az előadókat (és azért is, hogy minden metalheadet magamra haragítsak):
 - Black Sabbath – igazi ősmetál. Dallamos, borongós stílusú. A maga idejében az egyik legkeményebb zenét játszó zenekar, ma már a Walt Disney mesékben is keményebb zene szól.
 - Iron Maiden – a heavy metál egyik legismertebb zenekara, lendületes, energikus, kiabálós.
 - Slayer - trash metál zenekar, nagyon gyors ritmussal, dupla lábdobbal, kiabálós, üvöltős zene.
 - Six feet under – death metál zenekar, erős torzítással, üvöltős, halálhörgős énekkel, halálról, vérengzésről szóló témákkal.
- Szóval „durvaság” szempontjából kijelenthető, hogy a sorrend:
Black Sabbath < Iron Maiden < Slayer < Six feet under



- Öt független csoport között keresünk különbséget egy folytonos változó mentén, első ötletünk a független mintás ANOVA, melynek 5 feltétele van:
 - Független kitöltők, független csoportok, skála típusú függő változó, függő változó csoportonkénti normál eloszlása és szóráshomogenitás.
 - Az első három feltétel értelemszerűen teljesül. Ellenőrizzük a normalitást (Analyze / Descriptive / Explore)!



		Tests of Normality					
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Ruháinak hány százaléka fekete	Metálzene csecsemőkorban nem hallgatott	,300	10	,011	,891	10	,175
	Black Sabbath	,238	10	,116	,914	10	,310
	Iron Maiden	,207	10	,200*	,918	10	,337
	Slayer	,268	10	,040	,774	10	,007
	Six feet under	,406	10	,000	,624	10	,000

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

- A S-W teszt két csoport esetében is szignifikáns (Slayer és SFU). Ha ellenőrizzük ezeknél a csoportoknál a csúcossági és ferdeségi értékeket, láthatjuk, hogy a Z-tesztekre sem hivatkozhatunk, tehát a normalitás nem teljesül e két csoportnál, független mintás ANOVA nem használható az elemzéshez.

Kruskal-Wallis

- Mivel a normalitás a probléma, válasszuk helyette a Kruskal-Wallis tesztet!
- A tesztnek három feltétele van, független kitöltők, független csoportok és legalább ordinális mérési szintű függő változó, mely feltételek teljesülnek, a K-W teszt kikérhető.
- **Analyze / Nonparametric test / Legacy dialogs / K Independent Samples**

Függő változó

Független változó

Csoportosító változó
Megadhatjuk, melyik csoportok kerüljenek az elemzésbe

The image shows two screenshots from the SPSS software interface. The main screenshot is the 'Tests for Several Independent Samples' dialog box. It has a 'Test Variable List' on the right containing 'Ruháinak hány száz...' and a 'Grouping Variable' field containing 'zene(1 5)'. Below these is a 'Define Range...' button. At the bottom, under 'Test Type', the 'Kruskal-Wallis H' option is selected with a checked box. Red arrows point from labels to specific parts: 'Függő változó' points to the 'Test Variable List', 'Független változó' points to the 'Grouping Variable' field, and another arrow points from the 'Define Range...' button to a second, smaller screenshot. This second screenshot is the 'Several Independent Sample...' sub-dialog, showing 'Range for Grouping Variable' with 'Minimum' set to 1 and 'Maximum' set to 5, bracketed together. A red arrow points from the text 'Csoportosító változó' to this bracketed range.

Ranks

	Metálzene csecsemőkorban	N	Mean Rank
Ruháinak hány százaléka fekete	nem hallgatott	10	10,00
	Black Sabbath	10	14,05
	Iron Maiden	10	26,65
	Slayer	10	33,95
	Six feet under	10	42,85
	Total	50	

Test Statistics^{a,b}

	Ruháinak hány százaléka fekete
Chi-Square	35,478
df	4
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable:
Metálzene
csecsemőkorban

- Átlagos rangok az öt csoportban.
- Látható, hogy azoknál a legkisebb a rangátlag, akik nem hallgattak metált csecsemőkorukban, tehát sorban náluk a legkisebb a fekete ruhák aránya. A Legmagasabb átlagos ranggal pedig a SFU-t hallgató csoport rendelkezik, tehát náluk a legmagasabb a fekete ruhák aránya a szekrényben.
- A Kruskal-Wallis teszt táblázatában látjuk, hogy a csoportok között szignifikáns különbség van.
- Azonban ugyanúgy, ahogy az ANOVÁ-nál is, ebből csak azt tudjuk meg, hogy van különbség a csoportok között, de azt nem, hogy *melyik* csoportok különböznek egymástól.

Utóvizsgálat Kruskal-Wallis teszthez

- Nincs beépített parancs rá, helyette Mann-Whitney tesztek sorozatát végezzünk „kézzel számolt” Bonfferoni korrekcióval
 - 1. Elvégezzük a Mann-Whitney tesztek a minket érdeklő párosítások között
 - 2. A szignifikanciához használt .05 kritériumszintet elosztjuk annyival, ahány próbát végeztünk
 - 3. Csak azokat a teszteredményeket tekintjük szignifikánsnak, melyek szignifikanciaszintje a korrigált kritériumszint alatt van
- Hasonlítsunk össze minden csoportot a szomszédjával! Ez négy összehasonlítást jelent, azaz négy Mann-Whitney tesztet fogunk végezni. Mivel négy összehasonlítást végzünk, ezért a .05-ös szignifikancia szintet négygyel osztjuk, és így az új, Bonferroni-korrigált szignifikancia szintünk $\alpha_B = .0125$

Test Statistics ^a	
	Ruháinak hány százaléka fekete
Mann-Whitney U	35,000
Wilcoxon W	90,000
Z	-1,148
Asymp. Sig. (2-tailed)	,251
Exact Sig. [2*(1-tailed Sig.)]	,280 ^b

Semmi / Black Sabbath
 $p > .0125$ nem szig

Test Statistics ^a	
	Ruháinak hány százaléka fekete
Mann-Whitney U	8,000
Wilcoxon W	63,000
Z	-3,200
Asymp. Sig. (2-tailed)	,001
Exact Sig. [2*(1-tailed Sig.)]	,001 ^b

Black S. / Iron M.
 $p < .0125$ szig

Test Statistics ^a	
	Ruháinak hány százaléka fekete
Mann-Whitney U	21,000
Wilcoxon W	76,000
Z	-2,204
Asymp. Sig. (2-tailed)	,028
Exact Sig. [2*(1-tailed Sig.)]	,029 ^b

Iron M. / Slayer
 $p > .0125$ nem szig!!

Test Statistics ^a	
	Ruháinak hány százaléka fekete
Mann-Whitney U	26,500
Wilcoxon W	81,500
Z	-1,902
Asymp. Sig. (2-tailed)	,057
Exact Sig. [2*(1-tailed Sig.)]	,075 ^b

Slayer / SFU
 $p > .0125$ nem szig

- Kruskal-Wallis:
 - $\chi^2([df], N = [elemszám]) = [Chi_square_értéke] \ p = [p \text{ értéke}]$
 - $\chi^2(4, N = 50) = 35.478 \ p < .001$
- Utóvizsgálat:
 - Ahogy a Mann Whitney-t publikálnánk – effect-size-ot csak a post hoc tesztekhez számoljuk
- Szövegesen:
 - A csecsemőkori metálzene hallgatás szignifikánsan befolyásolja a pubertáskori fekete ruhák arányát $\chi^2(4, N = 50) = 35.478 \ p < .001$
 - A csoportok között előforduló különbségeket Mann-Whitney tesztek sorozatával elemeztük részletesebben. A többszörös tesztelés miatt az utóelemzés során Bonfferoni-korrekciót alkalmaztunk, az eredmények $\alpha_B=.0125$ szignifikancia szinten kerülnek értelmezésre.
 - Nem találtunk szignifikáns különbséget a kontrolcsoport és a Black Sabbath-ot hallgatók között ($U = 35 \ Z = -1.148 \ p=.251 \ r=.25$). Szignifikánsan több feketét hordanak az Iron Maident hallgatók, mint a Black Sabbath-ot hallgatók ($U = 8 \ Z = -3.200 \ p < .001 \ r = .71$). Nem találtunk szignifikáns különbséget az Iron Maident hallgatók és Slayert hallgatók között ($U = 21 \ Z = -2.204 \ p = .028 \ r = .49$) és a Slayert és Six Feet Undert hallgatók között ($U = 26.5 \ Z = -1.902 \ p = .057 \ r = .42$)

Több összefüggő mintát összehasonlító
nemparametrikus próbák

Összefüggő mintás ANOVA helyett

- ANOVA helyett használandó nem-parametrikus próba, ha
 - sérül a **szfericitás**:
 - ha csak szóráshomogenitás sérül, akkor használható az ANOVÁ-n belül beépített **Greenhous-Geisser** és **Hynh-Feldt** korrekció (lásd korábban)
 - sérül a **normalitás**, vagy **ordinális a függő változó**:
 - **Friedman-féle ANOVA** - személyenként a kondíciók közötti különbségeket rangsorolásából indul ki
 - **dichotóm a függő változó**:
 - **Cochran's Q** tesztet lehet végezni

- Greenhous-Geisser és Hynh-Feldt
 - Lásd a One-Way ANOVA ppt-ben
- Friedman-féle ANOVA teszt
 - Független kitöltők
 - Legalább ordinális adatok
- Cochran's Q teszt
 - Független kitöltők
 - Dichotóm függő változó
 - A nagy minta alapból nem feltétel, de a Cochran's Q teszt nem túl érzékeny próba, így kisméretű mintákon csak a viszonylag nagy hatások kimutatására alkalmas

- [statgyakGY2_05_nemparaXminta_shopping.sav](#)
- Az előző vizsgálat leginkább fekete csoportját a kutatók megpróbálják „megmenteni”, ezért elmennek velük több alkalommal shoppingolni. Azt feltételezzük, hogy a ruhavásárlás hatására csökkent a fekete ruhák aránya.
 - ruha0 – Fekete ruhák aránya kezdetben
 - ruha1 – Fekete ruhák aránya 1 hónap shoppingolást követően
 - ruha2 – Fekete ruhák aránya 2 hónap shoppingolást követően
 - ruha3 – Fekete ruhák aránya 3 hónap shoppingolást követően
- Arra is rákérdeztünk, hogy az adott hónapban volt-e koncerten. Azt feltételezzük, hogy a shoppingolás hatására csökken a koncertre járás is.
 - koncert0 – Volt-e koncerten a kutatás kezdő hónapjában
 - koncert1 – Volt-e koncerten 1 hónap shoppingolást követően
 - koncert2 – Volt-e koncerten 2 hónap shoppingolást követően



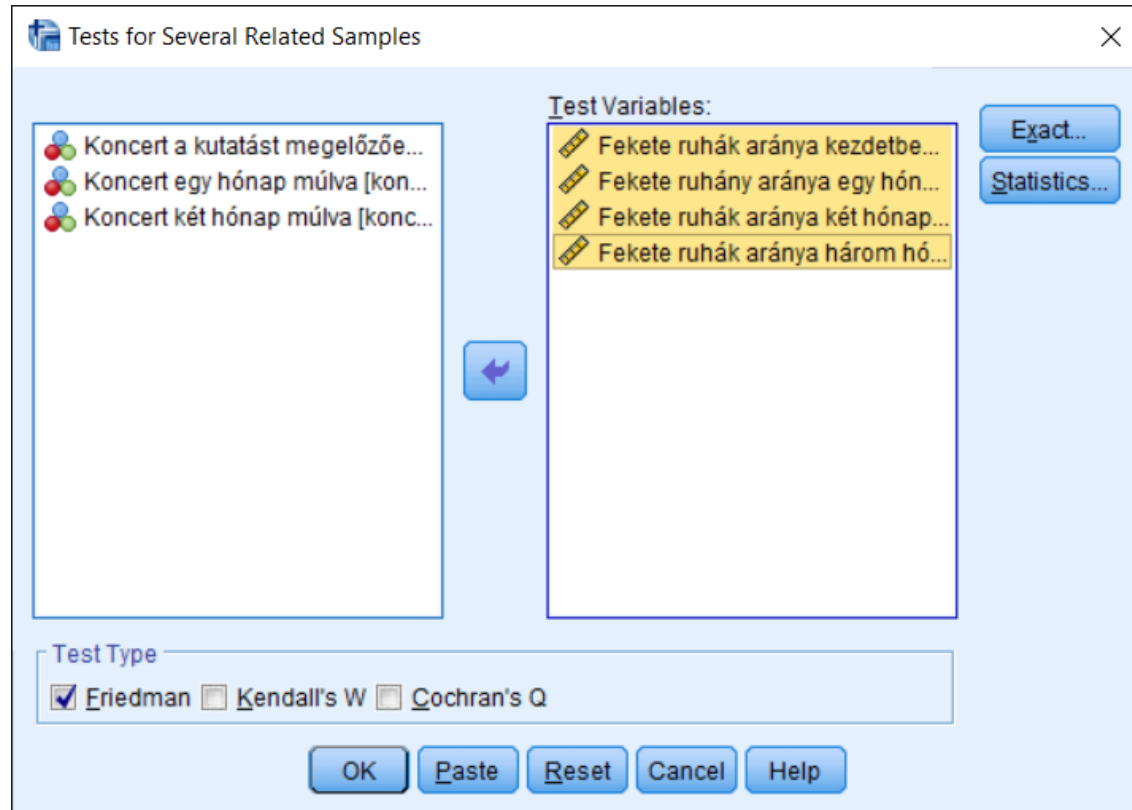
- *A shoppingolás hatására csökkent a fekete ruhák aránya.*
- Négy összefüggő minta között keresünk különbséget. Első ötletünk az ismételt méréses ANOVA, melynek négy feltétele van: független kitöltők, skála típusú változók, normál eloszlás, szfericitás. Az első két feltétel teljesül, a normalitással viszont meggyűlik a bajunk, A S-W tesztre nem hivatkozunk, és ha megnézzük a ferdeségi és csúcsossági mutatókat, láthatjuk, hogy azok alapján sem feltételezhető a normalitás.
- Mit válasszunk helyette?
 - A probléma a normalitás sérülésével volt, így válasszuk a Friedman-féle ANOVÁ-t.
 - A Friedman ANOVÁnak két feltétele van, a független kitöltők és a legalább ordinális adatok. Mindkét feltétel teljesül.

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Fekete ruhák aránya kezdetben	,406	10	,000	,624	10	,000
Fekete ruhák aránya egy hónap shoppingolást követően	,248	10	,083	,857	10	,071
Fekete ruhák aránya két hónap shoppingolást követően	,318	10	,005	,812	10	,020
Fekete ruhák aránya három hónap shoppingolást követően	,152	10	,200*	,972	10	,911

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

- Analyze / Nonparametric test / Legacy dialogs / K Related Samples*



Friedman – több összetartozó minta nemparametrikus tesztelése

Kendall's W – tipikusan raterek közötti egyezésre használjuk

Cochran's Q – dichotóm adatok elemzésére (McNemar több mintás megfelelője)

Ranks

Mean Rank

Fekete ruhák aránya kezdetben	4,00
Fekete ruhák aránya egy hónap shoppingolást követően	2,65
Fekete ruhák aránya két hónap shoppingolást követően	2,05
Fekete ruhák aránya három hónap shoppingolást követően	1,30

Test Statistics^a

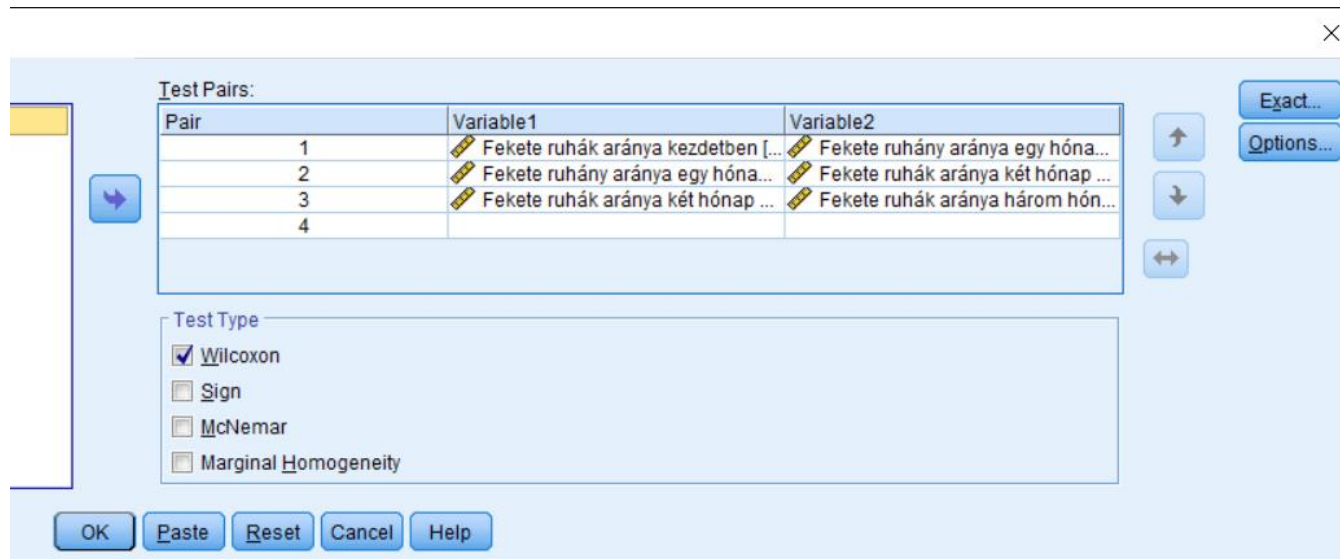
N	10
Chi-Square	23,727
df	3
Asymp. Sig.	,000

a. Friedman Test

- Átlagos rangok a négy csoportban
 - Látszik, hogy a legmagasabb rangok a kezdeti méréshez tartoznak, tehát kezdetben volt a legmagasabb a fekete ruhák mennyisége. A legalacsonyabb rangátlag a három hónapnyi shoppingolást követő mérésen van, ekkor volt a legalacsonyabb a fekete ruhák mennyisége.
- Teszt statisztika tábla:
 - A csoportok között szignifikáns különbséget találtunk.
 - Ugyanúgy, ahogy az ANOVÁ-nál is, most is csak azt tudjuk meg, hogy van különbség a csoportok között, de azt nem, hogy *melyik* csoportok különböznek egymástól

Utóvizsgálat Friedman teszthez

- Nincs beépített parancs, Wilcoxon tesztek sorozatát végezzük „kézzel számolt” Bonfferoni korrekcióval
 - 1. Elvégezzük a Wilcoxon tesztek a minket érdeklő párosítások között
 - 2. A szignifikanciához használt .05 kritériumszintet elosztjuk annyival, ahány próbát végeztünk
 - 3. Csak azokat a teszteredményeket tekintjük szignifikánsnak, melyek szignifikanciaszintje a korrigált kritériumszint alatt van
- Hasonlítsunk össze minden hónapot az előzővel! Ez három összehasonlítást jelent, azaz három Wilcoxon tesztet fogunk végezni. Mivel három összehasonlítást végzünk, ezért a .05-ös szignifikancia szintet hárommal osztjuk, és így az új, Bonferroni-korrigált szignifikancia szintünk $\alpha_B = .01666$



Egyszerre betehetjük az összes párosítást

Test Statistics^a

	Fekete ruhák aránya egy hónap shoppingolás t követően - Fekete ruhák aránya kezdetben	Fekete ruhák aránya két hónap shoppingolás t követően - Fekete ruhák aránya egy hónap shoppingolás t követően	Fekete ruhák aránya három hónap shoppingolás t követően - Fekete ruhák aránya két hónap shoppingolás t követően
Z	-2,829 ^b	-,892 ^b	-2,222 ^b
Asymp. Sig. (2-tailed)	,005	,372	,026

0 vs 1 hó
p<.0166
szig

1 vs 2 hó
p>.0166
nem szig

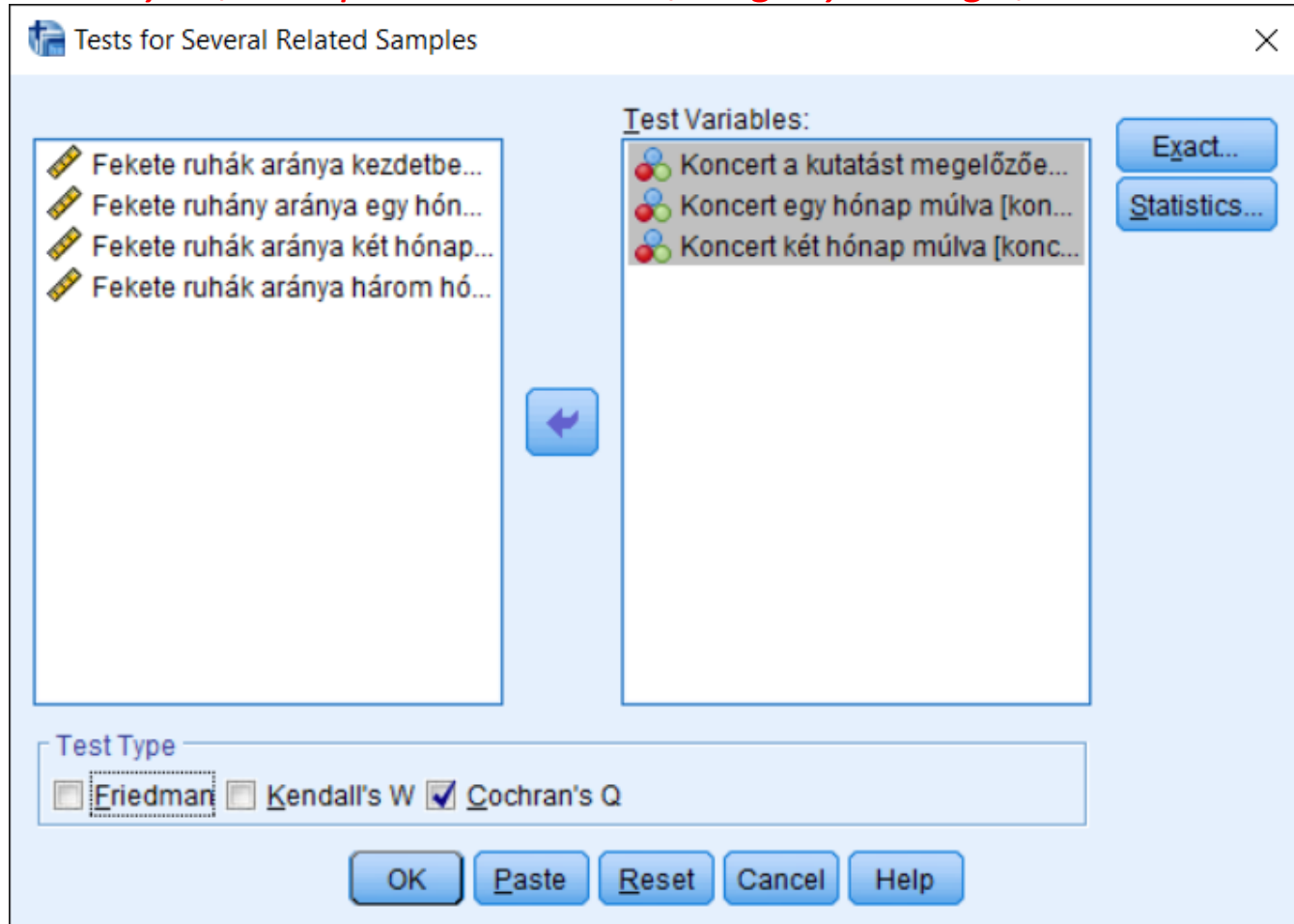
2 vs 3 hó
p>.0166
nem szig!!

- Friedman ANOVA:
 - χ^2 ([df], N = [elemszám] = [khi_érték] p = [p-érték]
 - χ^2 ([3], N = [10] = 23.727 p < .001
- Utóvizsgálat:
 - Ahogy a Wilcoxon tesztet publikálnánk – effect-size-ot csak a post hoc tesztekhez számoljuk
- Szövegesen:
 - Shoppingolással csökkenthető a metalheadek ruhatárában a fekete ruhák aránya. χ^2 (3, N = 10) = 23.727 p < .001
 - A mérések között előforduló különbségeket Wilcoxon tesztek sorozatával elemeztük részletesebben. A többszörös tesztelés miatt az utóelemzés során Bonfferoni-korrekciót alkalmaztunk, az eredmények α_B = .0166 szignifikancia szinten kerülnek értelmezésre.
 - Szignifikáns különbséget találtunk a kezdeti és egy hónapot követően mérhető fekete ruhák aránya között (Z = -2.829 p = .005 r = .895). Egy hónapot követően csökken a fekete ruhák aránya. Nem találtunk szignifikáns csökkenést a további hónapok során: egy illetve két hónap között Z = -0.892 p = .372 r = .282, nem szignifikáns; két illetve három hónap között Z = -2.222 p = .026 r = .702 szintén nem szignifikáns a korrigált szignifikancia szintet figyelembe véve.

Második hipotézis

- *A shoppingolás hatására csökkent a koncertre járás.*
- Három összefüggő minta között keresünk különbséget. Első ötletünk az ismételt méréses ANOVA, melynek négy feltétele van: független kitöltők, skála típusú változók, normál eloszlás, szfericitás. Könnyen látható, hogy a skála típusú változókra vonatkozó feltétel nem teljesül, hiszen a koncertre járás változói dichotómak.
- Mit válasszunk helyette?
 - A probléma az, hogy dichotóm változók vannak, ezért válasszuk a Cochran's Q tesztet.
 - A Cochran's Q tesztnek két feltétele van, a független kitöltők és a dichotóm változók. Mindkét feltétel teljesül.

- *Analyze / Nonparametric test / Legacy dialogs / K Related Samples*



Cochran's Q – dichotóm adatok elemzésére (McNemar több mintás megfelelője)

Frequencies

	Value	
	0	1
Koncert a kutatást megelőzően	2	8
Koncert egy hónap múlva	3	7
Koncert két hónap múlva	8	2

Test Statistics

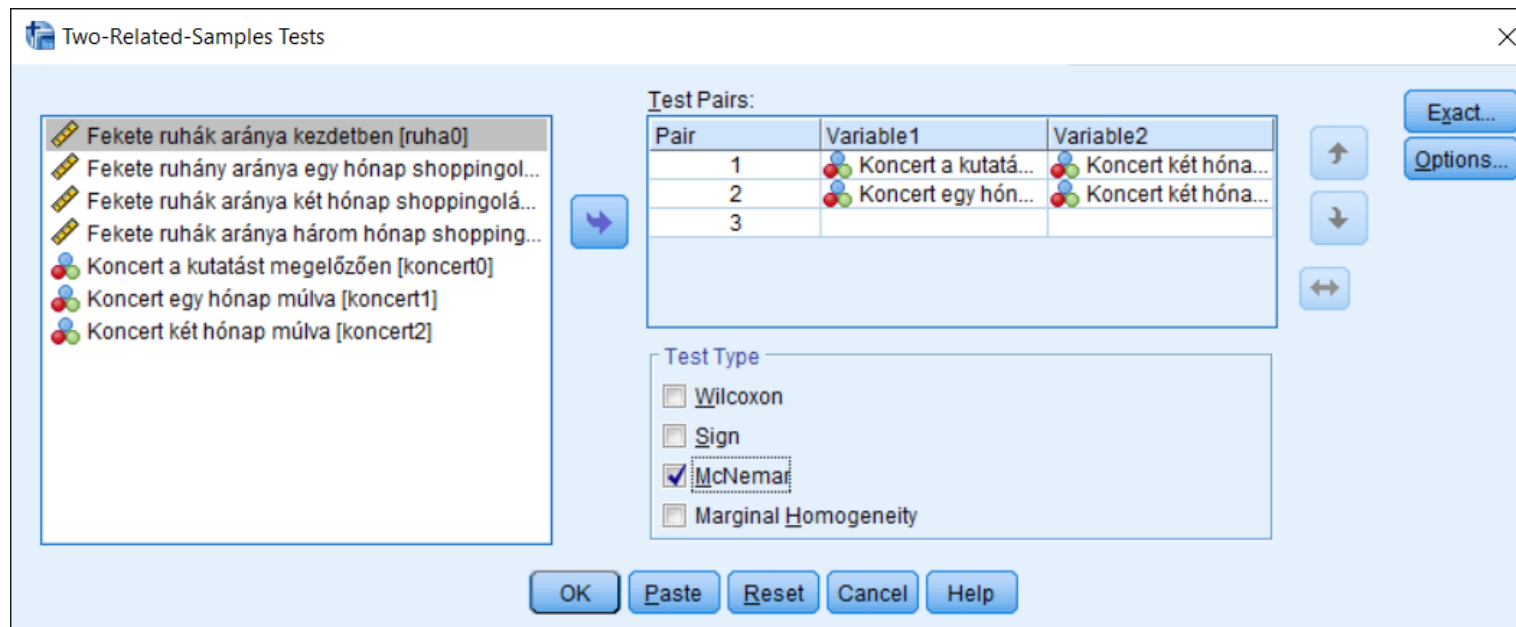
N	10
Cochran's Q	10,333 ^a
df	2
Asymp. Sig.	,006

a. 1 is treated as a success.

- Gyakoriság tábla
 - A kutatás kezdetén 2 személy nem járt el koncertre, 8 eljárt
 - Egy hónap múlva, 3 személy nem járt, 7 eljárt
 - Két hónap múlva 8 nem járt, 2 eljárt koncertre
- Teszt statisztika tábla:
 - A csoportok között szignifikáns különbséget találtunk.
 - Ugyanúgy, ahogy az ANOVÁ-nál is, most is csak azt tudjuk meg, hogy van különbség a csoportok között, de azt nem, hogy *melyik* csoportok különböznek egymástól

Utóvizsgálat Cochran's Q teszthez

- Nincs beépített parancs rá, helyette McNemar tesztek sorozatát végzünk „kézzel számolt” Bonfferoni korrekcióval.
- Az előző gyakoriságtábla alapján sejthetjük, hogy az utolsó hónapot érdemes összehasonlítani a többivel.
- Két összehasonlítást végzünk, ezért a Bonfferoni korrekció miatt a .05-ös szignifikancia szintet kettővel osztjuk, és így az új, Bonferroni-korrigált szignifikancia szintünk $\alpha_B = .025$



Test Statistics^a

	Koncert a kutatást megelőzően & Koncert két hónap múlva	Koncert egy hónap múlva & Koncert két hónap múlva
N	10	10
Exact Sig. (2-tailed)	,031 ^b	,063 ^b

0 vs 2 hó
p>.025
nem szig!!

1 vs 2 hó
p>.025
nem szig

Figyelj rá, most az utolsó mérést hasonlítottuk az előző kettővel

Utóvizsgálatban semmi sem szignifikáns???!?

- Hogy lehet, hogy az F próba (vagy valamely nem parametrikus párja) szignifikáns, de egyetlen utóvizsgálat eredménye sem az?
 - A posthoc vizsgálatok szigorúak (a Bonfferoni, amit most használtunk, méginkább)
 - ANOVA érzékenyebb, mint a páros összehasonlítások, mert az F statisztika a független változó függő változóra tett hatását vizsgálja, a posthoc párok közötti különbséget.
 - kicsi elemszám esetén ez fokozottan érvényes

