STATISZTIKA

Asszociáció

Ismérvek közötti kapcsolatok

Statisztikai ismérvek:

- Minőségi ismérvek
- Mennyiségi ismérvek
- Időbeli ismérvek
- Területi ismérvek

Eddig a sokaságokat egy ismérv szerint elemeztük, most a sokaságokat egyszerre két – egymással valamilyen kapcsolatban álló – megkülönböztető ismérv szerinti csoportosításban, azaz kombinációs táblába rendezve vizsgáljuk. A vizsgálat célja pedig az, hogy van-e és ha van, akkor milyen erősségű/jellegű a kapcsolat a vizsgált két ismérv között.

Ismérvek közötti kapcsolatok

- a két ismérv (x és y) **független egymástól**, ha x ismérv szerinti hovatartozás nem ad semmiféle többletinformációt az y szerinti hovatartozásról. (ezekkel nem kell foglalkoznunk)

(pl. gólyák száma vs születések száma; termésmennyiség vs mozilátogatók száma);

- a két ismérv között **sztochasztikus összefüggés** van, ha az egyik ismérvváltozathoz való tartozásból tendenciaszerűen, valószínűségi jelleggel következtethetünk a másik ismérv szerinti hovatartozásra. → Statisztika

(Pl. testsúly vs testmagasság; a föld aranykorona értéke vs termés mennyisége);

- a két ismérv **függvényszerű kapcsolatban** áll egymással, ha a vizsgált egységek x szerinti hovatartozásának ismeretében teljesen egyértelműen megmondható azok y szerinti hovatartozása is. (ezt a matematika vizsgálja)

(Pl. cukorrépa termése vs gyártott cukor mennyisége; mozilátogatók száma vs jegybevétel);

Sztochasztikus kapcsolatok

Különböző okozati jellege lehet az egyes ismérveknek:

x ismérv: ok (magyarázó változó)

y ismérv: okozat (eredményváltozó)

✓**ok-okozat** (pl. jövedelemnagyság és húsfogyasztás) (iskolai végzettség - munkanélküliség, biztonsági öv használata – baleset súlyossága);

? ok-okozat Vannak olyan esetek, amikor az ismérvek kölcsönösen hatnak egymásra, vagyis az ok-okozati viszony nem egyértelmű, az okság kölcsönös. (pl. ár és kereslet) (családi helyzet alakulása - alkoholfogyasztás)

Két változó oksági kapcsolatának feltételei

a) Az ok időben megelőzi az okozatot

b) a kettő között empirikus együttjárás van

c) ez a kapcsolat nem egy harmadik változó eredménye

Értelmezhető-e a kapcsolat iránya. Ha igen, pozitív vagy negatív irányú kapcsolatról van szó?

Nominális skála esetén, minőségi jellemzőknél nem értelmezhető.

1.pl. nem – dohányzás;

2.pl. iskolai végzettség – beosztás;

Rangsorok, mennyiségi ismérvek esetén értelmezhető.

1.pl. napi átlagos hőmérséklet – sörfogyasztás: minél melegebb van, általában annál több sör fogy és fordítva. *Pozitív kapcsolat*;

2.pl. ár és fogyasztás: minél drágább a termék, általában annál kevesebbet vásárolunk. *Negatív kapcsolat;*

Valós, vagy látszólagos kapcsolatról van szó?

A változók együttmozgásában, a látszólagos ok-okozati kapcsolatban egy harmadik ismérv játszik szerepet.

(Pl. TV eladások számának növekedése vs válások számának emelkedése)

Ismérvek közötti kapcsolatok

A két ismérv jellege szerint a következő sztochasztikus kapcsolatokat különböztethetjük meg:

- **asszociációs kapcsolat**: az egymással kapcsolatban álló ismérvek minőségi vagy területi ismérvek (pl.: nem (férfi,nő) – dohányzás ; nem - beosztás; iskolai végzettség - beosztás)
- **vegyes kapcsolat**: az egyik vizsgált ismérv területi vagy minőségi ismérv, a másik mennyiségi (pl.: iskolai végzettség -1 főre jutó bruttó havi jövedelem) **(nem - kereset; beosztás - életkor)**
- **korrelációs kapcsolat**: mindkét vizsgált ismérv mennyiségi ismérv (pl.: 1 főre jutó bruttó havi jövedelem-1 főre jutó élelmiszerfogyasztás) → egyszerre több ismérv között vizsgálható a sztochasztikus kapcsolat **(életkor - kereset; jövedelem - fogyasztás; tanulási idő - vizsgajegy)**
- **rangkorreláció**: mindkét ismérv ordinális mérési szintű, vagyis sorrendi skálán mérhető.

Ismérvek közötti kapcsolatok

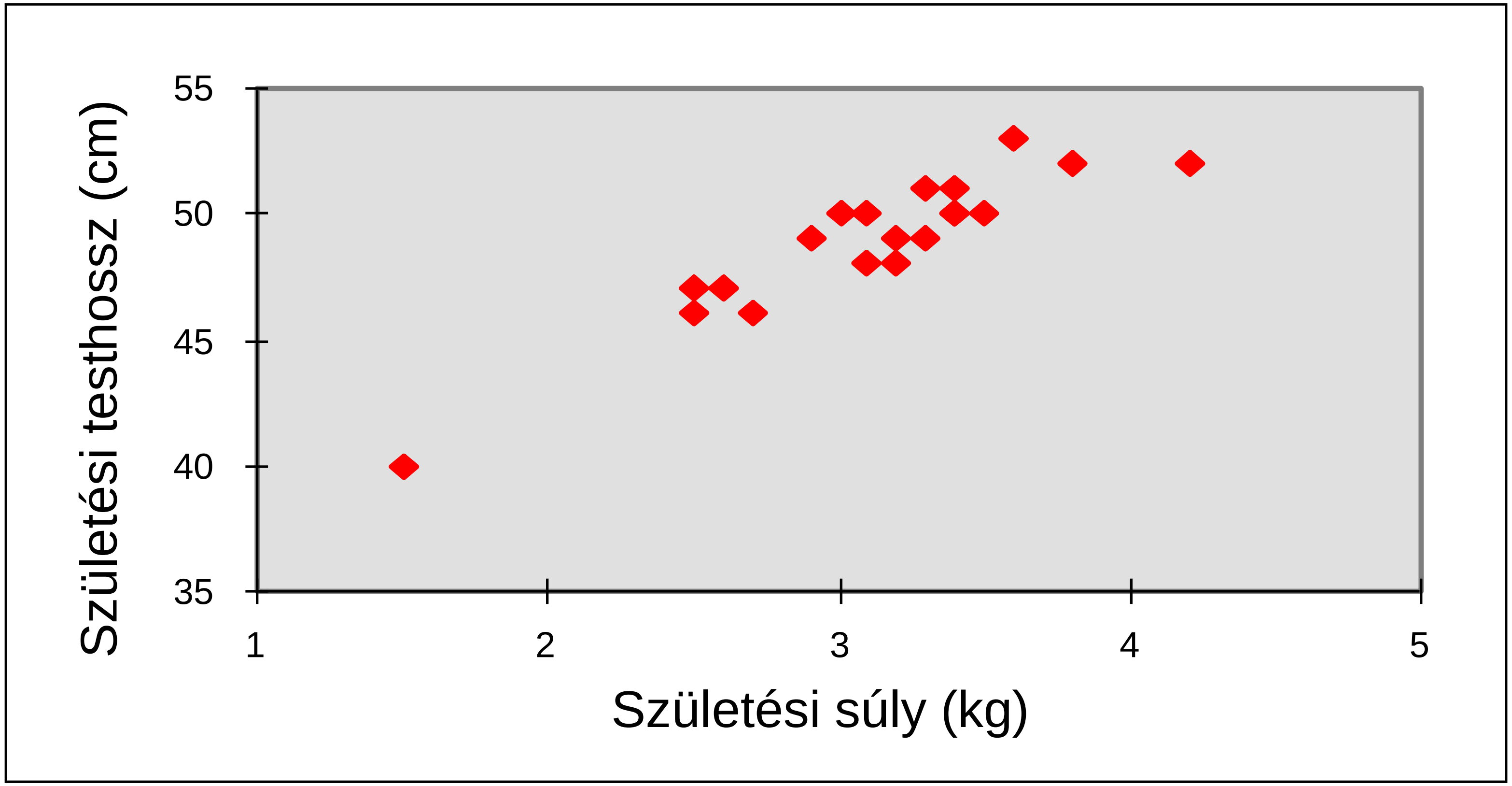
Az **asszociáció és a vegyes kapcsolat** esetén egyszerre csak két ismérv közötti kapcsolatot vizsgálhatjuk.

Arra keressük a választ, hogy a két ismérv között:

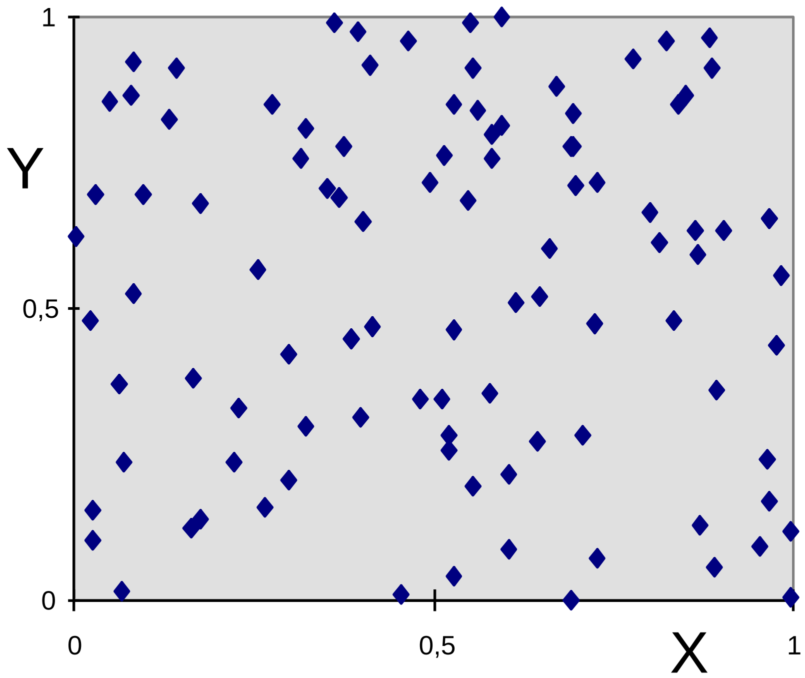
- van-e kapcsolat?
- ha van kapcsolat, akkor az milyen erős?

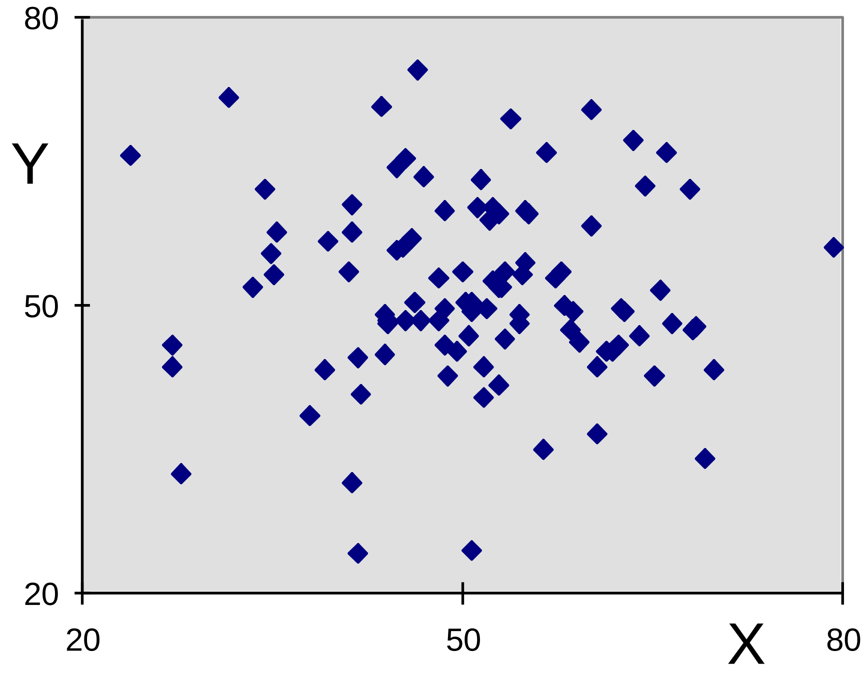
A **korrelációs kapcsolat** (mennyiségi ismérvek kapcsolatának a vizsgálata) több elemzési lehetőséget biztosít, hiszen itt azt is meg tudjuk vizsgálni, hogy az egyik ismérv milyen számszerű hatással van a másik (vagy több) ismérv alakulására.

Függ-e a születési testhossz a születési súlytól?
És fordítva?

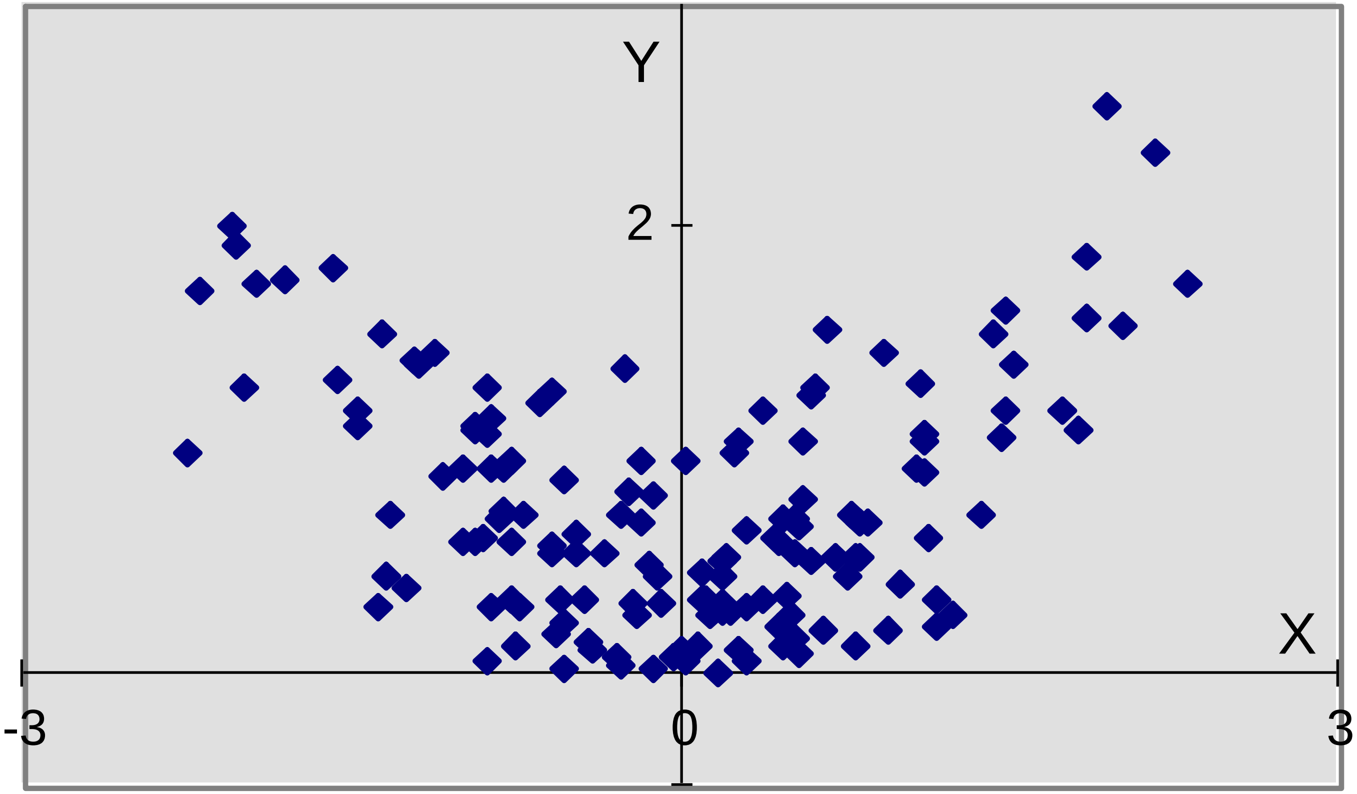


Függ-e az Y változó X-től?

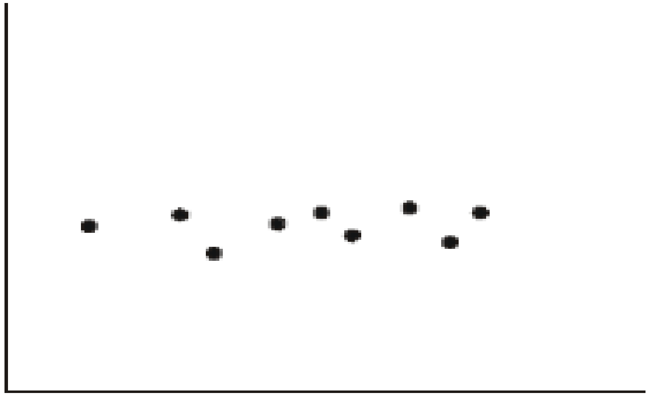




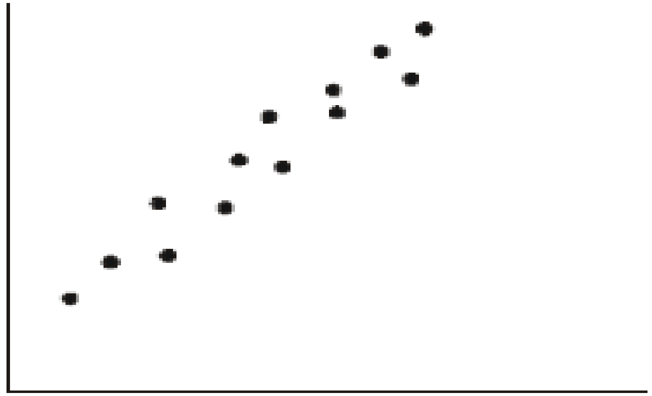
Függ-e az Y változó X-től?



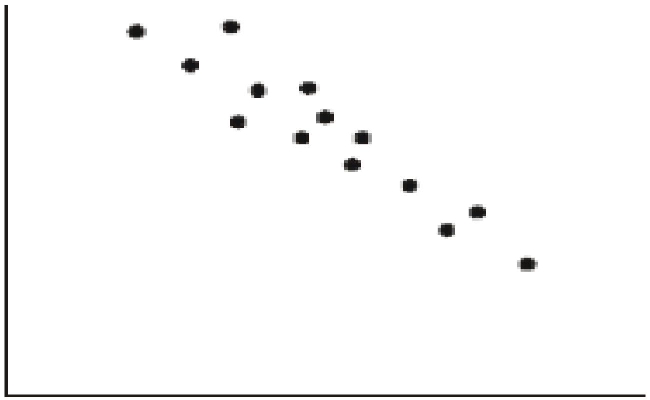
Pontdiagram



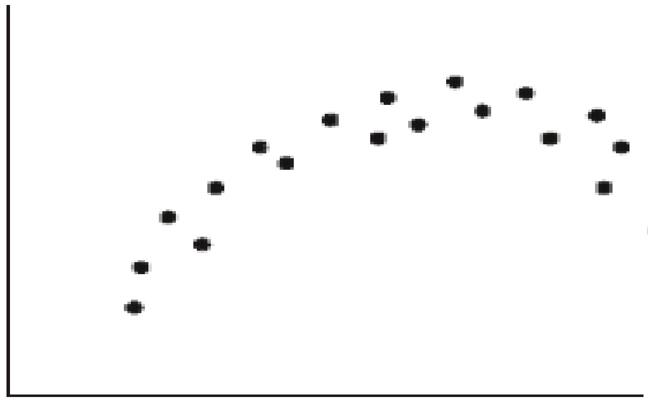
a) *X* és *Y* *független* egymástól – nincs az *Y* becslésére jól használható *f*(*X*) függvény



b) *X* és *Y* között *pozitív irányú* – alkalmasint lineáris – kapcsolat van



c) *X* és *Y* között *negatív irányú* – alkalmasint lineáris – kapcsolat van



d) *X* és *Y* között *nemlineáris* kapcsolat van

A függetlenség kölcsönös

FONTOS:

Ha Y független X-től,

akkor X is független Y-tól

Kontingencia tábla

X/Y	1	2	…	j	…	t	Σ
1	f_{11}	f_{12}	…	f_{1j}	…	f_{1t}	$f_{1\cdot}$
2	f_{21}	f_{22}	…	f_{2j}	…	f_{2t}	$f_{2\cdot}$
⋮	⋮	⋮		⋮		⋮	⋮
i	f_{i1}	f_{i2}	…	f_{ij}	…	f_{it}	$f_{i\cdot}$
⋮	⋮	⋮		⋮		⋮	⋮
s	f_{s1}	f_{s2}	…	f_{sj}	…	f_{st}	$f_{s\cdot}$
Σ	$f_{\cdot 1}$	$f_{\cdot 2}$	…	$f_{\cdot j}$	…	$f_{\cdot t}$	n

Kontingencia tábla

$f_{ij} =$ együttes gyakoriságok, tényleges gyakoriság a kontingencia tábla i sorában és j oszlopában

$f_{i.} =$ peremgyakoriságok, az összesen rovat gyakoriságai
Az x ismérvváltozattal rendelkező elemek száma

$f_{.j} =$ peremgyakoriságok, az összesen rovat gyakoriságai
az y ismérvváltozattal rendelkező elemek száma

n = a sokaság elemeinek a száma

A szochasztikus kapcsolat erősségének, szorosságának mérése

Általános séma a mutatószámokhoz:

0	0,5	1
Alsó határ függetlenség	**Közepes erősség**	**Felső határ függvényszerű**

Követelmény a szorossági mérőszámokkal szemben:

- legyen alsó és felső korlátja
- a kapcsolat teljes hiánya esetén értéke 0 legyen
- függvénykapcsolat esetén értéke 1 legyen
- a mérőszám értéke ne függjön a megfigyelések számától

Asszociációs kapcsolat szorosságának mérése

1) Alternatív ismérvek esetén:

2 x 2 kontingencia tábla:

X/Y	y_1	y_2	Összesen
x_1	f_{11}	f_{12}	$f_{1.}$
x_2	f_{21}	f_{22}	$f_{2.}$
Összesen	$f_{.1}$	$f_{.2}$	n

A két ismérv függetlensége esetén
$$\frac{f_{11}}{f_{21}} = \frac{f_{12}}{f_{22}}$$

Yule –együttható (Y):
$$Y = \frac{f_{11}f_{22} - f_{21}f_{12}}{f_{11}f_{22} + f_{21}f_{12}} \qquad -1 \leq Y \leq 1$$

Asszociációs kapcsolat szorosságának mérése

A Yule-mutató tulajdonságai:

$$0 \le |Y| \le 1$$

$Y = 0$ Teljes függetlenség, a kapcsolat teljes hiánya

$0 < |Y| < 1$ Sztochasztikus kapcsolat

$|Y| = 1$ Függvényszerű kapcsolat

b) YULE-együttható (Y)

Alternatív ismérvek esetén.

A dohányzás tiltásával kapcsolatban 800 fő véleménye, nemenkénti beosztásban

Megnevezés	Engedélyezzék		Ne engedélyezzék		Összesen	
Férfi (1)	440	f_{11}	60	f_{12}	500	$f_{1.}$
Nő (2)	160	f_{21}	140	f_{22}	300	$f_{2.}$
Összesen	600	$f_{.1}$	200	$f_{.2}$	800	n

$$\frac{f_{11}}{f_{21}} = \frac{f_{12}}{f_{22}}$$

$$f_{11} \times f_{22} - f_{21} \times f_{12} = 0$$

$$Y = \frac{f_{11} \times f_{22} - f_{21} \times f_{12}}{f_{11} \times f_{22} + f_{21} \times f_{12}}$$

$$Y = \frac{440 \times 140 - 160 \times 60}{440 \times 140 + 160 \times 60} = 0{,}73$$

A kapcsolat közepesnél erősebb, azaz: a férfiak inkább támogatják a dohányzás engedélyezését mint a nők.

Asszociációs kapcsolat szorosságának mérése

2) Általánosan alkalmazható mutatószám (alternatív és két ismérvváltozatnál több változattal rendelkező ismérvek esetén egyaránt): (ahol **s** az egyik ismérv változatainak, míg **t** a másik ismérv változatainak a számát jelenti): **azaz s = sorok száma; t = oszlopok száma és** n = az ismérv változatok együttes gyakoriságainak az összege, vagyis n = $\sum f_{i.} = \sum f_{.j}$

Csuprov-mutató (T):

$$T = \sqrt{\frac{\chi^2}{n \cdot \sqrt{(s-1)} \cdot \sqrt{(t-1)}}}$$,ahol $$\chi^2 = \sum_i \sum_j \frac{\left(f_{ij} - f^*_{ij}\right)^2}{f^*_{ij}}$$

,ahol $$f^*_{ij} = \frac{f_{i.} \cdot f_{.j}}{n}$$

$f_{i.}$ = az i-edik sor gyakoriságainak az összege;
$f_{.j}$ = a j-edik oszlop gyakoriságainak az összege;
$s \leq t$;

$f^*_{ij} =$ függetlenség esetén feltételezett gyakoriság a kontingencia tábla i sorában és j oszlopában

Asszociációs kapcsolat szorosságának mérése

A Csuprov-mutató tulajdonságai:

$0 \leq T \leq T_{max}$

$s \leq t$

A **Cramer-mutatót** (C) használjuk; **ha**

$$0 \leq C \leq 1 \qquad C = \frac{T}{T_{\max}}, ahol \;\; T_{\max} = \sqrt[4]{\frac{s-1}{t-1}}$$

$s \leq t$

$$s = t = 2$$

Alternatív ismérvek esetén, azaz, ha

Esetén Y és T mutatók is alkalmazhatók, a T mutató alakja ebben az esetben:

$$T = \frac{|f_{11} \cdot f_{22} - f_{12} \cdot f_{21}|}{\sqrt{f_{.1} \cdot f_{.2} \cdot f_{1.} \cdot f_{2.}}}$$

CRAMER mutató

Alapgondolat: a teljes függetlenségnek megfelelő (számított) gyakoriságokat f^*_{ij} összeveti a tényleges gyakoriságokkal f_{ij}

Minél nagyobb az eltérés, annál erősebb a kapcsolat.

Felhasználható χ^2 érték:

$$\chi^2 = \sum_{i=1}^{s} \sum_{j=1}^{t} \frac{\left(f_{ij} - f_{ij}^*\right)^2}{f_{ij}^*}$$

A Cramer együttható:

$$C = \sqrt{\frac{\chi^2}{n(s-1)}}$$

$$(s \le t)$$

$$0 \le C \le 1$$

C=0, ha a két ismérv független

C=1, ha a kapcsolatuk függvénykapcsolat

Példa:
(Szellemi foglakozásúaknak a foglalkozási főcsoport és nem szerinti megoszlása, 2000. III. negyedévben, ezer főben)

Foglalkozási főcsoport	Férfi (1)	Nő (2)	Összesen
I. (1)	159	83,2	242,2
II. (2)	185,2	254,4	439,6
III. (3)	179,4	325,1	504,5
IV. (4)	17,6	235,4	253,0
Összesen:	541,2	898,1	1439,3

A χ^2 számítása

Ismérv változó	f_{ij}	f_{ij}^*	$\left(f_{ij} - f_{ij}^*\right)^2 / f_{ij}^*$
1.1.	159	91,1	50,6
2.1.	185,2	165,3	2,4
3.1.	179,4	189,7	0,6
4.1.	17,6	95,1	63,2
1.2.	83,2	151,1	30,5
2.2.	254,4	274,3	1,4
3.2.	325,1	314,8	0,3
4.2.	235,4	157,9	38,0
Összesen:	1439,3	1439,3	187,0

$$\chi^2=187$$

$$C=\sqrt{\frac{187}{1439{,}3(2-1)}}=0{,}36$$

Közepesnél gyengébb kapcsolat van a szellemi foglalkozásúak munkaterülete és a nemhez való tartozás között.

Vegyes kapcsolat

Vegyes kapcsolat: az egyik vizsgált ismérv területi vagy minőségi ismérv (azaz nem mennyiségi ismérv), a másik mennyiségi ismérv (pl.: iskolai végzettség - 1 főre jutó bruttó havi jövedelem)

A vegyes kapcsolat elemzése során azt vizsgáljuk, hogy a mennyiségi ismérv szóródását mennyiben befolyásolja a minőségi vagy a területi ismérv szerinti csoportosítás.

Heterogén sokaságok

összetett, minőségileg különböző részekből állnak.
Heterogén sokaság átlaga a részsokaságokra számított átlagok súlyozott átlaga.

Jelölések:

$\bar{x}_j$: j-edik csoport átlaga

n_j : j-edik csoport tagszáma

$j = 1, \ldots, m$: a csoportok száma

$\frac{n_j}{n} = w_j$: súlyarány

$$\bar{x} = \frac{\sum_{j=1}^{m} n_j \bar{x}_j}{\sum_{j=1}^{m} n_j} = \sum_{j=1}^{m} w_j \bar{x}_j$$

$\bar{x}$: a teljes sokaságra számított átlag

Heterogén sokaságok

Jelölések:

n = a sokaság tagszáma

m = a csoportok száma

n_j = a j-edik sokaság tagszáma

$\overline{x}_j$ = a j-edik csoport átlaga

$\overline{x}$ = a sokaság átlaga (főátlag)

x_{ij} = a j-edik sokaság i-edik eleme ahol $\sum_i x_{i_j} = n_j$

Csoportok	Elemszám	Csoportátlag	Csoportok szórása
C_1	n_1	$\bar{x}_1$	σ_1
C_2	n_2	$\bar{x}_2$	σ_2
…			
C_k	n_k	$\bar{x}_k$	σ_k
…			
C_m	n_m	$\bar{x}_m$	σ_m
Összesen	n	$\bar{x}$	σ

Példa:
Egy vidéki nagyváros ingatlanügynökségén értékesítésre váró ingatlanok

Eladási ár (m Ft)	panellakások száma (db)	Nem-panelből készült lakások száma (db)	Összes lakás (db)
6,1 – 8,0	8	2	10
8,1 – 10,0	15	5	20
10,1 – 15,0	34	12	46
15,1 – 20,0	24	14	38
20,1 – 25,0	7	19	26
25,1 – 30,0	2	8	10
Összesen	90	60	150
	$\bar{x}_p$	$\bar{x}_{np}$	$\bar{x}$
	σ_P	σ_{NP}	σ

$$\bar{x}_P = \frac{8\cdot 7{,}0 + 15\cdot 9{,}0 + 34\cdot 12{,}5 + 24\cdot 17{,}5 + 7\cdot 22{,}5 + 2\cdot 27{,}5}{90} = \frac{1\,248{,}5}{90} = 13{,}872$$

$$\bar{x}_{NP} = \frac{2\cdot 7{,}0 + 5\cdot 9{,}0 + 12\cdot 12{,}5 + 14\cdot 17{,}5 + 19\cdot 22{,}5 + 8\cdot 27{,}5}{60} = \frac{1101{,}5}{60} = 18{,}358$$

$$\bar{x} = \frac{10\cdot 7{,}0 + 20\cdot 9{,}0 + 46\cdot 12{,}5 + 38\cdot 17{,}5 + 26\cdot 22{,}5 + 10\cdot 27{,}5}{150} = \frac{2\,350}{150} = 15{,}67$$

$$\sigma_P = \sqrt{\frac{8\cdot(7{,}0-13{,}872)^2 + 15\cdot(9{,}0-13{,}872)^2 + \ldots + 2\cdot(27{,}5-13{,}872)^2}{90}} = \sqrt{\frac{2\,006{,}3}{90}} = 4{,}72$$

$$\sigma_{NP} = \sqrt{\frac{2\cdot(7{,}0-18{,}358)^2 + 5\cdot(9{,}0-18{,}358)^2 + \ldots + 8\cdot(27{,}5-18{,}358)^2}{60}} = \sqrt{\frac{2\,112{,}55}{60}} = 5{,}93$$

$$\sigma = \sqrt{\frac{10\cdot(7{,}0-15{,}67)^2 + 20\cdot(9{,}0-15{,}67)^2 + \ldots + 10\cdot(27{,}5-15{,}67)^2}{150}} = \sqrt{\frac{4\,843{,}335}{150}} = 5{,}68$$

A könnyebb áttekinthetőség kedvéért **az eladott lakásokkal** kapcsolatos számításokat táblázatba foglaljuk:

Eladott lakások			
lakástípus	**ingatlanok száma**	**A lakások eladási árának csoportonkénti átlaga**	**A lakások eladási árának csoportonkénti szórása**
Panel	**90**	**13872**	**4,72**
Nem-panel	**60**	**18358**	**5,93**
Összesen	**150**	**15670**	**5,68**
	n	$\bar{x}$	σ

Jelölések

$x_{ij} - \overline{x}$ = teljes eltérés (d_{ij})

$(x_{ij} - \overline{x}_j)$ = belső eltérés (B_{ij})

$(\overline{x}_j - \overline{x})$ = külső eltérés (K_{ij})

σ^2 = teljes szórásnégyzet

σ_B^2 = belső szórásnégyzet

σ_K^2 = külső szórásnégyzet

Szórásnégyzetek kiszámítása

$$\sigma^2 = \frac{\sum_{i=1}^{n_j} \sum_{j=1}^{m} (x_{ij} - \overline{x})^2}{n} = \frac{S}{n}$$

S: teljes eltérés-négyzetösszeg

$$\sigma_B^2 = \frac{\sum\sum (x_{ij} - \overline{x}_j)^2}{n} = \frac{\sum n_j \sigma_j^2}{n} = \frac{S_B}{n}$$

S_B: belső eltérés-négyzetösszeg

$$\sigma_K^2 = \frac{\sum n_j (\overline{x}_j - \overline{x})^2}{n} = \frac{S_K}{n}$$

S_K: külső eltérés-négyzetösszeg

Összefüggések

$$d_{ij} = B_{ij} + K_{ij}$$

$$x_{ij} - \bar{x} = (x_{ij} - \bar{x}_j) + (\bar{x}_j - \bar{x})$$

Teljes eltérés — Belső eltérés — Külső eltérés

$$\sigma^2 = \sigma_B^2 + \sigma_K^2$$

Teljes szórásnégyzet — Belső szórásnégyzet — Külső szórásnégyzet

$$S = S_B + S_K$$

Teljes eltérés négyzet összeg — Belső eltérés négyzet összeg — Külső eltérés négyzet összeg

Feladat:

Egy főiskolán 4 szakon folyik bachelor képzés. Az alábbi táblázatban a hallgatók napi tanulásra fordított idejére vonatkozó adatok találhatók:

Szak	Napi tanulásra fordított idő (óra)		Hallgatók %-os megoszlása
	átlaga	szórása	
Emberi erőforrás	1,5	1,2	24
Gazdálkodás menedzsment	2,25	0,8	26
Nemzetközi gazdálkodás	1,75	1,5	20
Pénzügy-számvitel	2,75	1,3	30

Számítsa ki a $\sigma_B, \sigma_K, \sigma$ mérőszámokat és értelmezze azokat!

Megoldás

$$\bar{x} = 0{,}24 \cdot 1{,}5 + 0{,}26 \cdot 2{,}25 + 0{,}2 \cdot 1{,}75 + 0{,}3 \cdot 2{,}75 = 2{,}12$$

$$\sigma_K^2 = 0{,}24 \cdot (1{,}5 - 2{,}12)^2 + \ldots + 0{,}3 \cdot (2{,}75 - 2{,}12)^2 = 0{,}2431$$

$$\sigma_k = 0{,}49$$

$$\sigma_B^2 = 0{,}24 \cdot 1{,}2^2 + 0{,}26 \cdot 0{,}8^2 + 0{,}2 \cdot 1{,}5^2 + 0{,}3 \cdot 1{,}3^2 = 1{,}469$$

$$\sigma_B = 1{,}212$$

$$\sigma^2 = \sigma_B^2 + \sigma_K^2$$

$$\sigma^2 = 1{,}469 + 0{,}2431 = 1{,}7121 \rightarrow \sigma = 1{,}308$$

A vegyes kapcsolat mutatószámai

Szórásnégyzet-hányados: megmutatja, hogy a minőségi vagy területi ismérv szerinti csoportosítás hány %-ban befolyásolja a mennyiségi ismérv szóródását.

$$H^2 = \frac{\sigma_K^2}{\sigma^2} = 1 - \frac{\sigma_B^2}{\sigma^2} = \frac{S_K}{S} = 1 - \frac{S_B}{S}$$

Szóráshányados: a szórásnégyzet-hányados négyzetgyöke, amely megmutatja, hogy milyen szoros a kapcsolat a nem mennyiségi (csoportosító) és a mennyiségi ismérv között.

$$H = \sqrt{H^2} = \sqrt{\frac{\sigma_K^2}{\sigma^2}} = \frac{\sigma_K}{\sigma} = \sqrt{1 - \frac{\sigma_B^2}{\sigma^2}} = \sqrt{\frac{S_K}{S}} = \sqrt{1 - \frac{S_B}{S}}$$

A vegyes kapcsolat mutatóinak értelmezése

$$0 < H < 1$$

$$0 < H^2 < 1$$

Sztochasztikus kapcsolat

$$H = H^2 = 0$$

Teljes függetlenség, a kapcsolat teljes hiánya

$$H = H^2 = 1$$

Függvényszerű, determinisztikus kapcsolat

KORRELÁCIÓ- ÉS REGRESSZIÓSZÁMÍTÁS

Alapfogalmak

- A mennyiségi ismérvek közötti kapcsolatot **korrelációnak** nevezzük;
- A **korrelációszámítás**: a mennyiségi ismérvek közötti kapcsolat szorosságának mérése;
- A **regressziószámítás**: a mennyiségi ismérvek egymásra gyakorolt hatásának számszerűsítésével, e hatások irányának és mértékének megállapításával foglalkozik;

Ha a korreláció mögött egyirányú okozati összefüggés állapítható meg:

- az ok szerepét betöltő ismérvet ***magyarázó változónak*** *(X)*;
- az okozat szerepét játszó ismérvet pedig ***eredményváltozónak*** *(Y)* nevezzük;

A kapcsolat szorosságának mérőszámai

A kovariancia

Az X és Y mennyiségi változók közötti kapcsolat irányát mutatja meg.

A megfelelő átlagtól vett ($x - \bar{x}$) és ($y - \bar{y}$) eltéréseken alapszik.

$$d_x = x - \bar{x} \qquad d_y = y - \bar{y}$$

$$C = \frac{\Sigma d_x d_y}{n - 1} = \frac{\Sigma xy}{n - 1} - \bar{x} \cdot \bar{y}$$

$$C = r \cdot s_x \cdot s_y$$

s_x: x szórása; s_y: y szórása; r: az x és y közötti korrelációs együttható

Kovariancia tulajdonságai

- Kovariancia előjele a kapcsolat irányát mutatja.
- A kovariancia abszolút mértékének nincs határozott felső korlátja.
- *A kovariancia* abszolút értéke akkor maximális, ha x és y között *lineáris függvénykapcsolat* áll fenn, ez esetben

$$\left|C_{\max}\right| = \sigma_x \cdot \sigma_y$$

- A kovariancia a két változóban szimmetrikus, X és Y szerepe a formulában felcserélhető.
- Az ismérvek *függetlensége* esetén a C = 0. (Megfordítva nem áll: ha C=0, akkor a kapcsolat *korrelálatlan*, de nem feltétlenül független, a függetlenség szigorúbb feltételeket jelent, mint a korrelálatlanság)

- Mivel C mértékegység-függő, ezért célszerű elosztani a maximális értékkel, és akkor egy előjeles mutatót kapunk: a *lineáris korrelációs együtthatót*:

$$r_{xy} = \frac{C_{xy}}{\sigma_x \cdot \sigma_y}$$ kiszámítható így is: $$r_{xy} = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \cdot \sum d_y^2}}$$

- Ez a *lineáris* kapcsolat szorosságát méri, ha tehát 0-hoz közel álló értéket mutat, akkor lehet, hogy lineáris a kapcsolat, de gyenge; de az is lehet, hogy szoros a kapcsolat, de *nem lineáris*.
- Egy változó esetén az önmagára vonatkozó kovariancia a variancia, azaz szórásnégyzet. A szórásnégyzet tehát a kovariancia speciális esete:

$$C_{xx} = \sigma_x^2$$

Egy vállalat dolgozóinak keresete és havi megtakarítása

Dolgozó	Bér (Ft/fő)	Havi megtakarítás (Ft/hó)	d_x	d_y	$d_x d_y$	d_x^2	d_y^2
1	120000	13000	-13000	-3010	39130000	169000000	9060100
2	90000	10000	-43000	-6010	258430000	1849000000	36120100
3	220000	35000	87000	18990	1652130000	7569000000	360620100
4	150000	18000	17000	1990	33830000	289000000	3960100
5	100000	12000	-33000	-4010	132330000	1089000000	16080100
6	115000	12500	-18000	-3510	63180000	324000000	12320100
7	160000	20000	27000	3990	107730000	729000000	15920100
8	130000	13800	-3000	-2210	6630000	9000000	4884100
9	145000	14000	12000	-2010	-24120000	144000000	4040100
10	100000	11800	-33000	-4210	138930000	1089000000	17724100
Összesen	**1330000**	**160100**	**0**	**0**	**2408200000**	**13260000000**	**480729000**

$$\bar{y} = \frac{13000 + 10000 + \cdots + 14000 + 11800}{10} = 16010$$

$$\bar{x} = \frac{120000 + 90000 + \cdots + 145000 + 100000}{10} = 133000$$

Kovariancia

$$C = \frac{\Sigma d_x d_y}{n - 1} = \frac{\Sigma xy}{n - 1} - \bar{x} \cdot \bar{y} = \frac{2408200000}{9} = 267577777{,}8$$

Értelmezés: a dolgozók keresete és a havi megtakarított összege közötti kapcsolat pozitív irányú.

A korrelációs együttható

- A korrelációs együttható a lineáris korreláció szorosságának legfontosabb mérőszáma.
- A kapcsolat hiányát (korrelálatlanság) az $r = 0$ érték jelzi.
- Az r előjele a korreláció irányát mutatja. Függvényszerű lineáris kapcsolatnak – az iránytól függően – az $r = +1$, illetve $r = -1$ értékek felelnek meg.
- A szélsőséges helyzetek között az együttható abszolút értéke a kapcsolat szorosságáról tájékoztat.

Korrelációs együttható

$$\mathbf{r} = \frac{\mathbf{C}}{s_{\mathbf{x}} \cdot s_{\mathbf{y}}} = \frac{\sum \mathbf{d_x d_y}}{\sqrt{\sum \mathbf{d_x}^2 \sum \mathbf{d_y}^2}} = \frac{\sum \mathbf{xy - n \cdot \bar{x} \cdot \bar{y}}}{\sqrt{(\sum \mathbf{x^2 - n \cdot \bar{x}^2})(\sum \mathbf{y^2 - n \cdot \bar{y}^2})}}$$

$$\Sigma d_x \cdot d_y = \Sigma xy - n\,\bar{x}\,\bar{y}$$

$$\Sigma d_x^2 = \Sigma x^2 - n\,\bar{x}^2$$

$$\Sigma d_y^2 = \Sigma y^2 - n\,\bar{y}^2$$

Korrelációs együttható

Dolgozó	Bér (Ft/fő)	Havi megtakarítás (Ft/hó)	$\sum d_x$	$\sum d_y$	$\sum d_x d_y$	$\sum d_x^2$	$\sum d_y^2$
Összesen	**1330000**	**160100**	**0**	**0**	**2408200000**	**13260000000**	**480729000**

$$r = \frac{C}{s_x \cdot s_y} = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}} = \frac{2408200000}{\sqrt{13260000000 \cdot 480729000}} = 0{,}954$$

Értelmezés: a dolgozók keresete és a havi megtakarított összege közötti kapcsolat pozitív irányú és erős.

Determinációs együttható (r^2)

- A determinációs együttható megmutatja, hogy a magyarázó változó hány %-ban befolyásolja az eredményváltozó szóródását.
- Jele: r^2
- A determinációs együttható jellemzi:
  - ✓ a regressziós függvény illeszkedését,
  - ✓ a modell magyarázó erejét.

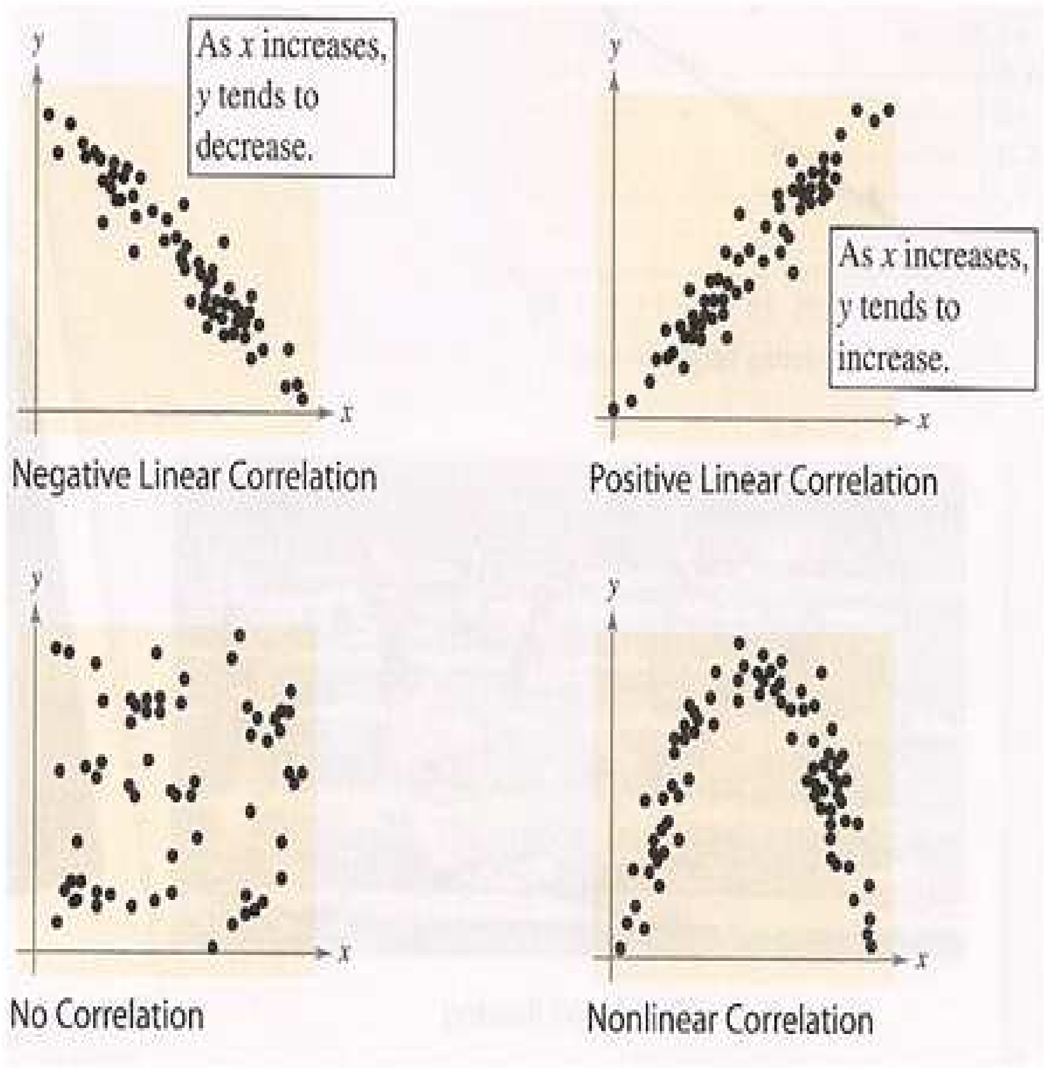
Determinációs együttható

$$r^2 = 0{,}954^2 = 0{,}9098 = 90{,}98\%$$

Értelmezés: a dolgozók keresete 90,98%-ban befolyásolja a havi megtakarított összeg szóródását.

Korreláció

- Az összetartozó (x, y) pontpárok ábrázolása;
- Ha létezik egy képzeletbeli egyenes, amely mentén helyezkednek el a pontpárok → lineáris korreláció;
- Az összefüggés irányától függően pozitív vagy negatív;
- Ha nincs ilyen egyenes → a változók korrelálatlanok (de nem feltétlenül függetlenek!)



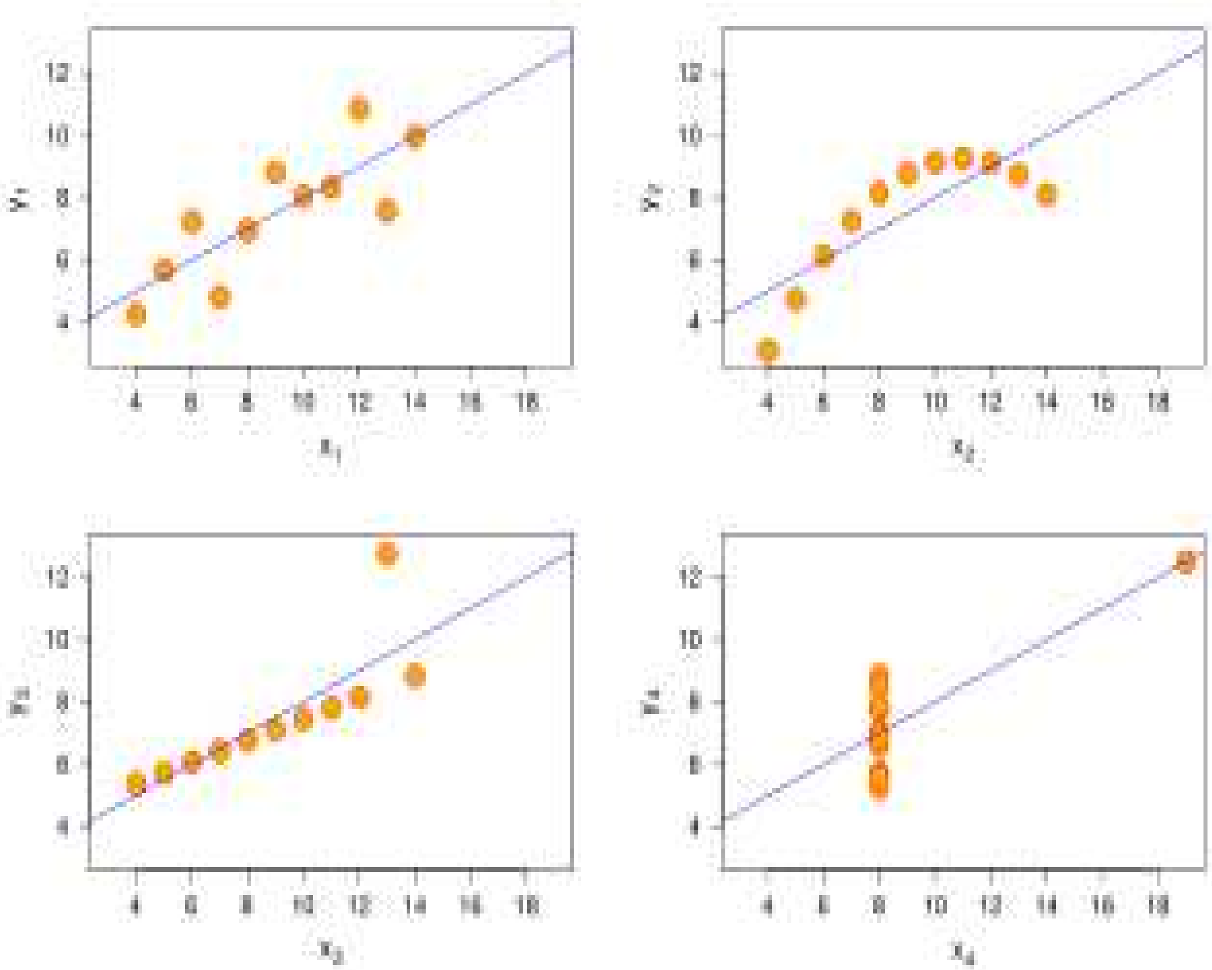
A korrelációs együttható legfontosabb tulajdonságai

1. **Ha nincs lineáris korreláció, akkor a korrelációs együttható értéke: 0, függvénykapcsolat esetén a korrelációs együttható értéke +1,00, illetve -1,00.**
2. **A korrelációs együttható értéke független a mértékegységektől [pl. a testmagasság és testsúly közötti korreláció független a változók mértékegységétől (kg, font, cm, inch)].**
3. **A korrelációs együttható szimmetrikus** (x korrelációja y-nal = y korrelációja x-el), azaz r(x, y) = r(y, x);
4. **A lineáris korrelációs együttható a lineáris összefüggést** méri, nem az összefüggést általában;
5. **A korrelációs együttható értékét az outlier (kiugró) értékek igen erősen befolyásolják. A kiugró érték lehet**
   - **szabálytalan, torzult eloszlás eredménye ⇒ a transzformáció;**
   - **mérési hiba** ⇒ a mérés meg**ismétlése / az érték kizárása;**
6. **A korreláció nem feltétlenül jelent ok-okozati kapcsolatot, mivel az**
   - **Az x változó befolyásolhatja az y változót;**
   - **Az y változó befolyásolhatja az x változót;**
   - **Egy harmadik tényező befolyásolhatja x-et és y-t**
     - ✓ **egy irányba (pozitív korreláció), vagy**
     - ✓ **különböző irányokba (negatív korreláció);**

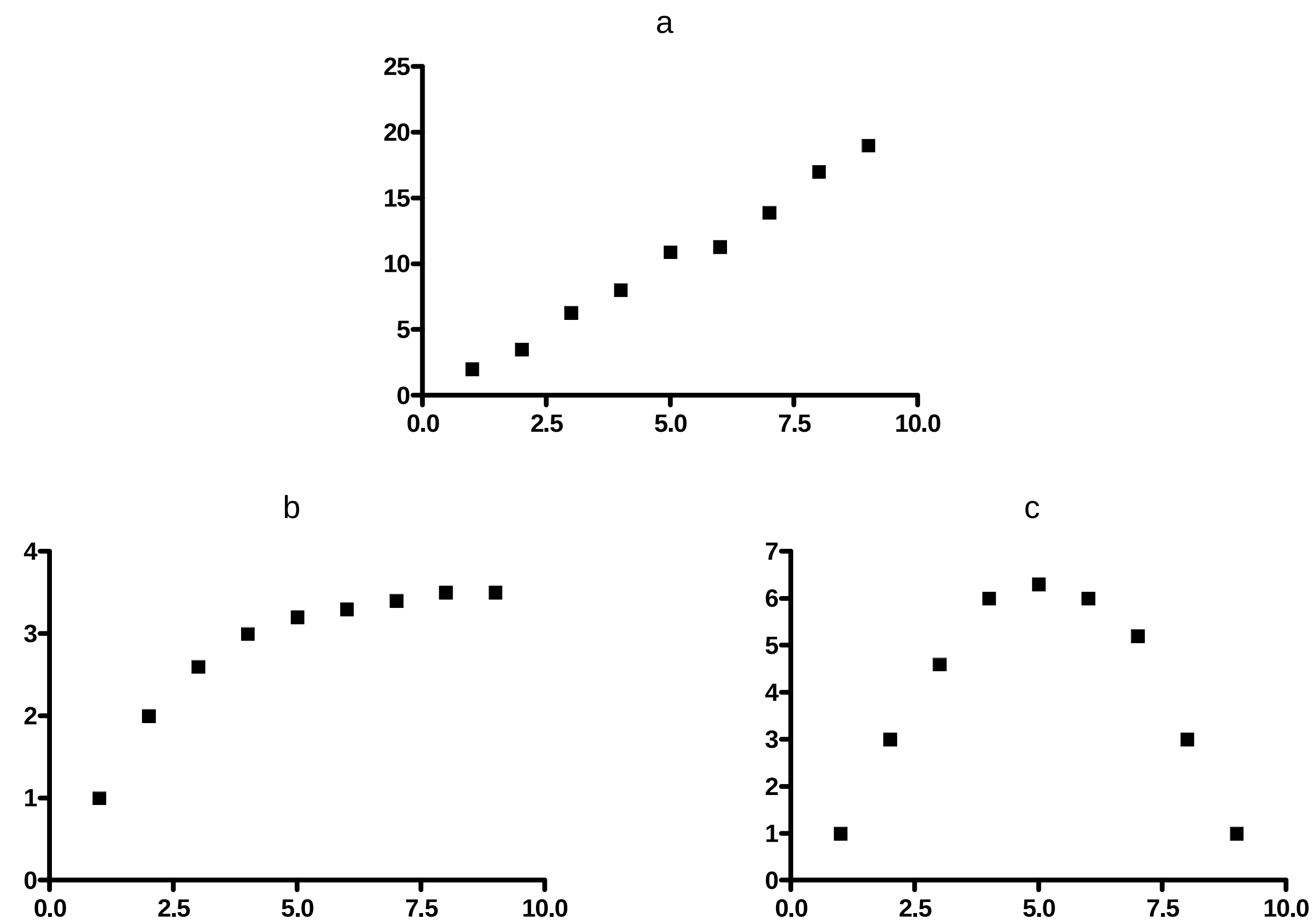
Outlierek, linearitás

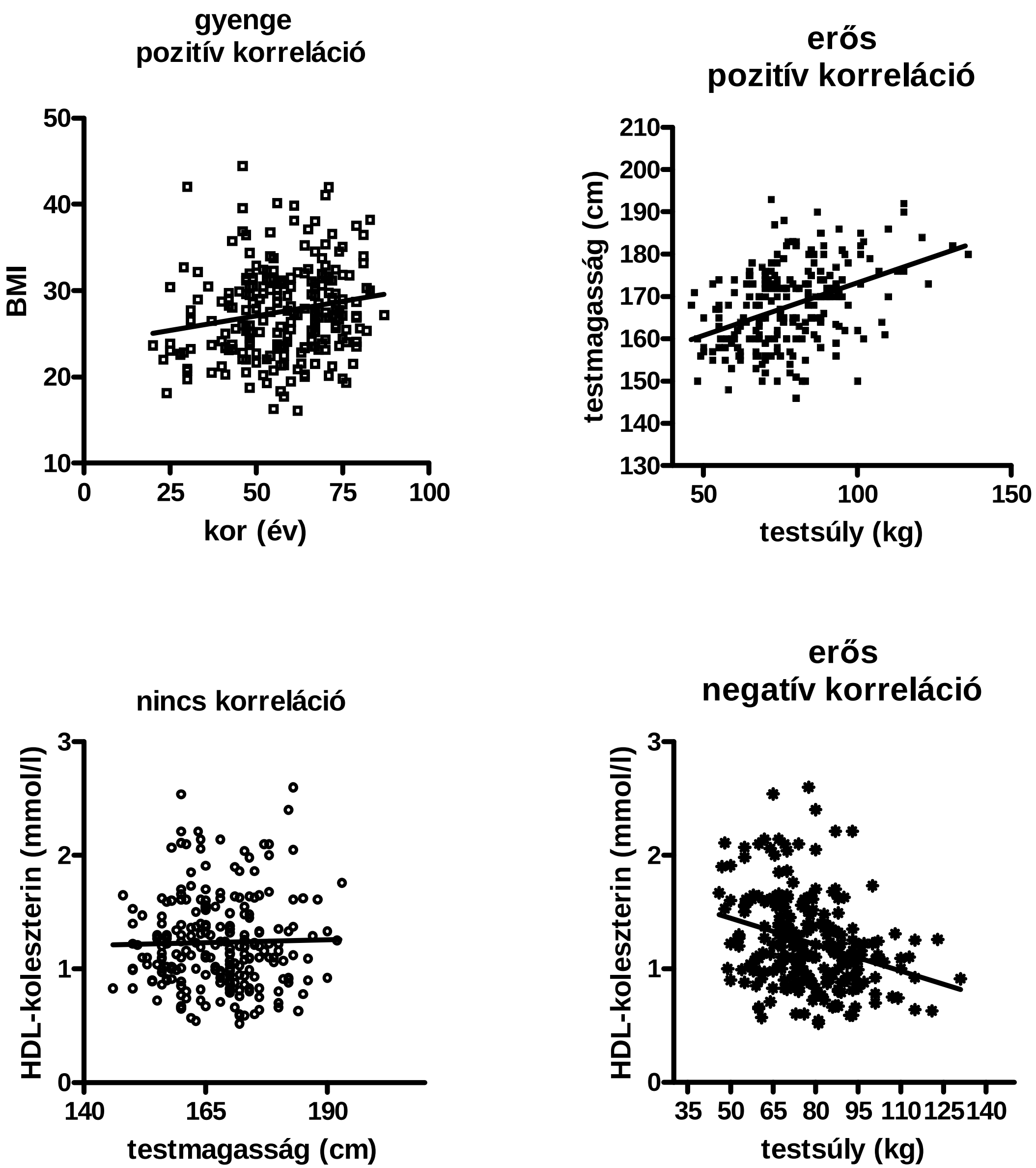
- Regressziós egyenes:

  y=0,5x + 3
- r = 0,816



- 2. nemlineáris kapcsolat!
- Outlier nélkül
  - 3. r=1
  - 4. r=0


a
25
20
15
10
5
0
0.0
2.5
5.0
7.5
10.0
b
4
3
2
1
0
0.0
2.5
5.0
7.5
10.0
c
7
6
5
4
3
2
1
0
0.0
2.5
5.0
7.5
10.0


gyenge
pozitív korreláció
BMI
kor (év)
erős
pozitív korreláció
testmagasság (cm)
testsúly (kg)
nincs korreláció
HDL-koleszterin (mmol/l)
testmagasság (cm)
erős
negatív korreláció
HDL-koleszterin (mmol/l)
testsúly (kg)

Mi a teendő?

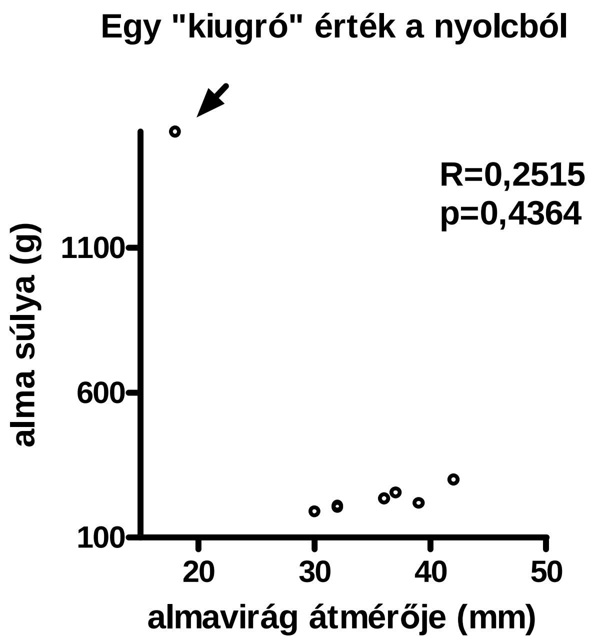
Outlierek

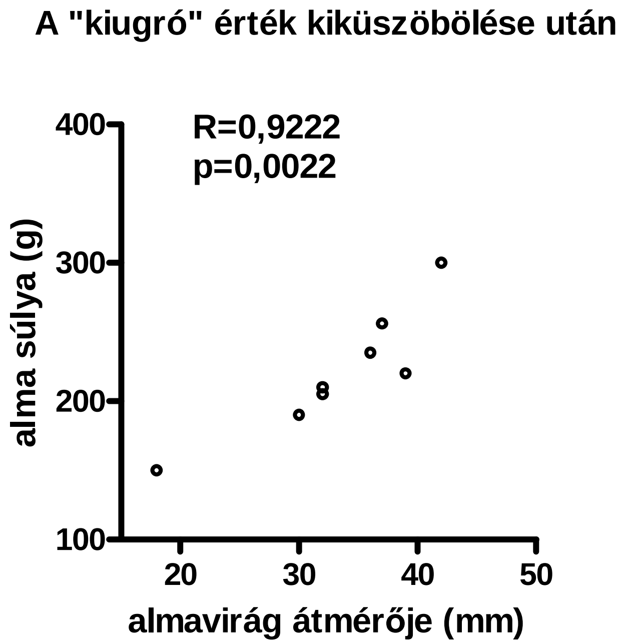
- Szubjektív, mi az outlier (általában 2 SD-n kívül)
- Ellenőrizni az adatbázist
  - Tényleg valós érték? Elütés? Mérési hiba?
- Ha valós adat – egyedi mérlegelés
  - Nem szerencsés automatikusan kizárni
  - Ha nagyon torzítja az összképet ez lehetséges;
- Ellenőrizni, nincs-e az outliereknek jelentősége? Biztos outlier, vagy csak nem passzol a mi teóriánkba?

Nemlinearitás

- Ha nem monoton, nincs értelme a korrelációnak. Ha monoton…
- Transzformációval lineárissá tehető? (pl. logaritmizálás – elsőként az ábra skálázását módosítva tesztelhető (Axis/Scaling))
- Nemparaméteres teszt végzése (Spearman rangkorrelációs teszt)
  - Kevésbé érzékeny
- Keresni egy függvényt, ami illeszkedik rá, helyesen leírja
- Az egyik változó mentén 4-5 egyenlő „szélességű” csoportra osztom a mintát. ANOVA-t végzek, úgy, hogy ez a csoportosító változó.

EGY KIUGRÓ (OUTLIER) ÉRTÉK HATÁSA A KORRELÁCIÓS EGYÜTTHATÓ NAGYSÁGÁRA ÉS SZIGNIFIKANCIÁJÁRA





A korreláció (a két változó közötti kapcsolat) erősségének megítélése. A leegyszerűsített megoldás.

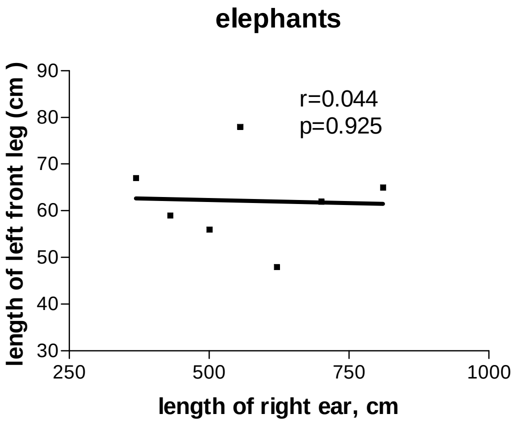
Korrelációs koefficiens	**A kapcsolat erőssége**
0-0,25	**Nincs vagy igen gyenge**
0,25-0,50	**Gyenge**
0,50-0,75	**Mérsékelten erős vagy erős**
0,75-1,00	**Igen erős**

VIGYÁZAT! A 0,95-nél nagyobb abszolút értékű korrelációs együttható gyanús, azt sugallja, hogy az egyik mért érték a másikból következik, illetve ezáltal determinált.

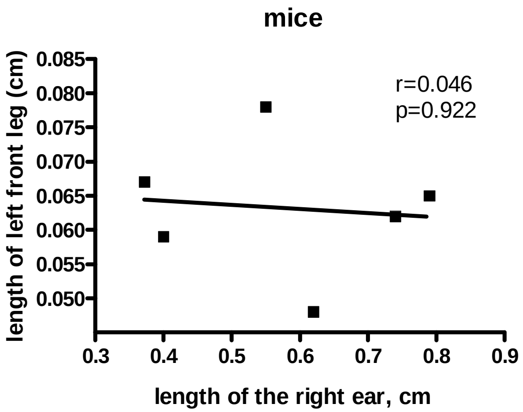
A lineáris (Pearson) korrelációs együttható kiszámíthatóságának feltételei

- A vizsgált **ismérvek változatai** egy nagyobb populációból **véletlenszerűen legyenek kiválasztva**;
- A **megfigyelések**nek **egymástól függetlenek**nek kell lenniük;
- **Sohasem szabad két populációból származó mintát kombinálni,** mert ez ál-szignifikáns korrelációt fog mutatni, noha sem az egyik, sem a másik mintában külön-külön nincs kapcsolat a két változó között;
- **Mind az x, mind az y mintáknak normál eloszlást mutató populációból kell származniuk**;
  - Ha ez nem áll fenn, akkor nemparaméteres eljárást (Spearman korrelációs együttható) kell végeznünk.

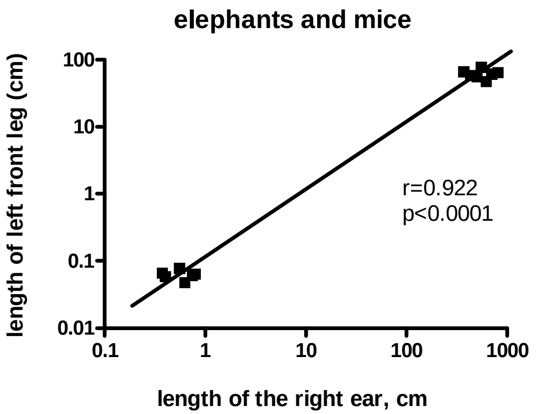
HOGYAN NEM SZABAD KORRELÁCIÓT SZÁMÍTANI?



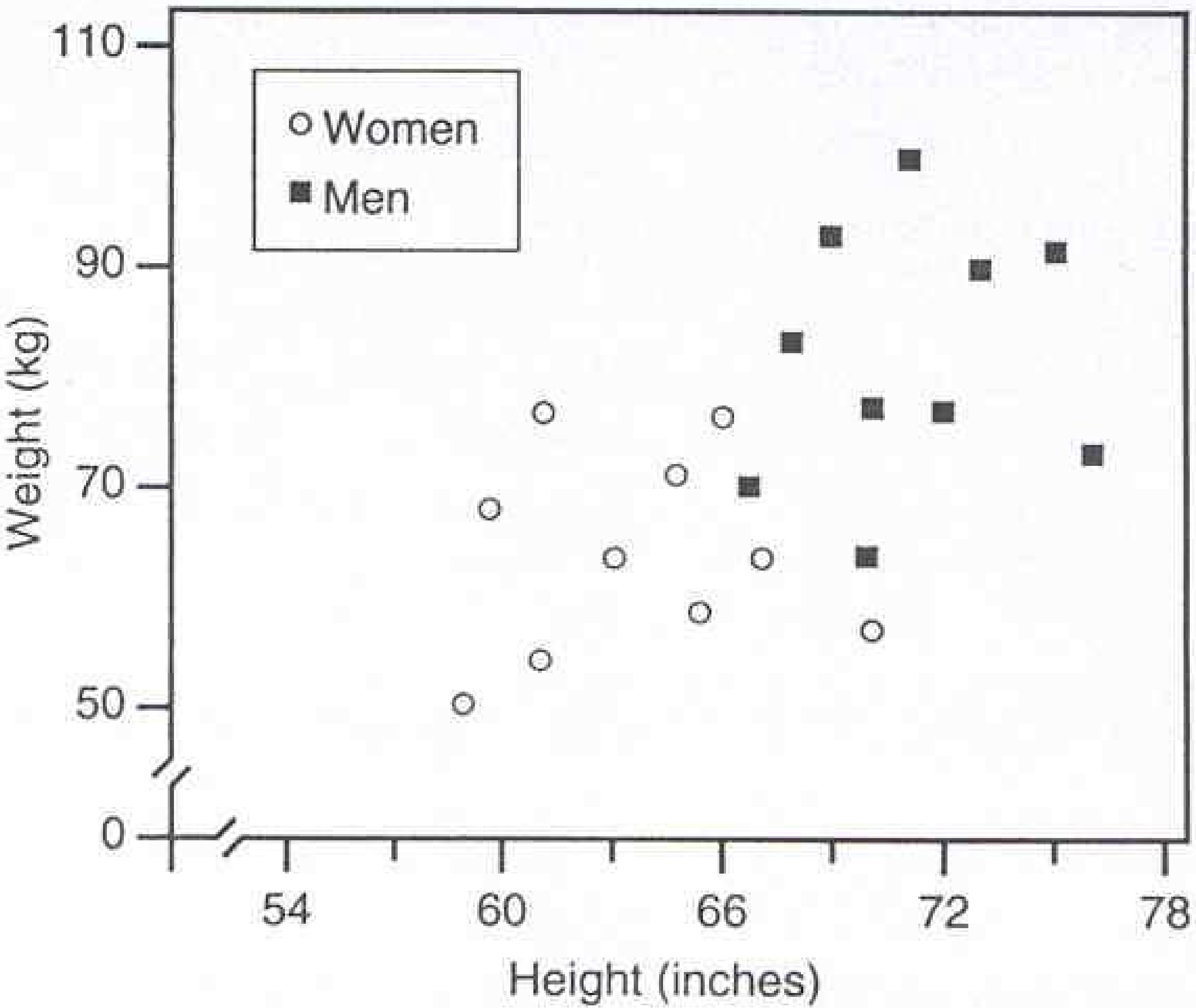
(L. E. Phant et al.: Big Animals, 2004;25:23-45)



(B. Hamster, P. Rat: Small Animals 1998;234:56-78)

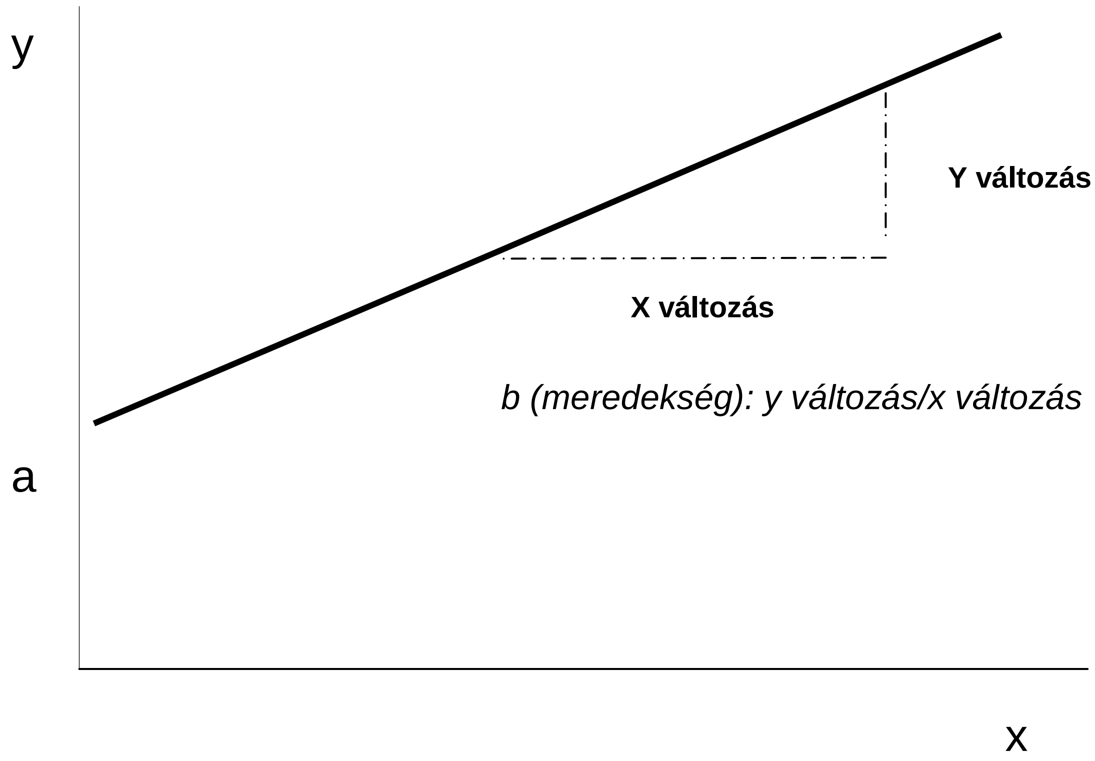


(G. Swine et al., unpublished)


110
90
70
50
0
Weight (kg)
Women
Men
54
60
66
72
78
Height (inches)

A REGRESSZIÓ

- A regresszió úgy mutatja meg két változó kapcsolatát, hogy egyúttal az egyik változónak (függő változó) a másik változótól (független változó) való függése mértékét is kifejezi;
- lineáris és nem-lineáris regresszió;
- egyszerű és többszörös regresszió;

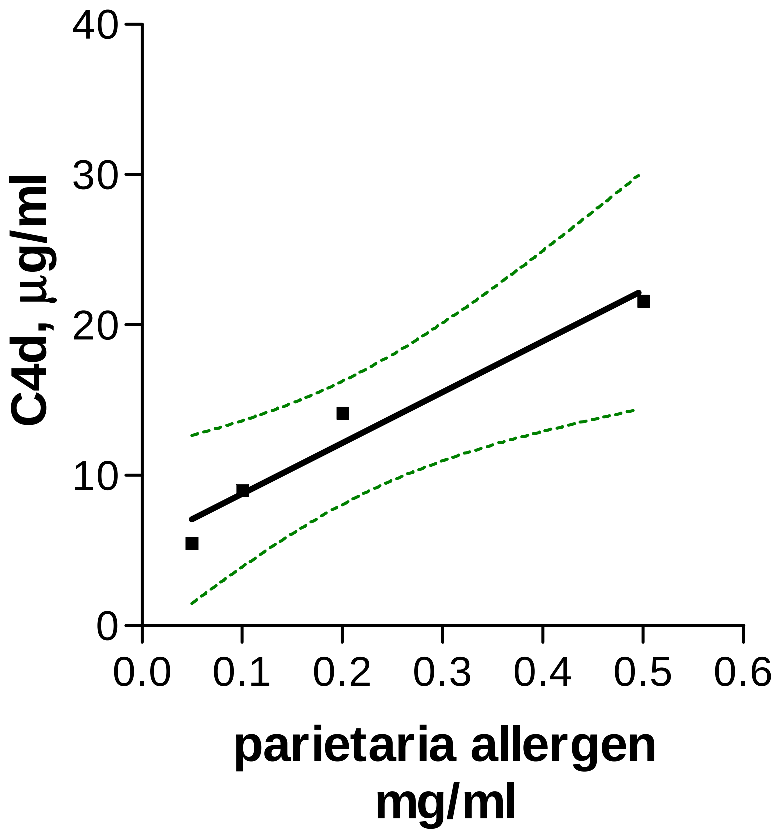

y
Y változás
X változás
b (meredekség): y változás/x változás
a
x

- A **lineáris regressziós számítás lényege** az, hogy egy **olyan egyenest húzunk, amely a mérési pontoktól a lehető legkisebb távolságban van**, ezeket a legjobban megközelíti (best fit regression line, azaz a **legkisebb négyzetek módszere**).

- A pontok és egyenes közötti függőleges távolságokat **reziduumok**nak (residual) nevezzük, ezek négyzetének összege a **reziduumok varianciája**, melynek négyzetgyöke a **reziduumok szórása**.

- A **regressziós egyenes** az az egyenes, amelynél **a reziduumok szórása a legkisebb**.

Meghatározható egy regressziós egyenes konfidencia intervalluma is.



A regressziós egyenes egyenlete

$$y = ax + b$$

$$a = r \cdot \frac{\sigma_y}{\sigma_x}$$

$$b = \overline{y} - a \cdot \overline{x}$$

A trendegyüttható szignifikánsan eltér-e 0-tól?

Ehhez a kapott trendegyütthatót a mintaelemek standard hibájával kell osztani, mely a *t*-próba elvégzéséhez szükséges *A* próbastatisztika értékét adja meg az alábbi képlet szerint:

Standard hiba, SE (átlag szórása):

$$SE = \frac{s}{\sqrt{n}}$$

$$A = \frac{a \cdot \sqrt{(n-2) \cdot \sum_{i=1}^{n} (x_i - \overline{X})^2}}{\sqrt{\sum_{i=1}^{n} (y_i - \overline{Y})^2}}$$

s = a minta szórása;
n = a minta elemszáma;

Az elemzés során az *A* próbastatisztika értékét hasonlítjuk össze az *n*–*2* (*n* = mintaelemszám) szabadsági fokú *t*-eloszlás kritikus értékével. Általában a 0,95 valószínűségi szintet választjuk.

Ha $A > \left| t \right|_{0,95}^{n-2}$ $\Rightarrow$ *a* (azaz a lineáris trendegyüttható) szignifikánsan eltér 0-tól;

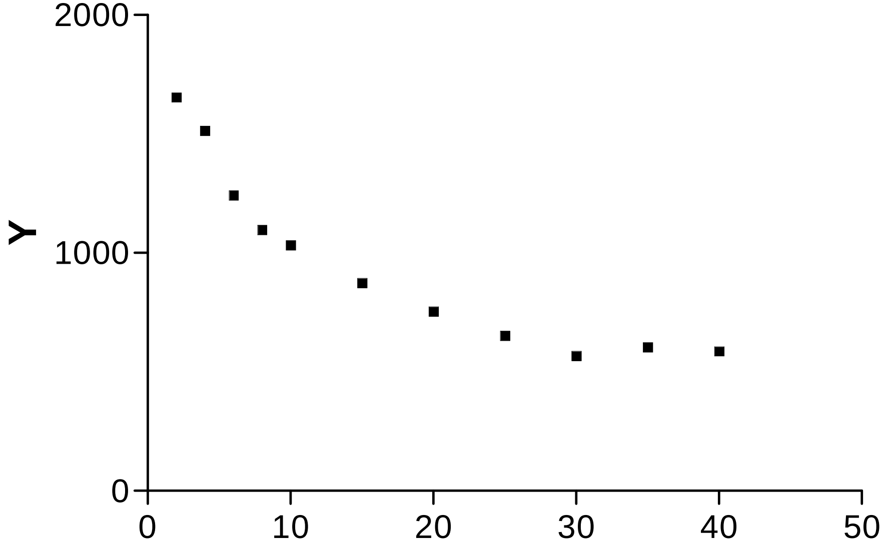
A Student (*t*) eloszlás táblázata

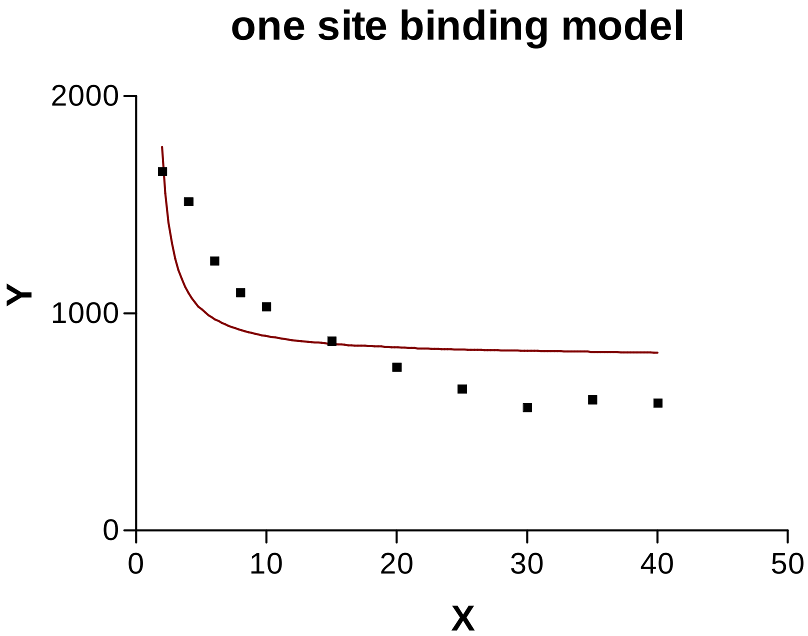
Sz.f.	0,55	0,6	0,7	0,75	0,8	0,9	0,95	0,975	0,99	0,995
1	0,158	0,325	0,727	1,000	1,376	3,078	6,314	12,706	31,821	63,656
2	0,142	0,289	0,617	0,816	1,061	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,584	0,765	0,978	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,569	0,741	0,941	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,559	0,727	0,920	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,553	0,718	0,906	1,440	1,943	2,447	3,143	3,707
7	0,130	0,263	0,549	0,711	0,896	1,415	1,895	2,365	2,998	3,499
8	0,130	0,262	0,546	0,706	0,889	1,397	1,860	2,306	2,896	3,355
9	0,129	0,261	0,543	0,703	0,883	1,383	1,833	2,262	2,821	3,250
10	0,129	0,260	0,542	0,700	0,879	1,372	1,812	2,228	2,764	3,169
11	0,129	0,260	0,540	0,697	0,876	1,363	1,796	2,201	2,718	3,106
12	0,128	0,259	0,539	0,695	0,873	1,356	1,782	2,179	2,681	3,055
13	0,128	0,259	0,538	0,694	0,870	1,350	1,771	2,160	2,650	3,012
14	0,128	0,258	0,537	0,692	0,868	1,345	1,761	2,145	2,624	2,977
15	0,128	0,258	0,536	0,691	0,866	1,341	1,753	2,131	2,602	2,947
16	0,128	0,258	0,535	0,690	0,865	1,337	1,746	2,120	2,583	2,921
17	0,128	0,257	0,534	0,689	0,863	1,333	1,740	2,110	2,567	2,898
18	0,127	0,257	0,534	0,688	0,862	1,330	1,734	2,101	2,552	2,878
19	0,127	0,257	0,533	0,688	0,861	1,328	1,729	2,093	2,539	2,861
20	0,127	0,257	0,533	0,687	0,860	1,325	1,725	2,086	2,528	2,845
21	0,127	0,257	0,532	0,686	0,859	1,323	1,721	2,080	2,518	2,831
22	0,127	0,256	0,532	0,686	0,858	1,321	1,717	2,074	2,508	2,819
23	0,127	0,256	0,532	0,685	0,858	1,319	1,714	2,069	2,500	2,807
24	0,127	0,256	0,531	0,685	0,857	1,318	1,711	2,064	2,492	2,797
25	0,127	0,256	0,531	0,684	0,856	1,316	1,708	2,060	2,485	2,787
26	0,127	0,256	0,531	0,684	0,856	1,315	1,706	2,056	2,479	2,779
27	0,127	0,256	0,531	0,684	0,855	1,314	1,703	2,052	2,473	2,771
28	0,127	0,256	0,530	0,683	0,855	1,313	1,701	2,048	2,467	2,763
29	0,127	0,256	0,530	0,683	0,854	1,311	1,699	2,045	2,462	2,756
30	0,127	0,256	0,530	0,683	0,854	1,310	1,697	2,042	2,457	2,750
40	0,126	0,255	0,529	0,681	0,851	1,303	1,684	2,021	2,423	2,704
60	0,126	0,254	0,527	0,679	0,848	1,296	1,671	2,000	2,390	2,660
100	0,126	0,254	0,526	0,677	0,845	1,290	1,660	1,984	2,364	2,626
120	0,126	0,254	0,526	0,677	0,845	1,289	1,658	1,980	2,358	2,617
50000	0,126	0,253	0,524	0,674	0,842	1,282	1,645	1,960	2,326	2,576

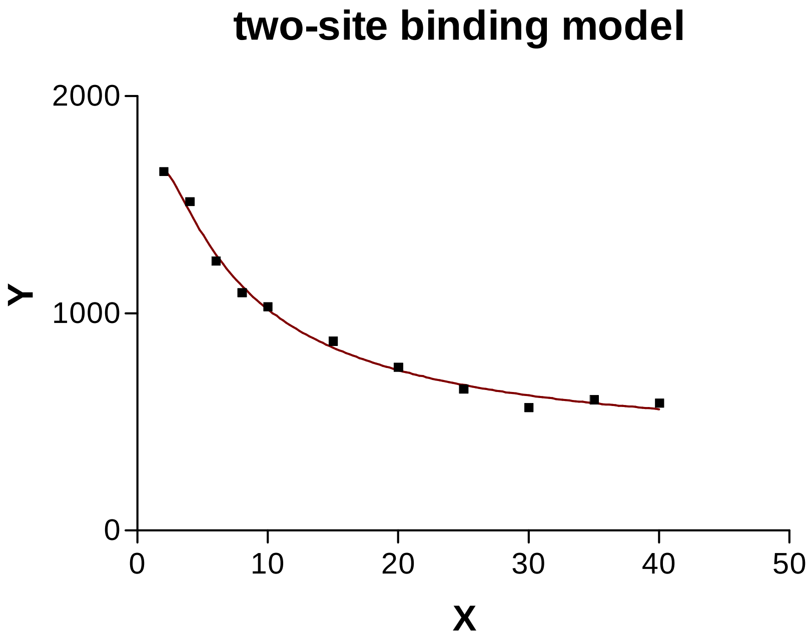
Mi tegyünk, ha az x és az y közötti összefüggés nemlineáris?

- 1. Meg kell próbálni úgy transzformálni az értékeket, hogy lineárissá váljon az összefüggés;
- 2. Ha ez nem lehetséges, a nemlineáris regresszióval kell dolgozni;

x	y
2.00	1654.00
4.00	1515.00
6.00	1243.00
8.00	1098.00
10.00	1032.00
15.00	874.00
20.00	754.00
25.00	653.00
30.00	567.00
35.00	604.00
40.00	587.00




one site binding model
2000
1000
0
Y
0
10
20
30
40
50
X


two-site binding model
2000
1000
0
Y
0
10
20
30
40
50
X

Spearman-féle rangkorrelációs együttható (r_s)

- A 20. sz. eleje óta ismert, ezt alkalmazzák leggyakrabban;
- Az egyik, vagy mindkét változó ordinális változó (pl. az alma íze és színezettsége közötti összefüggés);
- Értelmezési tartománya: [-1, +1];
    - Ha $r_s = 1$, $\Rightarrow$ a két rangsor ugyanaz;
    - ha $r_s = 0 \Rightarrow$ a két rangsor egymástól független, és
    - ha $r_s = -1 \Rightarrow$ a két rangsor egymás megfordításai.
- Jele: r_s ;
- Képlete:

$$r_s = 1 - \left[6 \sum_{i=1}^{n} \frac{d_i^2}{n(n^2 - 1)} \right]$$

- ✓ Az azonos elemszámú x_i és y_i minták elemeit külön-külön időrendben sorszámozzuk, majd őket a sorszámukkal együtt növekvő sorrendbe rendezzük;
- ✓ vesszük az összetartozó értékpárok rangszámainak (sorszámainak) az eltérés-négyzetösszegeit:

$$\sum_{i} (\mathrm{rang}(x_i) - \mathrm{rang}(y_i))^2 = \sum_{i} d_i^2$$

- ✓ d_i: a két minta összetartozó rangszámainak a különbségei;
- ✓ n: az elempárok száma;

10 dolgozatot két bíráló rangsorolt

Bírálók

A	1	2	3	4	5	6	7	8	9	10
B	2	3	1	4	6	5	8	9	7	10
különbség (d)	-1	-1	2	0	-1	1	-1	-1	2	0
d^2	1	1	4	0	1	1	1	1	4	0

$$r_s = 1 - \frac{6 \times 1}{1 \qquad -1} = 1 - \frac{8}{9} = 0,915$$

Nagyfokú hasonlóságot mutat a két bíráló rangsorolása

(-1: teljes ellentét, +1 teljes megegyezés esetén)

Rangkorreláció

Egy régió vállalatainak gazdálkodására vonatkozó adatok

Régió	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Árbevétel (MFt)	34	30	25	22	21	10	12	8	31	20
Nyereség (MFt)	16	10,5	10	12	7	4	2	1	9	11

x	10	8	7	6	5	2	3	1	9	4
y	10	7	6	9	4	3	2	1	5	8
d	0	1	-1	-3	1	-1	1	0	4	-4
d^2	0	1	1	9	1	1	1	0	16	15

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6 \cdot 45}{10(10^2-1)} = 0{,}7273$$

Értelmezés: a vállalatok árbevétele és nyeresége között közepesnél szorosabb, pozitív irányú kapcsolat van.

Összefüggés az almavirágok átmérője és az almák súlya között. Példa a rangkorrelációs eljárás elvének szemléltetésére.

Virág-alma párok sorszáma	Virág átmérője, mm	Rangszám	Alma súlya, g	Rangszám	Rangszám különbség
1	32	3,5	210	4	0,5
2	18	1	150	1	0
3	36	5	235	6	1
4	32	3,5	205	3	0,5
5	39	7	220	5	2
6	37	6	256	7	1
7	30	2	190	2	0
8	42	8	300	8	0
Spearman korrelációs koefficiens				r=0,9222	

Szignifikáns-e ez a korrelációs együttható?

A korelációs együttható szignifikánsan eltér-e 0-tól?

- ✓ a korrelációs együttható próbastatisztikája: Student-eloszlású, *n-2* szabadságfokkal;
- ✓ *t*-próbát végezhetünk H_0 elfogadására vagy elvetésére;

$$\boxed{\text{A } H_0: r = r_0 \text{ hipotézis vizsgálata}}$$

H_0: $r = 0$ esetén:

$$t_{(r)} = r\sqrt{\frac{n-2}{1-r^2}}$$ ⇒ a kritikus r-érték: $$r = \frac{t}{\sqrt{n-2+t^2}}$$

A transzformáció az *r* eloszlását *t*-eloszlássá alakítja

⇒ ha $r > |t|_{0,95}^{n-2}$ ⇒ *r* (azaz a lineáris korrelációs együttható szignifikánsan eltér 0-tól;

A Student (*t*) eloszlás táblázata

Sz.f.	0,55	0,6	0,7	0,75	0,8	0,9	0,95	0,975	0,99	0,995
1	0,158	0,325	0,727	1,000	1,376	3,078	6,314	12,706	31,821	63,656
2	0,142	0,289	0,617	0,816	1,061	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,584	0,765	0,978	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,569	0,741	0,941	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,559	0,727	0,920	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,553	0,718	0,906	1,440	1,943	2,447	3,143	3,707
7	0,130	0,263	0,549	0,711	0,896	1,415	1,895	2,365	2,998	3,499
8	0,130	0,262	0,546	0,706	0,889	1,397	1,860	2,306	2,896	3,355
9	0,129	0,261	0,543	0,703	0,883	1,383	1,833	2,262	2,821	3,250
10	0,129	0,260	0,542	0,700	0,879	1,372	1,812	2,228	2,764	3,169
11	0,129	0,260	0,540	0,697	0,876	1,363	1,796	2,201	2,718	3,106
12	0,128	0,259	0,539	0,695	0,873	1,356	1,782	2,179	2,681	3,055
13	0,128	0,259	0,538	0,694	0,870	1,350	1,771	2,160	2,650	3,012
14	0,128	0,258	0,537	0,692	0,868	1,345	1,761	2,145	2,624	2,977
15	0,128	0,258	0,536	0,691	0,866	1,341	1,753	2,131	2,602	2,947
16	0,128	0,258	0,535	0,690	0,865	1,337	1,746	2,120	2,583	2,921
17	0,128	0,257	0,534	0,689	0,863	1,333	1,740	2,110	2,567	2,898
18	0,127	0,257	0,534	0,688	0,862	1,330	1,734	2,101	2,552	2,878
19	0,127	0,257	0,533	0,688	0,861	1,328	1,729	2,093	2,539	2,861
20	0,127	0,257	0,533	0,687	0,860	1,325	1,725	2,086	2,528	2,845
21	0,127	0,257	0,532	0,686	0,859	1,323	1,721	2,080	2,518	2,831
22	0,127	0,256	0,532	0,686	0,858	1,321	1,717	2,074	2,508	2,819
23	0,127	0,256	0,532	0,685	0,858	1,319	1,714	2,069	2,500	2,807
24	0,127	0,256	0,531	0,685	0,857	1,318	1,711	2,064	2,492	2,797
25	0,127	0,256	0,531	0,684	0,856	1,316	1,708	2,060	2,485	2,787
26	0,127	0,256	0,531	0,684	0,856	1,315	1,706	2,056	2,479	2,779
27	0,127	0,256	0,531	0,684	0,855	1,314	1,703	2,052	2,473	2,771
28	0,127	0,256	0,530	0,683	0,855	1,313	1,701	2,048	2,467	2,763
29	0,127	0,256	0,530	0,683	0,854	1,311	1,699	2,045	2,462	2,756
30	0,127	0,256	0,530	0,683	0,854	1,310	1,697	2,042	2,457	2,750
40	0,126	0,255	0,529	0,681	0,851	1,303	1,684	2,021	2,423	2,704
60	0,126	0,254	0,527	0,679	0,848	1,296	1,671	2,000	2,390	2,660
100	0,126	0,254	0,526	0,677	0,845	1,290	1,660	1,984	2,364	2,626
120	0,126	0,254	0,526	0,677	0,845	1,289	1,658	1,980	2,358	2,617
50000	0,126	0,253	0,524	0,674	0,842	1,282	1,645	1,960	2,326	2,576

$$r = \frac{t}{\sqrt{n-2+t^2}}$$

Nézzük mindig a dolgok napos oldalát!

Mára befejeztük, viszontlátásra!