Asszociációs együtthatók

Statisztika II., 4. alkalom

Asszociációs mérőszámok ordinális változók esetén

Monotonitási együtthatókat is lehet használni ordinális változók esetén.

Monotonitás szempontjából megkülönböztetünk:
- -konkordáns
- -diszkordáns
- -kapcsolt párokat

Konkordáns (egyirányú) pár (P): Az értékek sorrendjét tekintve azonos irányban térnek el a két változóra vonatkozóan.

	Elolvasom a hirdetéseket	Megnézem a tv reklámokat
Személy 1	néha	gyakran
Személy 2	soha	néha

Diszkordáns (fordított) pár (Q): Az értékek sorrendjét tekintve ellentétes irányban térnek el a két változóra vonatkozóan.

	Elolvasom a hirdetéseket	Megnézem a tv reklámokat
Személy 1	néha	gyakran
Személy 3	gyakran	néha

Kapcsolt pár (T:lehet Tx vagy Ty): Valamely változót tekintve azonos értékkel rendelkező pár

	Elolvasom a hirdetéseket	Megnézem a tv reklámokat
Személy 1	néha	gyakran
Személy 4	néha	soha

Monotonitási mérőszámok

A P-Q és a P+Q arányát vizsgálják
és a mérőszámok lehetséges értékeit a -1 és 1 tartományba igyekszenek beszorítani.

Monotonitási mérőszámok:
-Goodman-Kruskal féle gamma
-Somers féle D
-Kendal féle tau, tau b, tau c

Goodman-Kruskal féle gamma

$$\Gamma = \frac{P - Q}{P + Q}$$

Azt fejezi ki, hogy mennyivel nagyobb a konkordancia valószínűsége a diszkordanciánál, ha nem lép fel értékegyezés.

A kapcsolt párokat nem veszi figyelembe. Értéke -1 és 1 közé esik.

Gamma csak akkor lehet 1, ha nincsennek diszkordáns párok és gamma csak akkor lehet -1, ha nincsennek konkondárs párok.

Függetlenség esetén értéke nulla, a nulla érték azonban csak 2x2-es kontingencia táblázat esetén jelent függetlenséget.

Somers féle D

$$D_{(Y|X)} = \frac{P-Q}{1-T_x} = \frac{P-Q}{P+Q+T_Y}$$

Tx annak az aránya, hogy az x értékek megegyeznek.
Ty annak az aránya, hogy az y értékek megegyeznek.

A Sommers féle D statisztika azt fejezi ki, hogy mennyivel nagyobb a konkordancia aránya a diszkordanciánál, ha a X értékei nem egyezőek.

Az egyik változót függő változóként kezeli (Y): a Sommers féle D egy aszimmetrikus mérőszám. Értéke csak akkor lesz 1 ha Y nagyságrendi viszonyai minden esetben követik X nagyságrendi viszonyait, és csak akkor lesz értéke -1, ha minden esetben ellentétes a nagyságrendi viszony.

Bármelyik változót tekinthetjük a függő változónak: $D_{(X|Y)} = \frac{P-Q}{1-T_Y} = \frac{P-Q}{P+Q+T_X}$

Továbbá, van szimmetrikus változata is a statisztikának: $D_{(sym)} = \dfrac{P-Q}{P+Q+\dfrac{T_Y+T_x}{2}}$

Kendal féle tau, tau b, tau c

$$\tau = \frac{P-Q}{\frac{N(N-1)}{2}} = \frac{2(P-Q)}{N(N-1)}$$

A Kendall féle tau azt fejezi ki, hogy mennyivel nagyobb a konkordáns párok valószínűsége a
diszkordáns párok valószínűségéhez képest, ha az összes párt figyelembe vesszük.

$$\tau_b = \frac{P-Q}{\sqrt{(P+Q+T_Y)(P+Q+T_X)}} = \sqrt{D_{X|Y} D_{Y|X}}$$

A Kendall féle tau b a két aszimmetrikus Sommers féle D mértani közepével egyenlő, azt méri, hogy átlagosan milyen mértékű a monoton függése a két változónak egymástól. Csak akkor lehet értéke 1 vagy -1 ha a két változó lehetséges értékeinek száma egyenlő.

$$\tau_c = \frac{2k(P-Q)}{N^2(k-1)}$$

k a két változó értékeinek száma közül a kisebbik. Bármely lehetséges értékek esetén lehet értéke 1 vagy -1.

Spearman féle rangkorreláció

A rangsorolásos eljárások közé tartozik, ahol a számítás nem az értékekkel történik, hanem a sorszámmokkal.

Ha korrelációt szeretnénk számolni, de a normalitás feltétele nem teljesül, akkor a Pearson féle r érték torzított eredményt adhat. Ilyen esetben használhatunk rangsorolásos eljárást, mely nem érzékeny a normalitás feltétel teljesülésére. Az értékeket nagyságrendben besoroljuk, amihez csak az szükséges, hogy a változó legalább ordinális típusú legyen.

Két N elemű változó esetén az értékekből rangsorokat képezünk.

X->R

Y->S

Ezután a Pearson féle r képletét a rangszámokra alkalmazzuk.

$$r_s = \frac{\sum_i (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_i (R_i - \bar{R})^2 \sum_i (S_i - \bar{S})^2}}$$