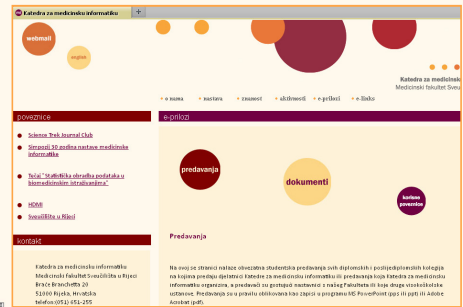


Statistička obradba biomedicinskih podataka: uvodna razmatranja

Prof. dr. sc. Mladen Petrovečki



Sva predavanja: mi.medri.hr



<http://mi.medri.hr/e-prilozi.html>



Raspored nastave

Biostatistika (studij na hrvatskom jeziku)

ponedjeljak 30. svibnja – petak 3. lipnja 2016., 15 – 19h30 sati

Popis tematskih cjelina:

- Statistička metodologija znanstvenog rada i statistička hipoteza
- Odabir statističkog testa
- Tablični i grafički prikaz podataka (deskriptivna statistika, prikaz: kategoričkih i brojničkih podataka)
- Vrste istraživanja
- Granice pouzdanosti
- Činjenice slučajne
- Logičke zakonitosti znanstvenog rada
- Najčešće statističke pogreške u istraživanjima iz područja dentalne medicine

Voditelj nastave: prof. dr. sc. Mladen Petrovečki

Suradnice u nastavi: prof. dr. sc. Ana-Maria Šimundić, prof. dr. sc. Lidija Bilić-Zulte



Ispit

- pisani seminarski rad
- kritička raščlamba statističkih postupaka objavljenog znanstvenog rada
 - područje dentalne medicine (biomedicine)
 - dostavljanje:
 - elektronički ⇒ mladenp@kbd.hr
 - u doc ili rtf formatu
 - zajedno s pdf-om izvornoga rada
 - poslati desetak dana prije usmene obrane
 - pomoćna pitanja:
http://mi.medri.hr/assets/pomocna_pitanja_za_seminar.pdf



Ispit

- rad koji se tumači:
 - iz časopisa citiranog u *Current Contents-u*
 - cjelokupno istraživanje (NE osvrt na drugi objavljeni rad)
 - najmanje dva priloga (tablica, slika)
 - jasan opis (tekst) statističkih postupaka
- **nema ispitnih razdoblja!**



Ispit

- seminarski rad
 - ime i prezime polaznika
 - studij, kolegij, godina nastave
 - naziv (referencija) rada koji se tumači
 - literatura (ako se rabi)
- primjer (ako zatreba):
<http://mi.medri.hr/assets/ispit-seminar.pdf>



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Literatura

- Dawson-Saunders B, Trapp RG. Basic & Clinical Biostatistics. Prentice-Hall Int. Inc., London, 2004.
- Petrie A, Sabin C. Medical statistics at a glance. Blackwell Science, Oxford, 2000.
- Marušić M i sur. Uvod u znanstveni rad u medicini. Medicinska naklada, Zagreb, 2013.
- Petz B, Kolesarić V, Ivanec D. Petzova statistika – Osnovne statističke metode za nematematičare. Naklada Slap, Jastrebarsko, 2012.
- StatSoft Inc. Electronic Statistics Textbook. Tulsa, OK: StatSoft, 2002. (<http://www.statsoft.com/textbook/stathome.html>)
- Moher D, Schulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. Lancet 2001;357:1191-4. (<http://www.consort-statement.org/revisestatement.htm>)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Literatura

Znanost je sumnja.
(N. Alegretti)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Statistička metodologija znanstvenog istraživanja



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Metodologija znanstvenog 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



Pokus

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Nasumični odabir

- funkcija slučajnog broja (uniformna raspodjela)
- npr. Excel → RAND()
- biranje slučajnih brojeva
 - <http://graphpad.com/quickcalcs/RandMenu.cfm>
 - <https://www.random.org/>
 - <http://www.randomizer.org/>



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Odluka, ždrijeb, slučajnost

Kada je zvao junaka na mejdan, sâm ja se s njime
Ogledah, a vi moliste tad, **da iziđe ždr'jeb moj**,
I bi po molitvi vašoj. A pitate l', kakav svršetak
Toga bješe mejdana: nijesam od Hektora svladan.
...
Ovidije, Metamorfoze, 13. pjevanje, 87. i dalje



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Slučajnost, značajnost

- slučajnost (*randomness*)
 - skup događaja (uzroka) koji dovode do posljedica koje je nemoguće predvidjeti njihovim poznavanjem (tj. poznavanjem uzroka)
 - događaj bez uzroka
 - svaki nedeterministički događaj
 - pseudonim kojim se služi Bog kad se ne želi potpisati (Anatole France)
- značajnost (*significance*)
 - sve što se ne događa slučajno
 - sve što može poći loše neće završiti dobro (Murpheyjev zakon, IV. zakon termodinamike)



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 (engl. *probability*)

$$0 < P < 1$$



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Izračun vjerojatnosti

- vjerojatnost događaja P
$$P = \frac{\text{broj povoljnih mogućnosti}}{\text{ukupni broj mogućnosti}}$$
- 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; 3. pogl.

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Statistička hipoteza

- elementarna tvrdnja
- točna (istinita) ili netočna (neistinita)
- provjera hipoteze \Rightarrow **traženje istine**



Ivana Brlić Mažuranić
Kako je Potjeh tražio istinu
Mladost, Zagreb: Albert Kinert, 1967.



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Statistička hipoteza

- istina \Rightarrow stvarno, objektivno stanje
- probabilistički sustav:
istina \Rightarrow **vjerojatnost**
- značajno \Rightarrow ono što se ostvaruje na svaki drugi način osim slučajno:
iskaz vjerojatnosti \Rightarrow **razina značajnosti**



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Testiranje statističke hipoteze

1. postavljanje hipoteze
2. odabir statističkog testa
3. određivanje razine značajnosti
4. izračunavanje statistike testa
5. zaključivanje



Testiranje statističke hipoteze, primjer

Samopozna halitoza kod diplomskih studenata medicine, dentalne medicine i zdravstvenih znanosti na Sveučilištu Sains Malaysia

Self-Reported Halitosis among Medical, Dental and Health Science Undergraduate Students at the University Sains Malaysia



Testiranje statističke hipoteze, primjer

Tablica 2. Čimbenici povezani sa samopoznaenom halitozom (n=300)

Razlikuje li se učestalost samopoznaene halitoze prema... pojavnosti naslaga na jeziku?

Self-Reported Halitosis among Medical, Dental and Health Science Undergraduate Students at the University Sains Malaysia

Samopoznaena halitoza:	DA		NE		Statistika	
	n	%	n	%	χ^2 (df)	P
Naslaga na jeziku • Tongue coating:						
Da • Ya	17	(19,3)	71	(8,07)	4,935 (1)	0,036
Ne • No	22	(10,4)	190	(89,6)		



1. Postavljanje hipoteze

- ništična:
→ 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 **odbaćena**



Ništična hipoteza

Nema razlike ≈ "Nisam kriv"



Ništična hipoteza

Nema razlike



2. Odabir testa

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



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Nastavak (mjerno obilježje)

- Mjerne ljestvice
 - nominalna
 - ordinalna
 - intervalna
 - omjerna



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Mjerenje i mjerne ljestvice

- postupak kojim primjenom jednoznačnog pravila ispitivanim jedinkama pridružujemo brojne vrijednosti ili oznake
- rezultat mjerenja → podatak
- instrument mjerenja → mjerna ljestvica
 - kvalitativne:
 - nominalna (kategorička)
 - ordinalna (uredbena)
 - kvantitativne (kardinalne):
 - intervalna
 - omjerna



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Nominalna ljestvica

- nazočnost bjelančevina u mokraći (testna traka):
 - da/ne, tj. pozitivan ili negativan nalaz
- klasifikacije, nomenklature
 - FAB-klasifikacija akutnih leukemija



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Ordinalna ljestvica

- nazočnost bjelančevina u mokraći (testna traka)
 - negativan nalaz, poz. +, poz. ++, poz. +++
- razredi u smislenom (rastućem) nizu
 - stupanj opekline (I – IV)
- širina razreda: nejednaka
- nedostatak: subjektivna procjena



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Intervalna ljestvica

- temperatura u °C, pH
- izražava količinu → brojna vrijednost odgovara vrijednosti obilježja
- svi intervali su jednake širine
- vrijednost "nula" ≠ nepostojanje obilježja
- "nula" → dogovorena vrijednost
- omjer dvaju mjerenja → besmislen



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Omjerna ljestvica

- mjerenje količine bjelančevina u mokraći (rezultat izražen u g/L)
- kvantitativni podatak – brojnost jedinice mjerenja
- vrijednost “nula” = nepostojanje obilježja
- omjer → omogućuje procjenu vrijednosti
 - npr: GUK=12,0 mmol/L ⇒ dvostruko veća koncentracija od gornje granice referentne vrijednosti



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Pretvorba podataka

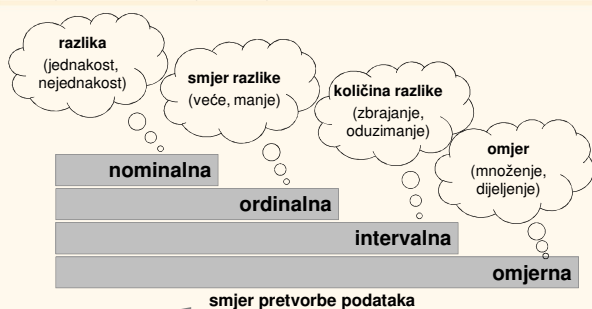
- samo jedan smjer
 - od složenijih k jednostavnijima
- kriterij: granična (*cut-off*) vrijednosti
- vrsta studije → utvrđuje vrstu podataka
- primjer:
 - načini bilježenja dobi (dob u godinama; dobna skupina; stariji ili mlađi od 50 godina)
- svako prevođenje podataka mora se jasno opisati i tumačiti



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Ljestvice mjerenja



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



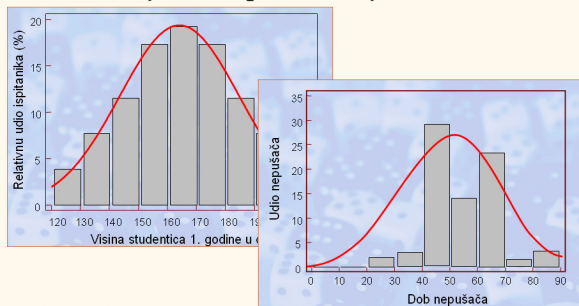
Nastavak (odabir testa)

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

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Parametrijski i neparametrijski testovi



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



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. $P < 0,05$



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. Izračun

- matematički račun
- $P \Rightarrow$ egzaktna vrijednost
 - 3 decimalna mjesta
 - npr. $P = 0,325$

	Sample 1	Sample 2
Sample size	19	19
Arithmetic mean	11,3000	11,2300
95% CI for the mean	2,6261 to 11,3640	6,4540 to 16,0197
Variance	17,3000	96,4713
Standard deviation	4,1593	9,8223
Standard error of the mean	1,0398	2,2766
T-test for equal variances	P = 0,063	

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



5. Zaključivanje

- mala vrijednost $P \Rightarrow$ mala vjerojatnost da odbacujemo (ne prihvaćamo) istinitost
- zaključivanje:
 - $P < \alpha$
 - vjerojatnost P mala
 - odbacujemo (ne prihvaćamo) H_0
 - prihvaćamo alternativnu hipotezu, H_1
 - potvrdimo je, *iskažemo je, uz $P = \dots$*



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Testiranje statističke hipoteze, primjer

	Samoopažena halitoza: DA		NE		Statistika	
	n	%	n	%	χ^2 (df)	P
Naslage na jeziku • Tongue coating						
Da • Yes	17	(19,3)	71	(8,07)	4,935 (1)	0,036
Ne • No	22	(10,4)	190	(89,6)		

Tablica 2. Čimbenici povezani sa samoopaženom halitozom (n=300)
Razlikuje li se učestalost samoopažene halitoze prema... pojavnosti naslaga na jeziku?

Da, jer je $0,036$ ($P < 0,05$) (a).

Rezultat: Samoopažanje halitoze češće je u osoba s naslagama na jeziku (17/88, 19%) od onih koje naslaga na jeziku nemaju (22/212, 10%; $P = 0,036$).

Rezultat: Naslage na jeziku češće se nalaze u osoba koje samoopažaju halitozu (17/39, 44%) od onih koje je ne opažaju (71/261, 27%; $P = 0,036$).



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Pogrješke testiranja hipoteze

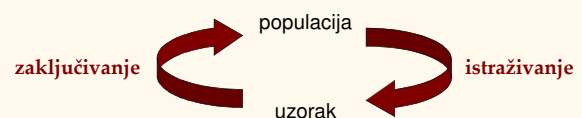
PRAVO STANJE		ZAKLJUČENO
RAZLIKA POSTOJI (H_1)	RAZLIKA NE POSTOJI (H_0)	
ISPRAVAN ZAKLJUČAK	α pogriješka (I. vrste) (H_0 odbacivanje)	RAZLIKA POSTOJI
β pogriješka (II. vrste)	ISPRAVAN ZAKLJUČAK (H_0 prihvaćanje)	RAZLIKE NEMA

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)

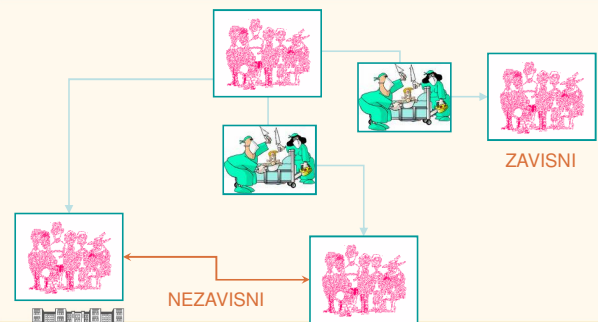
- jednostavni slučajni
- sustavni slučajni
- slojevit slučajni
- skupovni (klasterirani) slučajni



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Zavisni i nezavisni uzorci

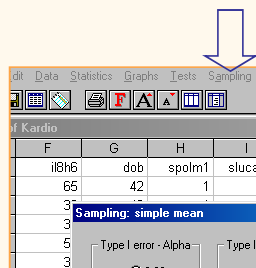


Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Uzorak (nastavak) – veličina

- uzorkovanje – MedCalc (engl. *sampling*)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



WWW

- <http://faculty.vassar.edu/lowry/VassarStats.html>
- <http://www.statserve.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>



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



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

