

„Az SZTE Kutatóegyetemi Kiválósági Központ tudásbázisának
kiszélesítése és hosszú távú szakmai fenntarthatóságának megalapozása
a kiváló tudományos utánpótlás biztosításával”



SZTE ETSZK Védőnői Munka és Módszertana Szakcsoporti Tudományos Diákköri Tanács Kutatásmódszertani alapismeretek kurzus

**Dr. Papp László
főiskolai docens**

Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Az adatfeldolgozás első lépése: az adatok kódolása

- A mérőmódszer tervezésénél érdemes végiggondolni a kódolási szisztémát, mert:
 - Ennek alapján történik az adatok rögzítése
 - (valamilyen adatbáziskezelő (MS Excel) vagy statisztikai szoftverbe (pl. SPSS, Sigmastat,...))
 - Szoftveres elemzés esetén szövegszerű kódolás nem értelmezhető
 - A jó kódolás segíti a megfelelő elemzési módszer megtalálását



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Az adatfeldolgozás második lépése: az adatok csoportosítása I.

Elvi megfontolások:

- Egy adat csak egy csoportban legyen elhelyezhető.
- Minden adat legyen elhelyezhető valamelyik csoportban.
 - Ha az „egyéb” kategóriába az adatok több, mint 5%-a esik, át kell dolgozni a csoportosítást.
 - A csoportok terjedelmét érdemes egyforma nagyságra meghatározni – a gyakorisági eloszlás felállításának feltétele
 - (a két szélső érték előfordul, hogy nagyobb is lehet)



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Az adatfeldolgozás második lépése: az adatok csoportosítása II.

- A csoportok számának meghatározása
 - Ha túl sok csoport – kezelhetetlen elemzés
 - Ha túl kevés csoport – nagy elemszám, csoporton belüli különbségek
 - Kis elemszámú minta (50 körül) esetén 8-9 csoport elegendő
 - Nagyobb elemszám esetén 10-20 csoport mérlegelhető
- Csoportintervallumok meghatározása
 - Az intervallumok nagysága általában 1,2,3,5,10...
 - Ajánlás: az intervallum alsó határai az egyes intervallumhosszok többszörösei.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Az adatfeldolgozás harmadik lépése: az adatok hitelessége

1. Validitás (érvényesség)

- Azt méri-e az módszer, ami a vizsgálat tárgya?
Pl. betegelégedettségi vizsgálatok: nem az ellátás minőségét, hanem az ellátással kapcsolatos beállítódást, véleményt méri

2. Reliabilitás (megbízhatóság)

- Mennyire pontosak az adatok?
 - mérőeszközfüggő,
 - adat-feldolgozás függő: alkalmazott számolási módszerek



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Reliabilitás (folyt.)

- Mennyire valóságűek az adatok?
 - Kérdező ill. megkérdezett szubjektivitása
 - Szimuláció / disszimuláció

A reliabilitás mértékét ellenőrizni lehet (ill. komolyabb kutatásokban elvárt):

- Split-half módszer: „Felezéses eljárás”.
 - Egy módszerrel kapott adatok halmazát tetszőlegesen két részre osztjuk, és megvizsgáljuk a két halmaz összegeinek összefüggését.
Pl. Az anyatejes táplálás vs tápszeres táplálást vizsgáljuk 2 éves gyermekeken. 148 adatunk van a testsúlyról és testmagasságról; Minden páratlan számú adat 1. halmaz; minden páros a 2. halmazba kerül. – az összegek összefüggését vizsgáljuk a módszerrel. – erős összefüggés megbízható adatra utal.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt

Az adatfeldolgozás negyedik lépése: az adatok elemzése I.

A minta jellemzőinek leírására – akár összehasonlítására – alkalmasak a:

LEÍRÓ STATISZTIKAI MÓDSZEREK

Leíró módszer lehet:

- » Gyakoriságok
 - » Középértékek
 - » Szóródások
- } számítása



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Leíró módszerek I. – Gyakorisági eloszlások

Kiinduló probléma: nagyszámú, rendezetlen adat. Pl. 346 fős minta pulzusértékei. → A (pl. nagyság szerinti) sorba rendezés körülményes, az adathalmaz nem válik áttekinthetőbbé.

Megoldás:

Gyakorisági eloszlás(ok) számítása

Lépései:

1. Értéktartomány meghatározása
2. Csoportok meghatározása
3. Gyakorisági eloszlás meghatározása



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Értéktartományok meghatározása

Értéktartomány:

A minta legkisebb és legnagyobb eleme által határolt zárt intervallum (amelybe a mintavételi eljárásból származó adatok mindegyike besorolható).



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Csoportok meghatározása

Lásd: 17.dia

- *„A csoportok számának meghatározása*
 - *Ha túl sok csoport – kezelhetetlen elemzés*
 - *Ha túl kevés csoport – nagy elemszám, csoporton belüli különbségek*
 - *Kis elemszámú minta (50 körül) esetén 8-9 csoport elegendő*
 - *Nagyobb elemszám esetén 10-20 csoport mérlegelhető*
- *Csoportintervallumok meghatározása*
 - *Az intervallumok nagysága általában 1,2,3,5,10...*
 - *Ajánlás: az intervallum alsó határai az egyes intervallumhosszok többszörösei. „*



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Gyakoriságok meghatározása

Csoporthoz tartozó gyakoriság:

Megmutatja, hogy a minta adataiból hány tartozik az adott csoportba.

Gyakorisági eloszlás:

A minta egyes elemei hogyan oszlanak meg az egyes csoportok között.

Formái:

- Abszolút gyakoriság: egy minta jellemzésére szolgál
- Relatív gyakoriság: több minta gyakorisági eloszlásának összehasonlítására szolgál.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Abszolút gyakoriság számítása

(Értéktartományok meghatározása → Intervallumok (csoportok)
létrehozása)



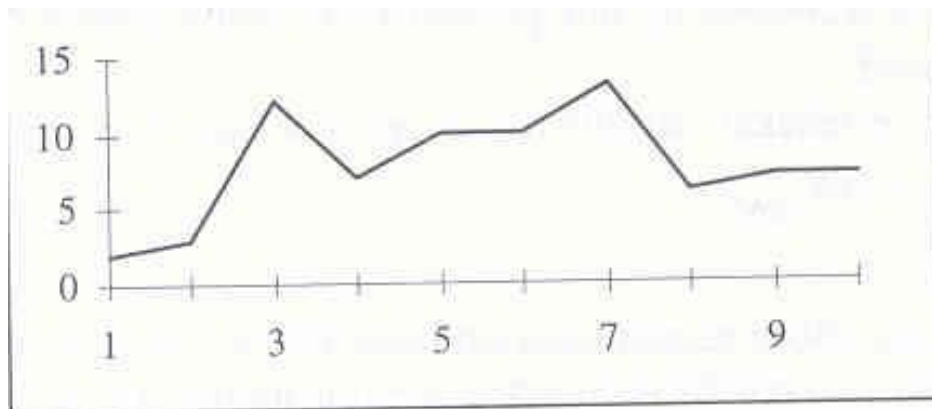
Annak összeszámolása, pontosan hány adat esik az adott
csoport tartományába.



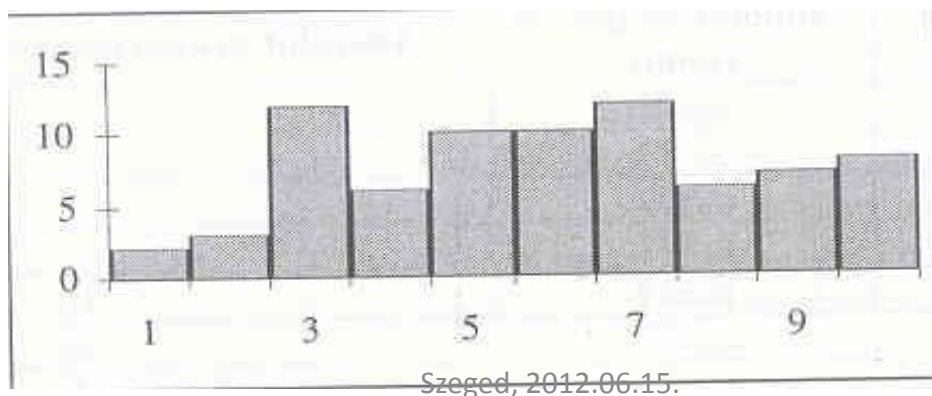
Abszolút gyakoriság

Abszolút gyakoriság ábrázolása

Leggyakrabban: gyakorisági poligon vagy hisztogram.



Poligon



Hisztogram

Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt

Relatív gyakoriság számítása

Két, különböző elemszámú minta esetén a gyakorisági eloszlás összehasonlítása **relatív gyakoriság** számításával lehetséges.

Számítási módszer:

1. A csoporthoz tartozó abszolút gyakoriság számítása.
2. Minden csoport abszolút gyakorisága x 100 = Relatív gyakoriság
elemszám

Ábrázolása (egy minta esetén):

- Kördiagram



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt

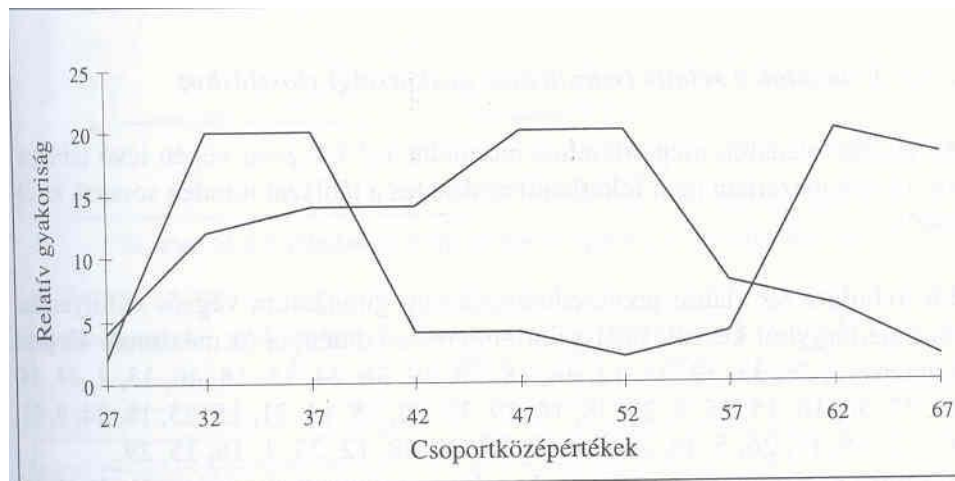


Két, különböző elemszámú minta összehasonlítása relatív gyakoriság számításával

Akkor alkalmazható, ha azonosak a csoport(intervallum)ok.

Ábrázolása:

- Gyakorisági poligon



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt

Leíró módszerek II. – Középértékek

Lényegük: a mintaeloszlás alapvető tendenciáinak jellemzése.

Típusaik:

1. Számítási közép
2. Medián
3. Módusz



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Az SZTE Kutatóegyetemi Kiválósági Központ tudásbázisának
kiszélesítése és hosszú távú szakmai fenntarthatóságának megalapozása
a kiváló tudományos utánpótlás biztosításával”



Számtani közép

- Számtani közép = átlag.
- Csak mérhető adatok esetén számítható.

Számítása:

$$\text{Átlag} = \frac{\text{Az összes adat összeadása}}{\text{elemszám}}$$



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Medián (Me)

~: A minta „közepét” kifejező középérték.

- Az az érték, melynél a minta egyik fele kisebb, a másik nagyobb.

Számítása:

- Páratlan számú adat esetén a középső adat.
- Páros számú adat esetén a két középső adat számtani közepe.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Módusz (Mo)

- ~: A minta elemei között leggyakrabban előforduló érték; vagy a legnagyobb gyakoriságú csoport számtani középértéke.
- Meghatározása akkor célszerű, ha az egyik érték gyakorisága jelentősen eltér a többitől.
 - Ábrázolás:
 - Gyakorisági poligonon a „csúcshoz” tartozó érték.
 - Ha a poligonnak két csúcsa van: **bimodiális eloszlás**.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Leíró módszerek III. – Szóródási mutatók

Szóródás: A minta azon tulajdonsága, hogy az egyes elemek eltérnek a középértékektől.

Jelentősége:

- Minta jellemzése pontosabb
- Különböző minták összehasonlítása

Fajtái:

- Szóródási terjedelem : a legnagyobb és legkisebb elem közötti távolság.
- Kvartilisek és interkvartilis félterjedelem
- Átlagos eltérés: A minta elemeinek számtani középétől való távolságának átlaga.
- **Variancia**
- **Szórás**
- Relatív szórás



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Variancia számítása

Jele: s^2

Számítás lényege:

1. A minta egyes elemeinek átlagtól való eltérését négyzetre emeljük, és a kapott értékeket összeadjuk. = ***négyzetösszeg***
2. $s^2 = \text{négyzetösszeg} / \text{a minta szabadságfoka}$

Variancia esetén a szabadságfok az elemszámnál eggyel kisebb szám.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Szórás számítása

Jele: s

Számítása: a variancia négyzetgyöke.

Értelmezése:

- A minta számtani középértékéhez hozzáadva vagy kivonva – normális eloszlás esetén – olyan értéktartományokat kapunk, melybe a minta meghatározott részei tartoznak.

Egyszeres eltérés: a minta 67 %-a.

Kétszeres eltérés: a minta 95 %-a.

Háromszoros eltérés: a minta 99%-a.

3 szigma szabály



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



A kutatás hipotéziseinek vizsgálata – matematikai statisztikai módszerekkel

MATEMATIKAI STATISZTIKA			
<i>Van-e szoros összefüggés?</i>			
Adatfajták → Változók száma ↓	Intervallum	Ordinális	Nominális
Kettő	Korrelációs számítás (7)	Rankkorreláció (14)	χ^2 -próba (15)
Kettő vagy több	Regresszióanalízis (8)		
Több	Parciális korreláció (7)		
	Faktoranalízis (9)		
	Klaszteranalízis (10)		

MATEMATIKAI STATISZTIKA			
<i>Jelentős-e a különbség?</i>			
Adatfajták → Minták száma ↓	Intervallum	Ordinális	Nominális
Egy	Egymintás t-próba (5)	Wilcoxon-próba (11)	χ^2 -próba (15)
Kettő	Kétmintás t-próba F-próba Welch-próba (5)	Mann-Whitney-próba (12)	χ^2 -próba (15)
Több	Variáncianalízis (6)	Kruskall-Wallis-próba (13)	χ^2 -próba (15)

Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Különbözőségvizsgálatok

Lehetséges alkalmazási területek:

- Egy minta esetén:
 - Előtte/utána (before & after) vizsgálatok (önkontrollos kísérletek): pl. egy csoportos oktatás hatása az ápolók tűszúrásos balesetekkel kapcsolatos ismereteire
- Két minta esetén:
 - Kontrollcsoportos vizsgálatok: pl. az új dolgozók egyik felét a hagyományos módon, a másikat új tematika szerint tanítjuk be.
- Több minta esetén:
 - Pl. Minden ápolónál ugyanazt a módszert alkalmazzuk az iv. injekciózás megtanítására. Van-e különbség az elsajátítás mértékében a korábbi középiskolai tanulmányok típusa szerint?



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Különbözőségvizsgálat egy minta esetén – egymintás t-próba

Egymintás t-próba során a két változó számtani középértéke
közötti szignifikáns különbség valószínűségét határozzuk meg.

Jele: t

Megjegyzés: t-próba végzésére MS Excel alkalmas!!



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Egymintás t-próba menete (összefoglalás)

1. t értékének meghatározása
2. t értékének szignifikancia-vizsgálata
 1. A minta szabadságfokának meghatározása
 1. Egymintás t-próba esetén a minta elemszámánál eggyel kisebb érték.
 2. a kiszámolt t érték összehasonlítása a t-eloszlás táblázatban található értékkel
 - Ha a számolt érték nagyobb, mint a táblázatban: a különbség szignifikáns. →érdeemes egy fokkal magasabb valószínűségi szintet is ellenőrizni (a táblázat azonos sorában, jobbra).
 - Ha a számolt érték kisebb, mint a táblázatban: a különbség nem szignifikáns.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Két minta összehasonlítása – kétmintás t-próba

Lényege: két, különböző személyektől származó adatsor számtani középértékeinek összehasonlítása.

Tudnivalók:

- csak mérhető adatok összehasonlítására alkalmas (mint minden t-próba)
- Kétmintás t-próba csak akkor alkalmazható, ha a két minta varianciája nem tér el szignifikánsan egymástól
 - A varianciák különbözőségének vizsgálata: F-próba.
- Ha a varianciák szignifikánsan különböznek: Welch (d) próbát kell végezni.



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Kétmintás t-próba elvégzése

1. A varianciák összehasonlítása F-próbával

1. F értékének meghatározása
2. F szignifikanciájának meghatározása:
 1. Minta szabadságfokának meghatározása
 2. Ha a kapott F érték nagyobb, mint az F táblázatban szereplő érték
 - szig. különbség: Welch próba
 3. Ha a kapott F érték kisebb, mint az F táblázatban szereplő érték
 - nincs szig. különbség: t-próba

2. t értékének meghatározása

3. t értékének szignifikancia-vizsgálata

1. Minta szabadságfokának meghatározása: két minta esetén az összelemszámnál 2-vel kisebb érték.
2. T szignifikanciájának meghatározása (hasonlóan, mint fenn)



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Kettőnél több minta összehasonlítása – varianciaanalízis

Lényege:

- Azt feltételezzük, hogy a minták adatai ugyanabból a „képzelt” populációból valók (nincs különbség a minták között)
nullhipotézis →
- Két, egymástól független módon megbecsüljük a „képzelt” populáció varianciáját → a két varianciát összehasonlítjuk egymással →
 - Ha nincs lényeges különbség: nincs jelentős különbség az egyes minták között
 - Ha van szignifikáns különbség: a csoportok szignifikánsan különböznek egymástól



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Variancia-analízis (ANOVA – analysis of variances)

~: A külső és a belső varianciát hasonlítjuk össze egymással.

- Belső variancia: feltételezése, hogy a minták között nincsenek lényeges különbségek; a különbségek a mintákon belüli egyes elemek eltéréseiből adódnak.
- Külső variancia: feltételezi, hogy a mintán belüli egyes elemek között nincsenek lényeges különbségek; az eltérések az egyes minták között vannak.
- A külső és belső variancia összehasonlítása: F-próba
 - A külső variancia szabadságfoka: minták számánál eggyel kisebb érték
 - A belső variancia szabadságfoka: az összelemszám és a minták számának különbsége
 - Mint előbb: Ha a számolt F nagyobb, mint az F táblázat megfelelő sorában szereplő érték: szignifikáns különbség van (ha kisebb – nincs különbség).



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Variancia-analízis: megjegyzés

~: Arra ad választ, hogy több minta egy dimenziója között van-e különbség → arra nem, hogy az egyes minták között van-e szignifikáns különbség → megoldás: Tukey-próba.

Pl. Minden ápolónál ugyanazt a módszert alkalmazzuk az iv. injekciózás megtanítására. Van-e különbség az elsajátítás mértékében a korábbi középiskolai tanulmányok típusa szerint?

- ANOVA eredménye: van szignifikáns különbség
- ANOVA és Tukey-próba együttes alkalmazásakor:
 - Van lényeges különbség a minták között (ANOVA)
 - Pl. az eü. szakközépben végzettek szignifikánsan jobban teljesítettek, mint a gimnáziumban végzettek (Tukey)



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Különbözőségvizsgálat megállapítható (nominális) adatok esetén

Ismételten:

- Az adat a skála megkülönböztetett pontjaira eshet. Pl. foglalkozás, érdemjegy, lakóhely,...
- Az egyes kódok (pl. számok) mindig ugyanazt a tulajdonságok jelölik, köztük matematikai/teljesítménybeli/rangsor különbség nincs. Pl. 1-férfi, 2-nő

A minták számától függetlenül használt statisztikai próba
nominális adatok különbözőségvizsgálatánál a

Chi^2 - (Khi^2 ; χ^2) próba



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Chi-négyzet próba kivitelezése

Kérdés: Van-e különbség a sepsis előfordulásában a szerint, hogy milyen osztályról érkezett a beteg:

Megoldás: (adatfeldolgozás)

1. Adatok kontingencia-táblázatba rendezése (példaszámokkal)

	Sepsis van	Sepsis nincs	Sorösszeg
Osztály 1.	50	20	70
Osztály 2.	30	20	50
oszlopösszeg	80	40	120

Chi-négyzet próba kivitelezése II.

2. Minden cellára vonatkozóan a várt érték kiszámítása
3. A várt és kapott érték különbségének összehasonlítása (Chi^2 kiszámítása)
4. A chi^2 szignifikanciájának meghatározása
 1. A Chi^2 szabadságfokának meghatározása
szf: $(\text{sorok száma}-1) \times (\text{oszlopok száma}-1)$
 2. Ha Chi^2 értéke nagyobb, mint a táblázatban: szignifikáns különbség
 3. Ha Chi^2 értéke kisebb, mint a táblázatban: nincs szignifikáns különbség



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



A Chi-négyzet próba alkalmazásai

1. X és y jelenség összefügg-e egymással:
függetlenségvizsgálat
2. **Változásvizsgálat:** pl. van-e hatása a transzfúziós tanfolyam elvégzésének a transzfúzió kivitelezésére?
3. **Homogenitás-vizsgálat:** pl. a hagyományos és új módszerrel betanított ápolók munkavégzése során van-e különbség?
4. **Illeszkedésvizsgálat:** pl. a mintában szereplő ápolók reprezentálják-e az ápolói populációt?



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Összefüggés vizsgálata matematikai statisztikai módszerekkel

Korreláció: A minta (legalább) két adata közötti összefüggés vizsgálata. Pl. összefügg-e az anya társadalmi helyzete és az anyatejes táplálás előfordulása?

Az eredmény lehet:

- Pozitív korreláció
- Negatív korreláció
- Korrelálatlanság (nincs összefüggés)



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Korrelációk típusai

1. Pozitív korreláció: az egyik változó növekvő értékével a másik változó növekvő értéke jár együtt. Pl. több gyakorlati óra – jobb gyakorlati teljesítmény
2. Negatív korreláció: az egyik változó növekvő értékével a másik változó csökkenő értéke jár együtt. Pl. több idő kómas betegek ápolásával – kevésbé hatékony kommunikáció éber betegekkel
3. Korrelálatlanság: nincs összefüggés a változók között. Pl. nyelvvizsgák száma – záróvizsga érdemjegye



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



A korrelációs együttható

Jele: r_{xy}

Értéke lehet: +1 és -1 között, ahol

- az előjel a korreláció irányát,
- az érték nagysága a korreláció erősségét mutatja meg.

Számítás menete:

1. Korrelációs együttható meghatározása
2. Korrelációs együttható szignifikanciájának meghatározása
 - A minta szabadságfokának meghatározása: a minták összelemszámánál a minták számával kisebb érték



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



A korrelációs együtttható szignifikanciájának meghatározása (folyt.)

3. A számított r abszolútértéke nagyobb, mint a táblázatban:
szignifikáns korreláció
4. A számított r abszolútértéke kisebb, mint a táblázatban:
nincs szignifikáns összefüggés



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Összefoglalás

- Az eredményes kutatómunka alapfeltétele a helyes adatelemzési módszer megválasztása.
- Helytelen elemzés helytelen következtetéseket eredményez; valamint pazarolja a kutatásba fektetett munkaórákat és más erőforrásokat.

Senki sem ért mindenhez ugyanolyan szinten a kutatási folyamat során – dolgozzunk csapatban, kérjünk segítséget!



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt



Az SZTE Kutatóegyetemi Kiválósági Központ tudásbázisának
kiszélesítése és hosszú távú szakmai fenntarthatóságának megalapozása
a kiváló tudományos utánpótlás biztosításával”



Köszönöm a figyelmet!



Szeged, 2012.06.15.

TÁMOP-4.2.2/B-10/1-2010-0012 projekt

