

Priprema podataka za statističku obradu

Doc. dr. sc. Gordana Brumini
Katedra za medicinsku informatiku

中華書局影印

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Koliko ispitanika treba tj.
koliki N treba?

- Procjena veličine uzorka
 - Treba 3 parametra
 - značajna razlika
 - snaga testa
 - razina značajnosti



Medicinska fakultet Sveučilišta u Rijeci Katedra za medicinsku informaticu

Najčešća pitanja?

- Koliki ispitanika treba za doktorat, je li dovoljno 100 ispitanika?
 - Kako treba unašati podatke?
 - Koje varijable koristiti?
 - Jesu li ove varijable dobre/dovoljne?



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Razina (statističke) značajnosti

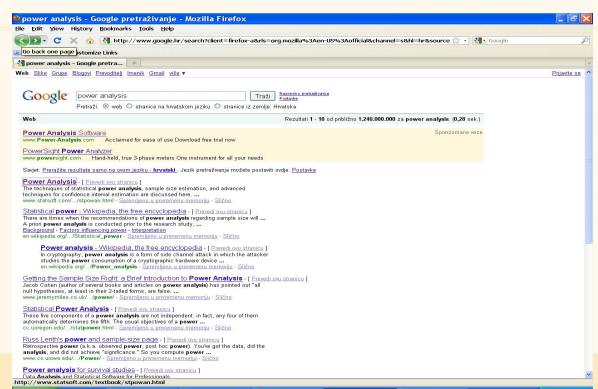
- $P < 0,05$
 - Nul-hipoteza se odbacuje, postoji statistička značajnost
 - u 95 % slučajeva statistička značajnost postoji, u 5 % slučajeva nema značajnosti
 - $P \geq 0,05$
 - Nul-hipoteza se prihvata, ne postoji statistička značajnost
 - u 95 % slučajeva statistička značajnost ne postoji, u 5 % slučajeva postoji



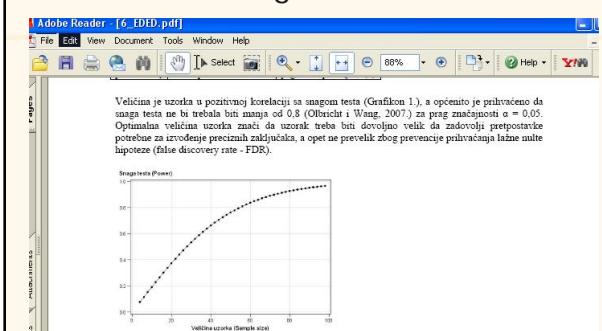
Medicinali Škola | Medicinali Škola je medicinska učionica

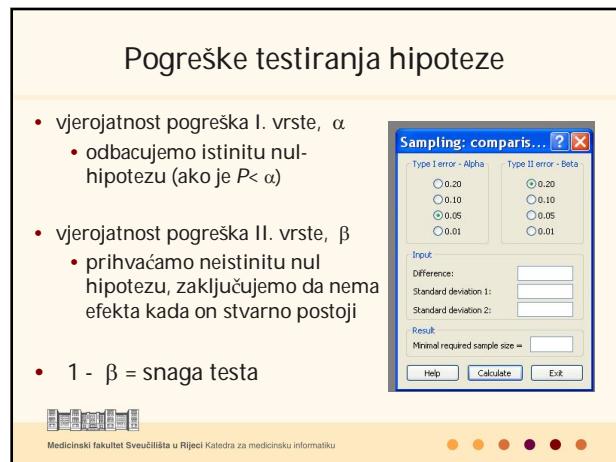
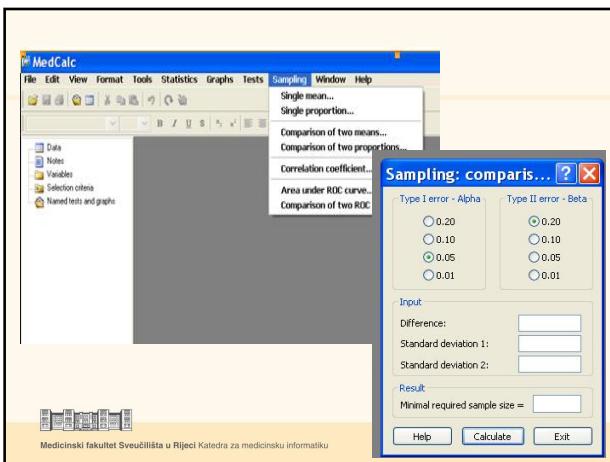
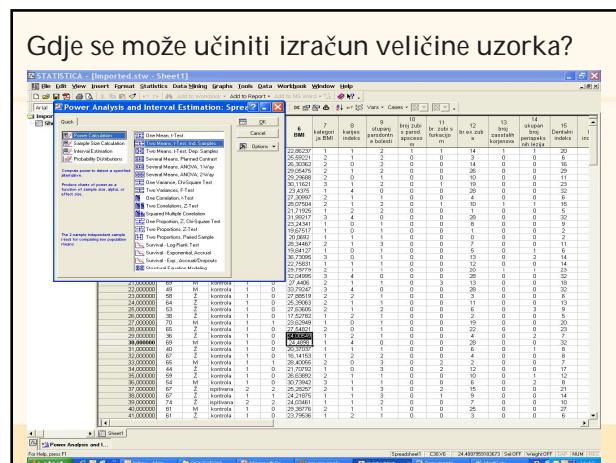
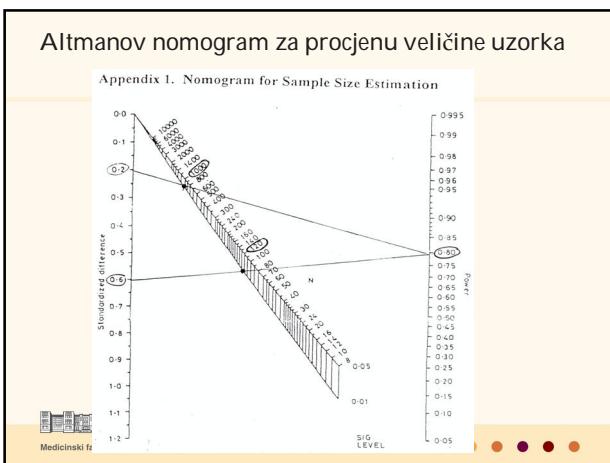
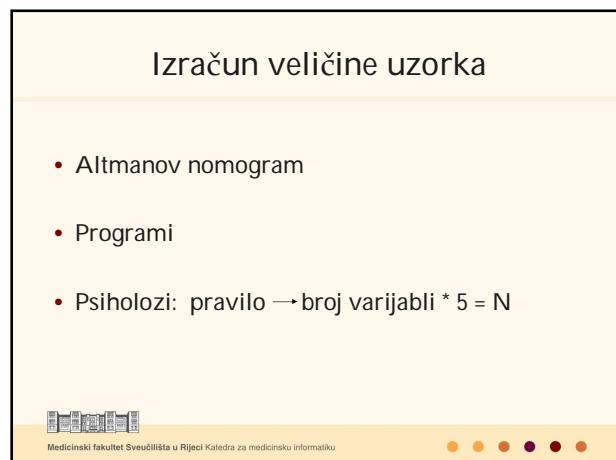
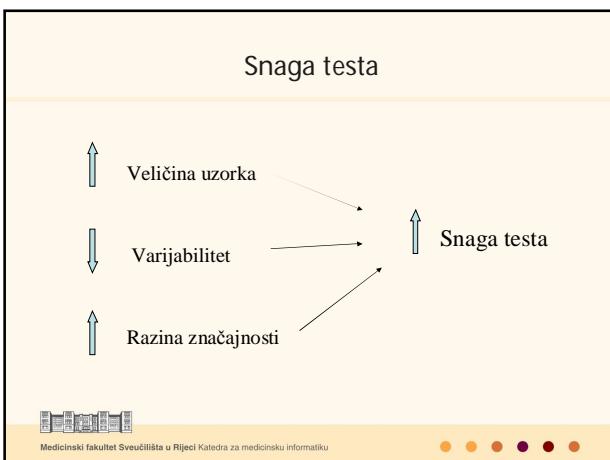


Power analysis



Snaga testa





Snaga testa

- snaga testa ocjenjuje kvalitetu samog testa
- definira koliki uzorak je potreban za statistički zaključak
- koliko je vjerojatno da se nekim statističkim testom određuje mjereno svojstvo



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Analiza snage testa

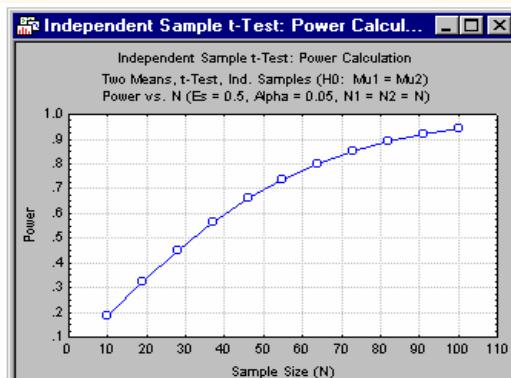
- "dobar test"- snaga bi trebala biti 70-80 %
- etički neprihvatljivo, gubitak vremenena i novaca provoditi istraživanje koje ima 40% vjerojatnost da otriva efekt koji se istražuje
- pr. mjerjenje stava s upitnikom (snaga testa 80 %)
 - u 80 % postoji vjerojatnost da se tim upitnikom doista mjeri stav
 - u 20 % postoji vjerojatnost da taj upitnik ne mjeri stav



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Krivulja snaga testa i veličina uzorka



Me

Koliko ispitanika treba tj. koliki N treba?

- Procjena veličine uzorka
- Treba 3 parametra
 - značajna razlika ✓
 - snaga testa ✓
 - razina značajnosti ✓



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Podaci

- | | |
|--|--|
| <ul style="list-style-type: none">• Kvalitativni• Nominalne i ordinalne mjerne ljestvice• Opisni, bez mjerne jedinice• 5 pušača i 2 nepušača• opelkine I. -IV. stupnja | <ul style="list-style-type: none">• Kvantitativni• Omjerene i intervalne mjerne ljestvice• Mjerljivi i imaju mjerну jedinicu<ul style="list-style-type: none">• 10 cigareta na dan• BMI• 45 godina |
|--|--|



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Nominalna ljestvica

- Nazočnost prisustva na nastavi
 - da/ne
- Pušenje
 - da/ne
- Kardiovaskularna bolest
 - ima/nema
- Spol
 - muški, ženski



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Ordinalna ljestvica

- ocjenjivanje studenata 1-5 ili A - E
- razredi u smislenom (rastućem) nizu
- širina razreda može biti nejednaka
- katkada se mogu tretirati kao podaci iz intervalne mjerne ljestvice
- opékline od I do IV stupnja
- stadiji bolesti (Ca,...)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Intervalna ljestvica

- izražava količinu → brojevna vrijednost odgovara vrijednosti obilježja
- može imati mjernu jedinicu
- temperatura, pH vrijednost, IQ,
- "nula" → dogovorena vrijednost (pr. 0 °C)
- $t_1 = -10 \text{ } ^\circ\text{C}$, $t_2 = 20 \text{ } ^\circ\text{C}$, nije 2 puta veća nego je razlika 30 °C



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Omjerna ljestvica

- kvantitativni podatak – brojnost jedinice mjerena
- masa, dob, koncentracija, broj studenata na I. godini
- vrijednost "nula" → nepostojanje obilježja
- omjer → omogućuje procjenu vrijednosti
 - npr: BMI ... dvostruko veći od ...



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Pravila za prikupljanje podataka

- što više varijabli
 - za multivarijantne analize treba više zavisnih varijabli
 - izbjegavati kategoriske varijable
- kvantitativni podaci mogu se prevesti u kvalitatitivne
- čuvati izvorne podatke
- po završetku EU projekta revizija i nakon 5 godina, podaci iz doktorata do 10 godina



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Kako upisati podatke



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Programi u kojima možete upisati podatke

- MS Excel
- MedCalc
- Statistica



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Kako pravilno upisati podatke u Excel

- Sve podatke upisati u jedan radni list (Sheet)
- Izbjegavati boje
- Objasniti legendu
- Što više varijabli imati
 - posebice kontinuirane podatke
- voditi računa o ciljevima istraživanja



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Primjer loše tablice

	KUT NARUDZ VALUOSA	KUT između I i II METATARZALNE KOSTILRELLEROVA OP.	KUT između I i IV METATARZALNE KOSTILRELLEROVA OP.
1	45	20	15
2	50	19	17
3	50	11	23
4	40	10	10
5	51	10	14
6	2	20	20
7	35	12	10
8	28	14	16
9	42	11	15
10	40	10	15
11	51	10	14
12	2	20	18
13	35	12	10
14	28	14	16
15	40	11	15
16	35	12	20
17	41	14	16
18	40	17	15
19	34	12	15
20	15	10	15
21	35	14	16
22	10	17	20
23	10	16	16
24	20	10	15
25	20	14	20
26	20	10	15
27	20	14	16
28	20	10	15
29	20	14	16
30	24	10	17
31	20	10	15
32	20	17	15
33	20	15	17
34	20	17	15
35	20	15	17
36	20	17	15
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			
101			
102			
103			
104			
105			
106			
107			
108			
109			
110			
111			
112			
113			
114			
115			
116			
117			
118			
119			
120			
121			
122			
123			
124			
125			
126			
127			
128			
129			
130			
131			
132			
133			
134			
135			
136			
137			
138			
139			
140			
141			
142			
143			
144			
145			
146			
147			
148			
149			
150			
151			
152			
153			
154			
155			
156			
157			
158			
159			
160			
161			
162			
163			
164			
165			
166			
167			
168			
169			
170			
171			
172			
173			
174			
175			
176			
177			
178			
179			
180			
181			
182			
183			
184			
185			
186			
187			
188			
189			
190			
191			
192			
193			
194			
195			
196			
197			
198			
199			
200			
201			
202			
203			
204			
205			
206			
207			
208			
209			
210			
211			
212			
213			
214			
215			
216			
217			
218			
219			
220			
221			
222			
223			
224			
225			
226			
227			
228			
229			
230			
231			
232			
233			
234			
235			
236			
237			
238			
239			
240			
241			
242			
243			
244			
245			
246			
247			
248			
249			
250			
251			
252			
253			
254			
255			
256			
257			
258			
259			
260			
261			
262			
263			
264			
265			
266			
267			
268			
269			
270			
271			
272			
273			
274			
275			
276			
277			
278			
279			
280			
281			
282			
283			
284			
285			
286			
287			
288			
289			
290			
291			
292			
293			
294			
295			
296			
297			
298			
299			
300			
301			
302			
303			
304			
305			
306			
307			
308			
309			
310			
311			
312			
313			
314			
315			
316			
317			
318			
319			
320			
321			
322			
323			
324			
325			
326			
327			
328			
329			
330			
331			
332			
333			
334			
335			
336			
337			
338			
339			
340			
341			
342			
343			
344			
345			
346			
347			
348			
349			
350			
351			
352			
353			
354			
355			
356			
357			
358			
359			
360			
361			
362			
363			
364			
365			
366			
367			
368			
369			
370			
371			
372			
373			
374			
375			
376			
377			
378			
379			
380			
381			
382			
383			
384			
385			
386			
387			
388			
389			
390			
391			
392			
393			
394			
395			
396			
397			
398			
399			
400			
401			
402			
403			
404			
405			
406			
407			
408			
409			
410			
411			
412			
413			
414			
415			
416			
417			
418			
419			
420			
421			
422			
423			
424			
425			
426		</td	

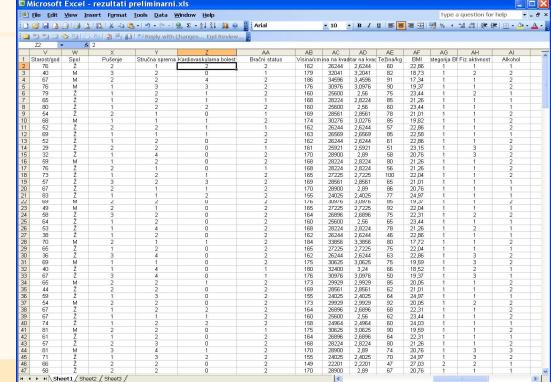
Zašto su bolji kontinuirani podaci

- Parametrijski testovi
- Veća snaga testa
- Multiple analize



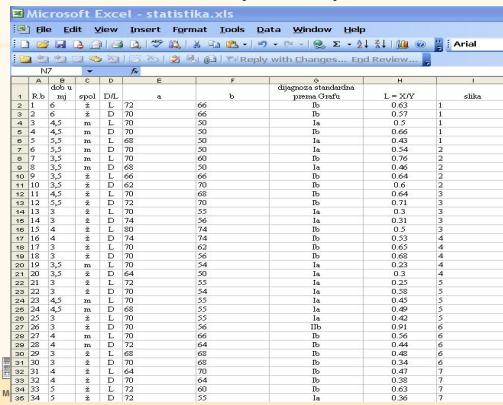
Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Dobro upisani podaci



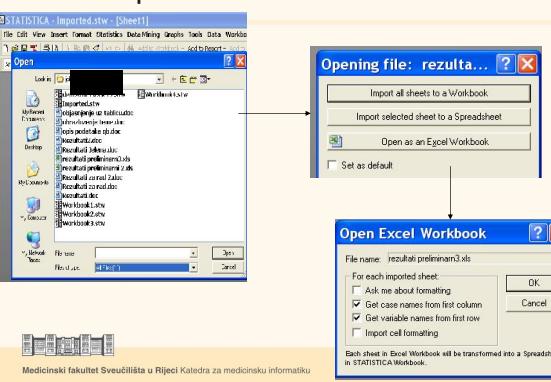
Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Dobro upisani podaci



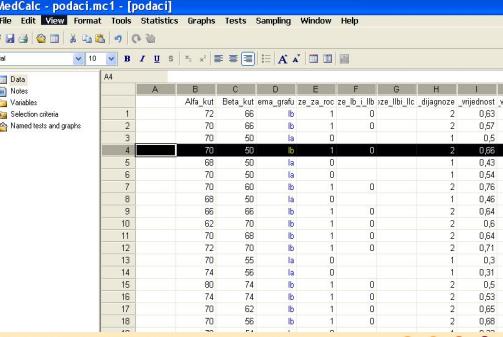
Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Unos podataka iz Excela u Statisticu



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Podaci u MedCalcu



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

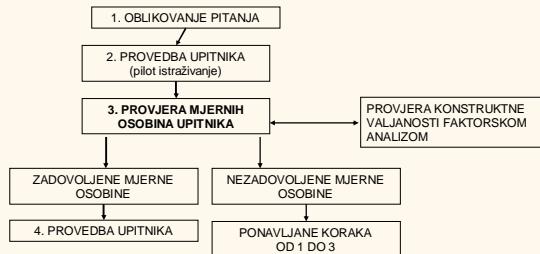
Oblikovanje upitnika

- Definiranje ciljeva
- Odabir uzorka ispitanika
- Oblikovanje pitanja
 - otvorenenog i/ili zatvorenog tipa- Likertova skala
- Provedba
- Provjera mjernih osobina



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Koraci u oblikovanju upitnika



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Oblikovanje pitanja

- Jasna, kratka i jednoznačna pitanja
- Izbjegavati pristrane ponuđene odgovore
Pušite li? Da/Ne
- Obratiti pažnju na izraze
sinonimi, više značnice, pozitivna/negativna pitanja
- Izbjegavati neugodna i hipotetska pitanja
- Poštovati privatnost i anonimnost

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Primjeri pitanja

- Pušite li:
 - da
 - ne
 - svakodnevno
 - povremeno
- Računalom se koristite u svrhu:
 - posla
 - igre
 - kupovine
- Pušite li:
 - ne
 - povremeno
 - svakodnevno
- Koristite li računalao za rad:
 - da
 - ne

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Primjer pitanja u upitniku 1

1. Dob (godina)
2. Spol M Ž
3. Ustanova u kojoj radite (ukoliko radite u više ustanova, upišite onu u kojoj ste zaposleni u većinskom udjelu):
 - Fakultet
 - Znanstveni institut
 - Bolница ili druga zdravstvena ustanova

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Primjer pitanja u upitniku 2

1.	Kadkad sam u iskušenju da plagiram jer svi ostali (studenti, istraživači, lječnici) to čine.	1	2	3	4	5
2.	Plagiranje osvrta me istraživački duh.	1	2	3	4	5
3.	Prekratki rokovi daju mi pravo na malo plagiranja.	1	2	3	4	5
4.	Samoplagiranje NE bi trebalo kažnjavati kao plagiranje.	1	2	3	4	5

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Pisana uputa 1

Molim Vas da budete sudionici ovog istraživanja. Istraživanje će obuhvatiti sve učenike prvog razreda srednje škole Primorsko-goranske županije. Ovim se upitnikom ispituju stavovi ljudi o pušenju.
Unaprijed vam hvala na suradnji!

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Pisana uputa 2

Molim Vas da budete sudionici u istraživanju **kojega provodim u sklopu svojeg doktorskog rada.**

Istraživanje će obuhvatiti sve učenike prvog razreda srednje škole Primorsko-goranske županije. Ovim se upitnikom ispituju **razmišljanja** ljudi o pušenju.

Ispitanje je anonimno, ne zanima nas Vaše ime i prezime, nitko neće biti upoznat s Vašim odgovorima.

Unaprijed vam hvala na suradnji!



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Pouzdanost upitnika

- Određuje pogreške mjerenja u opaženom rezultatu.
- Stupanj u kojem su rezultati dosljedni, sigurni ili ponovljivi (Croanbach koeficijent $\alpha > 0,7$)
- Savršeno pouzdan test - rezultati svakog mjerjenja su identični (Croanbach koeficijent $\alpha=1$)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



ROC analiza

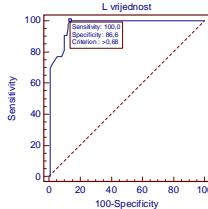
- engl. *Receiver Operating Characteristic Curve*
- za određivanje valjanosti dijagnostičke metode
- valjanost je sposobnost dijagnostičkog testa da pravilno klasificira ispitanike u bolesne i zdrave
- određuje se temeljem
 - osjetljivost testa
 - specifičnost testa



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



ROC krivulja

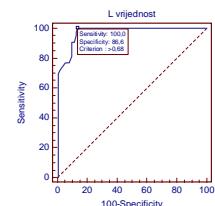


Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

- daje podatke o:
 - osjetljivosti i specifičnosti testa
 - definira vrijednost kriterijskog prediktora (najmanji broj LP i LN)
- Kvadratič opisuje:
 - osjetljivost = 100 %, (SP/SP+LN)
 - specifičnost 86,6 % (LP)
 - kriterijski prediktor $>0,68$



Površina ispod ROC krivulje



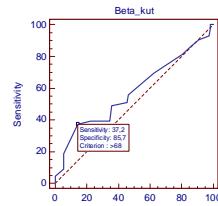
- Površina ispod ROC krivulje = 0,967; $P < 0,001$
- $0,9 - 1 =$ izvrstan test
- $0,8 - 0,9 =$ dobar test
- $0,7 - 0,8 =$ osrednji test
- $0,6 - 0,7 =$ slabiji test
- $0,5 - 0,6 =$ test bez uspjeha



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



ROC krivulja



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

- Površina ispod ROC krivulje = 0,579; $P = 0,133$

Površina ispod ROC krivulje	0,579
Standardna pogreška	0,052
95% granice pouzdanosti CI	0,497 to 0,857
Razina značajnosti P	0,133



ROC analiza - izračun vrijednosti kriterijskog prediktora						
Kriterij	Osjetljivost/%	95% CI	Specifičnost/%	95% CI	+LP	-LN
>0,34	100,00	91,7 - 100,0	0,00	0,0 - 3,3	1,00	0,00
>0,34	100,00	91,7 - 100,0	0,89	0,1 - 4,9	1,01	0,00
>0,35	100,00	91,7 - 100,0	1,79	0,3 - 8,3	1,02	0,00
>0,36	100,00	91,7 - 100,0	2,68	0,6 - 7,6	1,03	0,00
>0,43	100,00	91,7 - 100,0	3,57	1,0 - 8,9	1,04	0,00
>0,44	100,00	91,7 - 100,0	4,46	1,5 - 10,1	1,05	0,00
>0,45	100,00	91,7 - 100,0	6,25	2,6 - 12,5	1,07	0,00
>0,47	100,00	91,7 - 100,0	11,61	6,3 - 19,0	1,13	0,00
>0,48	100,00	91,7 - 100,0	13,39	7,7 - 21,1	1,15	0,00
>0,5	100,00	91,7 - 100,0	16,07	9,8 - 24,2	1,16	0,00
>0,61	100,00	91,7 - 100,0	17,88	11,3 - 36,2	1,28	0,00
>0,52	100,00	91,7 - 100,0	20,21	15,1 - 37,1	1,30	0,00
>0,53	100,00	91,7 - 100,0	27,68	19,6 - 36,9	1,38	0,00
>0,64	100,00	91,7 - 100,0	29,46	21,2 - 38,8	1,42	0,00
>0,55	100,00	91,7 - 100,0	37,50	28,5 - 47,1	1,60	0,00
>0,56	100,00	91,7 - 100,0	41,98	32,7 - 51,7	1,72	0,00
>0,57	100,00	91,7 - 100,0	52,68	43,0 - 62,2	2,11	0,00
>0,58	100,00	91,7 - 100,0	55,36	45,7 - 64,8	2,24	0,00
>0,59	100,00	91,7 - 100,0	58,04	48,3 - 67,3	2,38	0,00
>0,6	100,00	91,7 - 100,0	62,50	52,9 - 72,1	2,61	0,00
>0,61	100,00	91,7 - 100,0	64,93	54,1 - 73,1	2,70	0,00
>0,63	100,00	91,7 - 100,0	67,86	58,4 - 76,4	3,11	0,00
>0,64	100,00	91,7 - 100,0	69,64	60,2 - 78,0	3,29	0,00
>0,65	100,00	91,7 - 100,0	75,00	65,9 - 82,7	4,00	0,00
>0,66	100,00	91,7 - 100,0	83,04	74,8 - 89,5	5,89	0,00
>0,67	100,00	91,7 - 100,0	83,93	75,8 - 90,2	6,22	0,00
>0,68 *	100,00	91,7 - 100,0	86,61	78,9 - 92,3	7,47	0,00
M	0,69	97,67	87,7 - 99,6	87,50	79,9 - 93,0	7,81
SE	0,00	0,00	0,00	0,00	0,027	0,00

Mutipla regresijska analiza				
Čimbenik	β-koeficijent	P	Udio doprinosa (%)	
Životna dob	-0,044	0,191	0,5	
Spol	0,019	0,506	0,2	
Školska spremja	0,079	0,028	1,0	
Informatičko obrazovanje	0,085	0,012	1,5	
Uporaba računalnika	0,302	<0,001	10,1	

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

