

ROLP

Povezanost podataka

Ksenija Baždarić

akad. god. 2016./2017.



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Povezanost (korelacija)

- matematički postupak za izračunavanje stupnja povezanosti dvaju brojčanih pokazatelja
- brojčano se iskazuje koeficijentom povezanosti "r" čije se vrijednosti kreću u rasponu od -1 do +1 (stupanj povezanosti).
- s ciljem utvrđivanja značajnosti koeficijenta izračunava se značajnost korelacije (P).



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Uvjeti za izračunavanje korelacije

- Brojčana mjerna ljestvica
 - nominalna ☹
 - ordinalna ☺
 - intervalna ☺
 - omjerna ☺



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Pojmovnik - vrste povezanosti

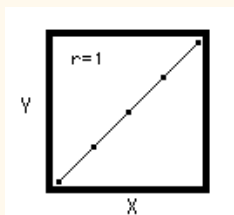
- potpuna povezanost
- pozitivna povezanost
- negativna povezanost
- nelinearna povezanost



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Potpuna povezanost (korelacija)

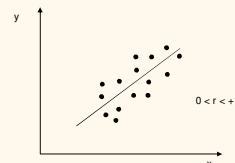
- u prirodnim sustavima ne postoji



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Pozitivna povezanost

- s porastom jedne varijable raste i vrijednost druge varijable



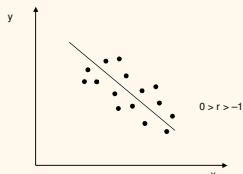
- Primjer: duljina trajanja šećerne bolesti i stupanj oštećenja kapilara u oku



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Negativna povezanost

- s porastom jedne varijable padaju vrijednosti druge varijable ili obratno



- Primjer: porast nadmorske visine i koncentracija kisika u zraku

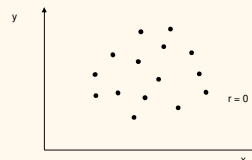


Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Nepostojanje povezanosti

- vrijednostima jedne varijable ne mogu se pridružiti vrijednosti druge varijable – nema linearne povezanosti



- Primjer: povezanost veličine zjenice oka i koncentracije kalcijevih iona u krvi

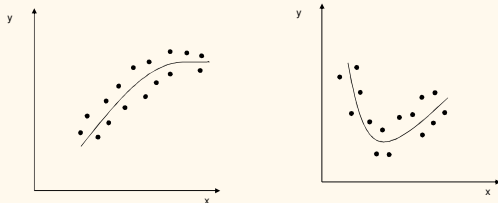


Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Vrste povezanosti

- Nelinearne korelacije



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Vrste koeficijentata povezanosti

- *Pearsonov koeficijent povezanosti*
- *Spearman koeficijent povezanosti (rang korelacija)*



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Pearsonov koeficijent korelacije

Uvjeti za izračunavanje:

- jedna od promatranih varijabli ima normalnu raspodjelu (podatci simetrični)
- veći uzorak ($N > 35$)
- povezanost je linearna (točkasti grafikon)
- podatci na intervalnoj ili omjernoj ljestvici



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Spearmanov koeficijent korelacije

Kada izračunavamo?

- podatci na ordinalnoj ljestvici
- kada je raspodjela varijabli odstupajuća
- mali uzorak ($N < 35$)
- prisutnost podataka koji značajno odskakuju od ostalih rezultata (engl. *outliers*).

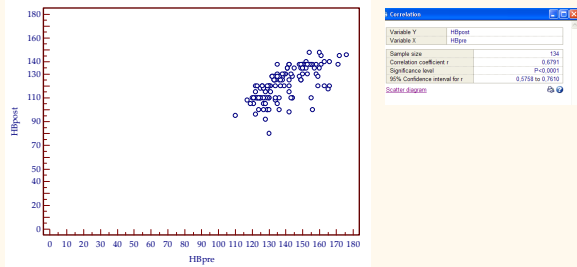
U slučaju rezultata $r_s = 0$ možemo zaključiti da varijable nisu povezane.



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Točkasti grafikon (engl. *scatter diagram*)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Značajnost povezanosti (P)

- Uz koeficijent korelacije treba naznačiti i njegovu statističku značajnost $\rightarrow P$
- Tumačiti se smiju samo statistički značajni koeficijenti korelacije (Colton, 1974.)

r	POVEZANOST
0 do 0,25 0 do -0,25	nema povezanosti
0,25 do 0,50 -0,25 do -0,50	slaba povezanost
0,50 do 0,75 -0,50 do -0,75	umjerena do dobra povezanost
0,75 do 1 -0,75 do -1	vrlo dobra do izvrsna povezanost



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Povezanost i uzročno-posljedična veza

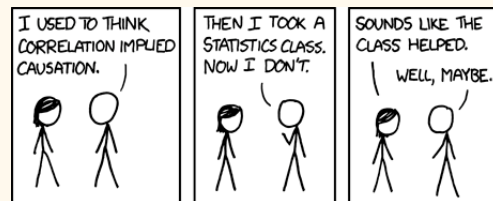
- Povezanost \neq uzročnost
- Primjer: određivanje koncentraciju glukoze s pomoću dva biokemijska analizatora $\rightarrow r=0,97$; $P<0,001 \rightarrow$ je li jedan od tih analizatora djelovao na drugi?
- Uzročno-posljedična veza dviju pojava može se dokazati jedino pokusom.



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Povezanost i uzročno-posljedična veza



<http://xkcd.com/552/>



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

