



**STATOKOS JEGYZET: TÖBBVÁLTOZÓS STATISZTIKA:
KORRELÁCIÓ ÉS REGRESSZIÓ**

StatOkos – Statisztikai és Módszertani Adatbázis

2018

TÖBBVÁLTOZÓS STATISZTIKA: KORRELÁCIÓ ÉS REGRESSZIÓ

szerzők: Kazinczi Csaba, Alter Emese

A korreláció

- Változók kapcsolatának vizsgálatára alkalmas olyan esetekben, amikor nincs szilárd előfeltételezésünk a kapcsolatra vonatkozóan
- Nem alkalmas ok-okozati kapcsolat feltárására, csupán az együttjárás meglétét, erősségét és irányát tudjuk ellenőrizni vele
- Metrikus változók esetén Pearson-korrelációt, ordinális változók esetén Spearman-féle (nemparaméteres) korrelációt alkalmazhatunk
- A korrelációs együttható (r) értéke -1 és 1 közé eshet. Minél közelebb esik r abszolút értéke 1 -hez, annál erősebb együttjárásról beszélhetünk. Ha $r = 1$, egyenes arányosságról, ha $r = -1$, fordított arányosságról beszélhetünk. Ha r negatív, fordított kapcsolatról beszélhetünk, vagyis minél kisebb x értéke, annál nagyobb y értéke, és fordítva. Pozitív irányú együttjárás esetén minél magasabb x értéke, annál magasabb y is, illetve minél alacsonyabb a egyik érték, annál alacsonyabb a másik is.
- Jelentős együttjárásról csak szignifikáns eredmény esetén beszélhetünk.
- Hivatkozása: $r(n\text{-elemszám}) = r$ értéke, $p =$ szignifikancia érték.

A lineáris regresszió

- Olyan esetekben alkalmazzuk változók kapcsolatának vizsgálatára, amikor a kapcsolatról szilárd előfeltételezésünk van.
- Nem csak a kapcsolat meglétét, de irányát is vizsgálhatjuk a segítségével.
- Annak a vizsgálatára használjuk, hogy több paraméteres változó milyen mértékben befolyásol egy szintén paraméteres függő változót. A függő és független változókat mi adjuk meg a próba lefuttatása során, tehát előfeltételezésünknek kell lennie a kapcsolat irányára vonatkozóan.
- Hivatkozása: $R^2 = R$ square értéke, $F(\text{regresszió szabadságfoka, reziduális szabadságfoka}) = F$ értéke, $p =$ szignifikancia értéke.