



Doktorski studij  
"Farmaceutске i medicinsko-biokemijske znanosti"  
Kolegij "Biostatistika"

## Statistička obradba podataka: uvodna razmatranja

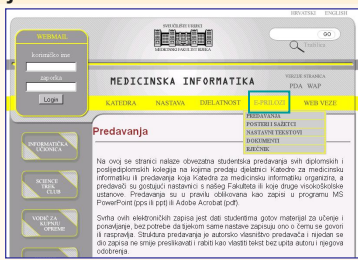
Mladen Petrovečki



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku 2009./10.

## Ovo predavanje...

- PDF
- ≈ 5 MB
- [mi.medri.hr](http://mi.medri.hr) (E-prilozi ⇒ predavanja)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

## Raspored nastave

| pon.     | 14.6. | 15:00 – 16:30 h | 2 | Uvodna razmatranja (S)                                                                |
|----------|-------|-----------------|---|---------------------------------------------------------------------------------------|
| utorak   | 15.6. | 15:00 – 18:30 h | 4 | Priprema podataka za obradbu<br>Osobine raspodjele podataka<br>Vrste istraživanja (S) |
| srijeda  | 16.6. | 15:00 – 18:30 h | 4 | Obradba kvalitativnih podataka                                                        |
| četvrtak | 17.6. | 15:00 – 18:30 h | 4 | Obradba brojčanih podataka                                                            |
| petak    | 18.6. | 15:00 – 18:30 h | 4 | Završna razmatranja (S)<br>Rasprava, pitanja, konzultacije                            |



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

## Raspored

Vježba, tema

1. uvod, *MedCalc*, vrste podataka (mjerne ljestvice), unos podataka, prikaz podataka
2. odnos dvaju brojčanih obilježja, male skupine, velike skupine, odnos više brojčanih obilježja
3. korelacija i regresija, usporedba kvalitativnih podataka, ROC, multivarijatni postupci



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

## Program



- MedCalc  
computer program for statistic  
in biomedicine
- Copyright, 1993-2010  
Medcalc Software  
Mariakerke,  
Belgija
- [www.medcalc.be](http://www.medcalc.be)





Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

## Program – postavljanje:

- otvoriti mapu "C:\MCWin"
- preslikavanje svih podataka sa diskete u zadanu mapu
- otvoriti mapu
- pokrenuti program "Install"
- po završetku pokrenuti program "MedCalc"





Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

## Literatura

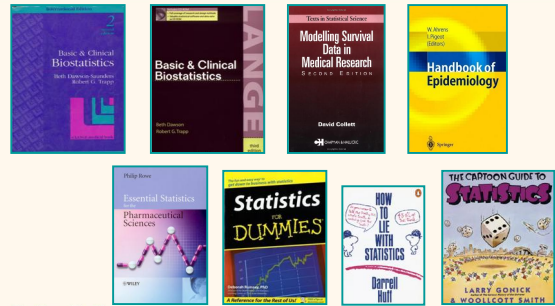
- Dawson-Saunders B, Trapp RG. Basic & Clinical Biostatistics. Prentice-Hall Int. Inc., London, 2001.
- Rowe P. Essential Statistics for the Pharmaceutical Sciences. West Sussex: Wiley; 2007.
- Petrie A, Sabin C. Medical statistics at a glance. Blackwell Science, Oxford, 2000.
- Marušić M. Uvod u znanstveni rad u medicini. III. obnov. izd. Medicinska naklada, Zagreb, 2004.
- Petz B. Osnovne statističke metode za nematematičare. III. dopunjeno izd. Naklada Slap, Jastrebarsko, 1997.
- StatSoft Inc. Electronic Statistics Textbook. Tulsa, OK: StatSoft, 2002. (<http://www.statsoft.com/textbook/stathome.html>)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Literatura



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Literatura

Znanost je sumnja.  
(N. Alegretti)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Ispit

- nema ispitnih razdoblja!
- položiti do 31. prosinca 2011.

- kritička raščlamba statističkih postupaka objavljenog znanstvenog rada
  - pismeno (elektronički ⇒ [mp@kdb.hr](mailto:mp@kdb.hr))
  - poslati desetak dana prije usmene obrane
  - [http://mi.medri.hr/assets/pomocna\\_pitanja\\_za\\_seminar.pdf](http://mi.medri.hr/assets/pomocna_pitanja_za_seminar.pdf) ⇄ ⇄ ⇄
- časopis – citiran u *Current Contents-u*
- rad:
  - cjelokupno istraživanje (ne osvrn na drugi objavljeni rad!)
  - najmanje dva priloga (tablica, slika)
  - jasan opis statističkih postupaka



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Pregled

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Metodologija znan. istraživanja

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

TVRDNJA  
OBRAZLOŽENJE  
PROVJERA



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Neznanstveni postupci istraživanja

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Pokus

- pokus – istraživanje / dio istraživanja
- planiranje pokusa
- nasumični odabir (randomizacija)
- prikrivanje (maskiranje)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Nasumični odabir

- RND (MedCalc)
- funkcija slučajnog broja (uniformna raspodjela)
- "RND(x) returns a computer-generated random number:
  - when  $x \leq 1$  then the result is a number between 0 and 1, or
  - when  $x > 1$  then the result is a number between 1 and x."
- (usp. funkcija RNDNORM)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Prikrivanje

- jednostruko
- dvostruko
- trostruko
- četverostruko



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Vjerojatnost

- modeli vjerojatnosti:
  - deterministički
    - predvidivi događaji
  - probabilistički
    - slučajni događaji



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Vjerojatnost jest...

- izračun matematičke vrijednosti ostvarivanja nekog događaja
- matematički → teorija vjerojatnosti
  - statistika
  - matematika
  - znanstvena metodologija
  - logika i filozofija
- zaključivanje o ostvarivosti događaja



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Probabilistički model vjerojatnosti

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

$$0 < P < 1$$



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## 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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Hipoteza

- znanstvena hipoteza
- statistička hipoteza



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Znanstvena hipoteza

- Uvod; Poglavlje 3, str 19.

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Statistička hipoteza

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Statistička hipoteza

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Testiranje statističke hipoteze

| CD  | Kontrolna skupina  | Ispitanici s PTSD | Statistika |       |
|-----|--------------------|-------------------|------------|-------|
|     |                    |                   | H          | p     |
| CD3 | 39,0 (15,1 – 61,1) | 44 (32,3 – 57,7)  | 0,59       | 0,442 |
| CD4 | 3,5 (2,1 – 5,8)    | 10,3 (6,5 – 17,4) | 15,9       | 0,018 |

Razlikuje li se razina CD4 u ispitanika u istraživanju?

Škarpa I et al. Croat Med J 2001;42:551-5.



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## 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 preostala biti **odbačena**



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## 2. Odabir testa

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Nastavak (mjereno obilježje)

- Mjerne ljestvice
  - NOMINALNA
  - ORDINALNA
  - INTERVALNA
  - OMJERNA



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Nastavak (odabir testa)

| Ljestvica  | Jedan uzorak               | Dva uzorka              |                       | Tri i više uzoraka |                 |
|------------|----------------------------|-------------------------|-----------------------|--------------------|-----------------|
|            |                            | Ovisni                  | Neovisni              | Ovisni             | Neovisni        |
| Nominalna  | binomni test<br>hi-kvadrat | McNemara                | Fisher<br>hi-kvadrat/ | Cohran             | hi-kvad         |
| Ordinalna  | Kol/Smir.<br>homologni     | Wilcoxon<br>MW<br>Moses |                       | Friedman           | p/medijan<br>KW |
| Intervalna |                            |                         |                       |                    |                 |
| Omjerna    | ...                        |                         |                       |                    |                 |



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## 3. Određivanje razine značajnosti

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## (nastavak)

- zašto upravo to:  $P < 0,05$ ?

### POKUS

bacanje novčića: pismo/glava

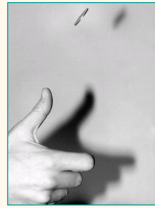
2x isto uzastopce = 0,5

3x = 0,25

4x = 0,125

5x = 0,0625

...

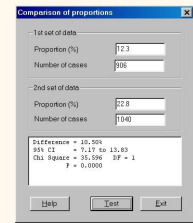


Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## 4. Izračunavanje

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## 5. Zaključivanje

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Pogrješke testiranja hipoteze

| PRAVO STANJE                 |                                              | ZAKLJUČENO      |
|------------------------------|----------------------------------------------|-----------------|
| RAZLIKA POSTOJI ( $H_1$ )    | RAZLIKA NE POSTOJI ( $H_0$ )                 |                 |
| ISPRAVAN ZAKLJUČAK           | $\alpha$ pogreška (I. vrste) ( $H_0$ odbac.) | RAZLIKA POSTOJI |
| $\beta$ pogreška (II. vrste) | ISPRAVAN ZAKLJUČAK ( $H_0$ prihvat.)         | RAZLIKE NEMA    |

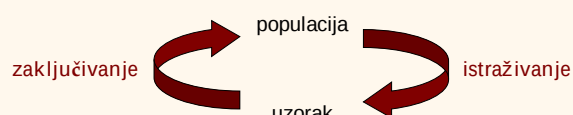


Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Statistički pojmovnik

- istraživanje donosi zaključke o populaciji



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Uzorak i populacija



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Zašto uzorak? Zašto ne populacija?

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Uzorak

- *reprezentativan*
- *mjerljiv*
- *slučajni (probabilistički)*

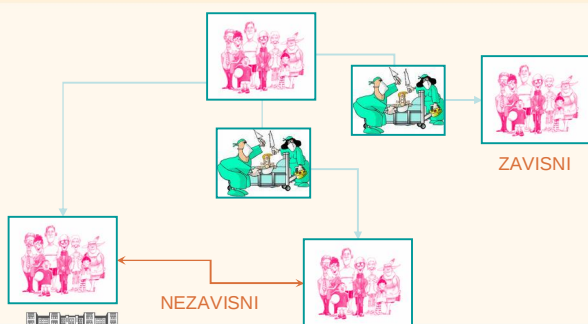
- jednostavni slučajni
- sustavni slučajni
- slojevit slučajni
- skupovni (klasterirani) slučajni



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Zavisni i nezavisni uzorci



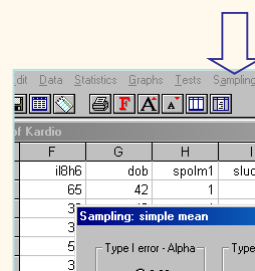
Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## Uzorak (nastavak) – veličina

- uzorkovanje - MedCalc (engl. sampling)

- primjer:  
sampling  
simple mean  
...



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



## WWW...

- <http://faculty.vassar.edu/lowry/VassarStats.html>
- <http://www.staterv.com/>
- <http://www.randomizer.org/>
- <http://www.bettycjung.net/Statsites.htm>
- <http://shazam.econ.ubc.ca/flip/>
- <http://statpages.org/>
- <http://www.stat.psu.edu/~resources/index.htm>
- ...
- <http://mi.medri.hr/webveze.htm>



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



[mladenp@kbd.hr](mailto:mladenp@kbd.hr)

Klinička bolnica Dubrava  
Klinički zavod za laboratorijsku dijagnostiku  
Avenija G. Šuška 6, 10000 Zagreb  
☎ 01 290 3379

Katedra za medicinsku informatiku  
Medicinski fakultet Sveučilišta u Rijeci  
Ulica braće Branchetta 20, 51000 Rijeka  
☎ 051 651 255



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

