

Medicinska informatika i biostatistika

Uvod u biostatistiku

prof. dr. sc. Mladen Petrovečki
dr. sc. Lidija Bilić-Zulle
Vanja Pupovac, prof.

Katedra za medicinsku informatiku
Medicinski fakultet Sveučilišta u Rijeci



Rijeka University School of Medicine Department of Medical Informatics

Podsjetnik



Rijeka University School of Medicine Department of Medical Informatics



Nastava

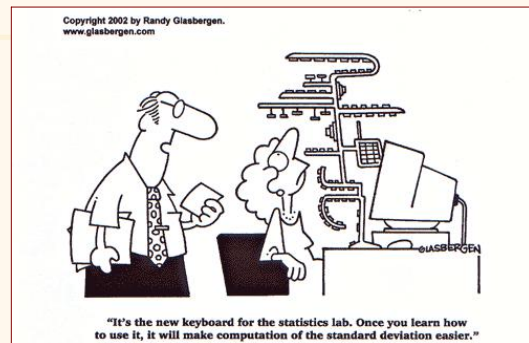
- Biostatistika
 - Predavanja -- 3
 - Seminari -- 2
 - Vježbe -- 5

Raspored na web stranicama Katedre

<http://mi.medri.hr>



Rijeka University School of Medicine Department of Medical Informatics



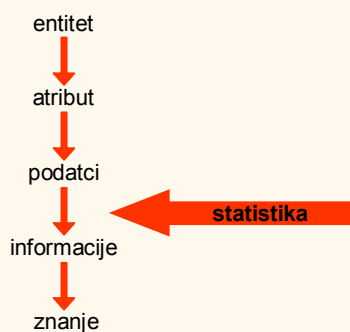
Rijeka University School of Medicine Department of Medical Informatics



Kako do znanja? Što je statistika?

Obradba podataka radi jasnijeg prikazivanja.

Sažimanje velikog broja podataka u manji.



Rijeka University School of Medicine Department of Medical Informatics



Teme

- temeljni statistički pojmovi
- mjerenje i prikaz podataka
- obradba podataka i vrste statističkih testova
 - obradba kvalitativnih podataka
 - obradba brojčanih podataka
 - opća pravila, neparametrijski testovi, parametrijski testovi, složene analize
- vrste istraživanja

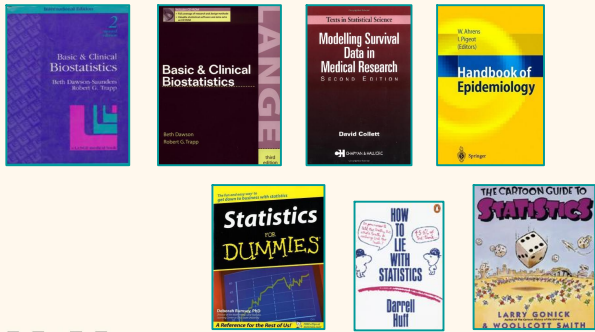
nastava: interaktivna
aktivno sudjelovanje, rasprava



Rijeka University School of Medicine Department of Medical Informatics



Literatura



Rijeka University School of Medicine Department of Medical Informatics



Literatura

Znanost je sumnja.
(N. Alegretti)



Rijeka University School of Medicine Department of Medical Informatics



Znanost

Znanost je čežnja čovječanstva
k istinskom znanju.



Rijeka University School of Medicine Department of Medical Informatics



Pregled

- metodologija znanstvenog istraživanja
- postupci istraživanja
- neznanstveni i znanstveni postupci



Rijeka University School of Medicine Department of Medical Informatics



Metodologija znanstvenog
istraživanja

- znanost kao način razmišljanja
- znanstveno djelovanje
 - privremeno tumačenje
 - hipoteza
 - zaključak
- provjerljivost
- postavke → činjenice → znanje



TVRDNJA
OBRAZLOŽENJE
PROVJERA



Rijeka University School of Medicine Department of Medical Informatics



Neznanstveni postupci istraživanja

- ustrajnost
(navika, stav, vjerovanje, inercija)
- autoritet
- intuicija (očiglednost)



Rijeka University School of Medicine Department of Medical Informatics



Predrasude

Laž, velika laž, statistika.

Ukoliko želite nešto statistički dokazati potrebno je samo iznaći test s kojim ćete postići željeni rezultat



Kako promatramo svijet?

- nesreća... sreća...
- svaki puta... desetak puta...
- rijetko... često...
- nemoguće... nevjerovatno...



Vjerojatnost

- modeli vjerojatnosti:

- deterministički
 - predvidivi događaji



- probabilistički
 - slučajni događaji



Probabilistički model vjerojatnosti

- tri sastavnice:
 - ishod mjerenja
 - ostali događaji koji nastaju (njihova brojnost i ovisnost)
 - vrijednost vjerojatnosti P (*probability*)

$$0 < P < 1$$



Izračun vjerojatnosti

- vjerojatnost događaja P
- broj povoljnih mogućnosti
$$P = \frac{\text{broj povoljnih mogućnosti}}{\text{ukupni broj mogućnosti}}$$
- vjerojatnost događaja:
 - vrijednost u rasponu 0-1
 - 0 – vjerojatnost nemogućeg događaja
 - 1 – vjerojatnost sigurnog događaja



Znanstvena hipoteza

- Uvod; Poglavlje 3, str 27.

- Značajke dobre hipoteze
 - svrhovita
 - provjerljiva
 - suglasna s općim znanjem
 - jednostavna



Statistička hipoteza

- elementarna, jednoznačna i eksplicitna tvrdnja
- točna (istinita) ili netočna (neistinita)
- proizlazi iz znanstvene hipoteze
- statističke hipoteze dokazujemo (testiramo)
- provjera hipoteze → traženje istine
- istina → stvarno, objektivno stanje
- probabilistički sustav:
istina → vjerojatnost
- značajno → **ono što se ostvaruje na svaki drugi način osim slučajno**



Statistička hipoteza

- postupak dokazivanja:
 1. postavljanje hipoteze
 2. biranje statističkog testa
 3. određivanje razine značajnosti
 4. izračunavanje
 5. zaključivanje



1. Postavljanje hipoteze

- nulta – H_0 -hipoteza: razlike nema!
- suprotna – H_1 -hipoteza: razlike ima!
- samo jedna može biti **istinita**
- samo jedna može biti **prihvaćena**, dok će ona druga biti **odbačena**



2. Odabir testa

- testovi dokazivanja statističkih hipoteza
- ovisnost:
 - osobine obilježja - mjerne ljestvice
 - osobine uzoraka
 - veličina
 - povezanost
 - osobine raspodjele
 - parametrijski
 - neparametrijski
 - broj obilježja - uni/bi/multivarijantni testovi



3. Određivanje razine značajnosti

- *vjerojatnost "P"*
- alfa (α)
 - vjerojatnost odbacivanja H_0 kad je ona stvarno točna i istinita
- tzv. α -pogrješka (pogrješka I. vrste)
- što manja!
- uobičajene vrijednosti
 - npr. $P < 0,05$



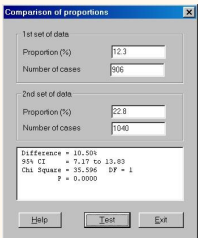
(nastavak)

- zašto upravo to: $P < 0,05$?
POKUS
bacanje novčića: pismo/glava
 $2x$ isto uzastopce = 0,5
 $3x$ = 0,25
 $4x$ = 0,125
 $5x$ = 0,0625
...



4. Izračunavanje

- matematički račun
- računalni statistički programi:
 - MedCalc
 - Statistica
 - SPSS
 - NCSS
 - GraphPad InStat
 - ...



5. Zaključivanje

- mala vrijednost “ P ” = mala vjerojatnost da ne prihvaćama ono što je istinito
- zaključivanje:
 - $P < \alpha$
 - vjerojatnost istinitosti H_0 je mala
 - odbacujemo (ne prihvaćamo) nultu hipotezu
 - prihvaćamo alternativnu, H_1
 - potvrdimo je, iskažemo, uz $P=...$



Pogrješke testiranja hipoteze

PRAVO STANJE		ZAKLJUČENO
RAZLIKA POSTOJI (H_1)	RAZLIKA NE POSTOJI (H_0)	
ISPRAVAN ZAKLJUČAK	α pogreška (I. vrste) (H_0 odbac.)	RAZLIKA POSTOJI
β pogreška (II. vrste)	ISPRAVAN ZAKLJUČAK (H_0 prihvat.)	RAZLIKE NEMA



Planiranje istraživanja

- Rad, red i stega ☺
- Poglavlje 6, str 48.
- Plan postavljanja problema
 - pretraživanje literature
 - definiranje problema
 - određivanje strategije
 - oblikovanje hipoteze
 - utvrđivanje praktičnih problema



(nastavak)

- Plan provedbe istraživanja
 - izbor vrste istraživanja
 - odabir uzorka
 - oblikovanje skupina
 - planiranje postupka mjerenja
 - odabir statističkog testa



Strategija istraživanja

- opis stanja u populaciji (procjena učestalosti bolesti)
- usporedba učinkovitosti dvaju postupaka
- dokazivanje povezanosti između nekog čimbenika i pojave bolesti
- str 66, sl. 6-1: nacrt plana istraživanja



Financijsko planiranje

- mora obuhvatiti sve faze istraživanja
- NE smije biti površno i "od prilike"
- važan dio prijave projekata i programa npr. pri MZOS-u i EU (FP7)
- moraju se navesti svi izvori



Važne napomene prilikom planiranja

Planiranje istraživanja je temeljni preduvjet uspješnog istraživanja.

Ne postoji dobar znanstveni rad koji je planiran tijekom provođenja.

NIŠTA se ne podrazumijeva, sve je potrebno opisati i poduprijeti valjanim dokazima!

Valjano istraživanje jednostavno je i transparentno stručnjaku, ali i laiku.

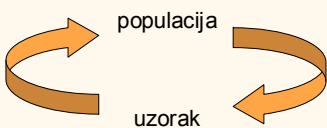
Svako dobro planirano istraživanje može se "nacrtati" tj. prikazati dijagramom toka.



Populacija i uzorak

• Poglavlje 5, str 50

- istraživanje donosi zaključke o populaciji

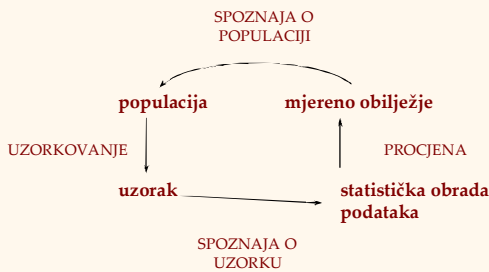


Zašto uzorak, a ne populacija?

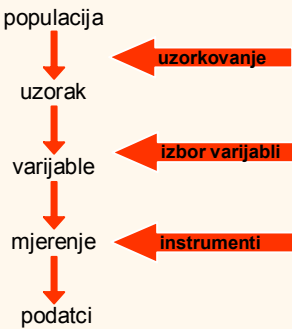
- brzina dobivanja rezultata
- cijena istraživanja
- dostupnost uzoraka i jediniki
- stvarna nemogućnost ispitivanja populacije



Uzorak i populacija



Uzorci i varijable



Uzorak

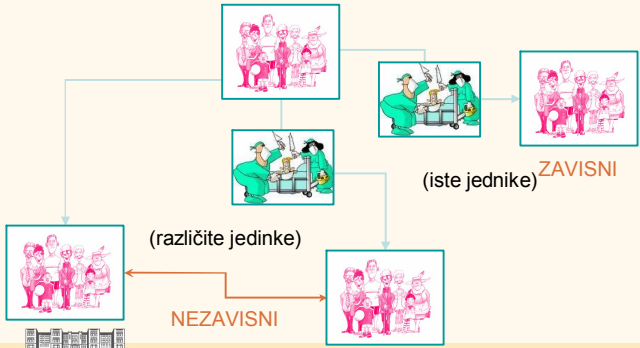
- reprezentativan
- velik
- mjerljiv
- ostvariv
- ekonomičan

• Poglavlje 5, str 54

- **PROBABILISTIČKI - SLUČAJAN**
 - jednostavan
 - sustavan
 - slojevit



Zavisni i nezavisni uzorci



Uzorak

- Što je prigodan uzorak?
- Zašto je prigodan uzorak loš? (reprezentativnost!)
- Kako ipak iskoristiti prigodan uzorak?
 - Ukoliko je reprezentativan (ne razlikuje se od populacije u bitnim značajkama)
- Kada je prigodan uzorak dostatan?
 - iznimke: prethodna istraživanja!



Prikupljanje podataka i mjerenje

- podatci: vrijednost varijabli
- valjanost istraživanja ⇔ izbor varijabli
- točnost zaključaka ⇔ kvaliteta podataka

PLANIRANJE!

