

ANALIZA TABLICA KONTINGENCIJA

Mirjana Kujundžić Tiljak i Davor Ivanković

Za prikaz obilježja mjerena nominalnom ili ordinalnom skalom mjerenja koristimo $n \times m$ tablice kontingencije .

Za testiranje razlika u distribucijama obilježja mjerenih nominalnom ili ordinalnom skalom između neovisnih uzoraka koristi se χ^2 -test. Na taj način ocjenjujemo da li naša empirijska distribucija pokazuje samo slučajna odstupanja kao rezultat uzorkovanja ili su odstupanja nastala pod utjecajem nekih snažnih faktora, tj. da li neka empirijska razdioba odgovara razdiobi koju bismo mogli očekivati na osnovu hipoteze.

Dobiveni rezultat se interpretira ili kao ocjena povezanosti uspoređivanih obilježja, ili kao ocjena statističke značajnosti razlika među distribucijama.

Računa se:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Dakle, χ^2 je suma svih kvadrata razlike opaženog i očekivanog broja jedinki u ćeliji tablice podijeljenih s očekivanim brojem jedinki u istoj ćeliji tablice.

Pri tome, očekivana frekvencija jedinki u svakoj pojedinoj ćeliji mora biti najmanje 5.

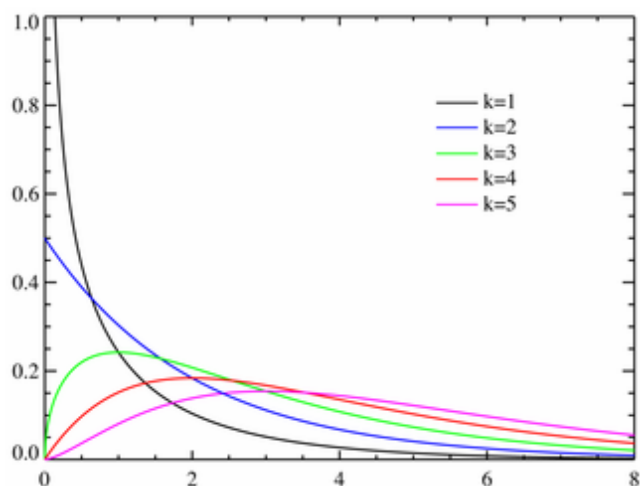
Stupnjevi slobode (DF) računaju se kao (broj redaka - 1) \times (broj stupaca-1).

Moguće vrijednosti rezultata χ^2 -testa slijede χ^2 -distribuciju - distribuciju pozitivnih vrijednosti, iskrivljenu prema desno. Njezin karakterističan parametar su *stupnjevi slobode*. Oblik distribucije ovisi o stupnjevima slobode. χ^2 -distribucija postaje sve više simetrična i sličnija normalnoj kako raste broj stupnjeva slobode.

Interpretacija χ^2 - testa:

Ukoliko je dobivena vrijednost χ^2 manja od granične χ^2 vrijednosti za datu vjerojatnost tada je $P > \alpha$, tj. prihvaćamo H_0 .

Ukoliko je dobivena vrijednost χ^2 veća od granične χ^2 vrijednosti za datu vjerojatnost tada je $P < \alpha$, tj. odbacujemo H_0 .



Slika 1. Krivulje χ^2 - distribucije

Izvor slike: http://en.wikipedia.org/wiki/Chi-square_distribution

Primjer: Povezanost HLA-B8 i idiopatske Addisonove bolesti

HLA – B8 (OPAŽENE FREKVENCije, OČEKIVANE FREKVENCije)			
	da	ne	ukupno
bolesnici	22 9	16 29	38
kontrola	467 480	1500 1487	1967
ukupno	489	1516	2005

$$\chi^2 = \frac{(22-9)^2}{9} + \frac{(16-29)^2}{29} + \frac{(467-480)^2}{480} + \frac{(1500-1487)^2}{1487}$$

$$\chi^2 = 23,67$$

$$df = 1$$

$$\chi^2_{\text{granično}} = 6,635 \dots \alpha = 0,01 \text{ (99\% vjerojatnost)}$$

Kada su ispirivani uzorci zavisni (uzorci parova), za testiranje razlika u distribucijama upotrebljavju se McNemarov χ^2 kada je obilježje nominalno te Stuart-Maxvellov test kada je obilježje ordinalno.

Računa se na način:

		Uzorak I Karakteristika A	
		DA	NE
Uzorak II Karakteristika B	DA	a	b
	NE	c	d

$$\chi^2 = \frac{(|b-c|-1)^2}{b+c} \quad \text{stupnjevi slobode} = 1$$

Primjer:

		Standardna terapija	
		poboljššan	nepoboljššan
Nova terapija	poboljššan	40	25
	nepoboljššan	10	0

$$\chi^2 = 5,06, \quad \text{df}=1; \quad P<0,05$$

OCJENA RIZIKA

Intenzitet povezanosti ispitivanih obilježja naziva se **rizik** za češće zajedničko pojavljivanje nekih stanja ispitivanih obilježja.

Relativni rizik je rizik izloženih osoba podijeljen s rizikom skupine koja nije izložena “osumnjičenom uzročnom faktoru”.

Atributivni rizik je rizik izloženih osoba umanjen za rizik neizložene skupine.

„ODDS RATIO“ (“omjer izgleda”) je omjer vjerojatnosti (izgleda) posjedovanja svojstva (“being a case”) u izloženoj skupini i vjerojatnosti (izgleda) posjedovanja svojstva (“being a case”) u neizloženoj skupini.

	Faktor		Ukupno
	DA	NE	
Bolesni	a	b	a+b
Zdravi	c	d	c+d
Ukupno	a+c	b+d	a+b+c+d

$$\text{Relativni rizik (RR)} = \frac{a/(a+c)}{b/(b+d)}$$

$$\text{Atributivni rizik (AR)} = \frac{a}{a+c} - \frac{b}{b+d}$$

$$\text{Odds ratio} = \frac{a/c}{b/d} = \frac{a \times d}{b \times c}$$

Primjer: Ocjena rizika HLA-B8 i idiopatske Addisonove bolesti

HLA – B8 (OPAŽENE FREKVENCije, OČEKIVANE FREKVENCije)			
	da	ne	ukupno
bolesnici	22	16	38
kontrola	467	1500	1967
ukupno	489	1516	2005

$$RR = 4,26$$

$$AR = 0,0449 - 0,01067 = 0,0344 \text{ (77\%)}$$

$$OR = 4,42$$

VALJANOST DIJAGNOSTIČKOG TESTA

Rezultati dijagnostičkih testova služe za postavljanje dijagnoze. Većina testova su nesavršeni, tj. zdrave jedinice mogu klasificirati kao bolesne, a bolesne kao zdrave).

Rezultati dijagnostičkih testova mogu biti:

“POZITIVNI” = veća vjerojatnost bolesti

“NEGATIVNI” = veća vjerojatnost odsustva bolesti

		STVARNO STANJE		
		Bolestan	Zdrav	
TEST	Pozitivan	a	b	$a + b$
	Negativan	c	d	$c + d$
		$a + c$	$b + d$	

$$\text{OSJETLJIVOST TESTA} = \frac{a}{a + c}$$

Osjetljivost testa je proporcija bolesnih osoba ispravno prepoznatih testom od ukupno stvarno bolesnih.

$$\text{SPECIFIČNOST TESTA} = \frac{d}{b + d}$$

Specifičnost testa je proporcija zdravih osoba ispravno prepoznatih testom od ukupnog broja zdravih.

$$\text{POZITIVNA PREDIKTIVNA VRIJEDNOST} = \frac{\mathbf{a}}{\mathbf{a + b}}$$

Pozitivna prediktivna vrijednost je proporcija bolesnih osoba ispravno prepoznatih testom od ukupno pozitivnih.

$$\text{NEGATIVNA PREDIKTIVNA VRIJEDNOST} = \frac{\mathbf{d}}{\mathbf{c + d}}$$

Negativna prediktivna vrijednost je proporcija zdravih osoba ispravno prepoznatih testom od ukupno negativnih

Primjeri:

Slučaj A

		STVARNO STANJE		
		Bolestan	Zdrav	
TEST	Pozitivan	450	50	500
	Negativan	50	450	500
		500	500	1000

$$\text{prevalencija} = \frac{500}{1000} = 0,5$$

$$\text{osjetljivost} = \frac{450}{500} = 0,9$$

$$\text{specifičnost} = \frac{450}{500} = 0,9$$

$$\text{pozitivna prediktivna vrijednost} = \frac{450}{500} = 0,9$$

$$\text{negativna prediktivna vrijednost} = \frac{450}{500} = 0,9$$

Opisani slučaj odgovara situaciji u bolničkoj populaciji.

Slučaj B

		STVARNO STANJE		
		Bolestan	Zdrav	
TEST	Pozitivan	90	90	180
	Negativan	10	810	820
		100	900	1000

$$\text{prevalencija} = \frac{100}{1000} = 0,1$$

$$\text{osjetljivost} = \frac{90}{100} = 0,9$$

$$\text{specifičnost} = \frac{810}{900} = 0,9$$

$$\text{pozitivna prediktivna vrijednost} = \frac{90}{180} = 0,5$$

$$\text{negativna prediktivna vrijednost} = \frac{810}{820} = 0,99$$

Opisani slučaj odgovara situaciji u epidemiološkoj (“otvorenoj”) situaciji. Prema rezultatima možemo očekivati mnogo lažno pozitivnih (niska PPV).

ROC ANALIZA

Valjanost testa je složeni pokazatelj. Ima dvije komponente: osjetljivost i specifičnost.

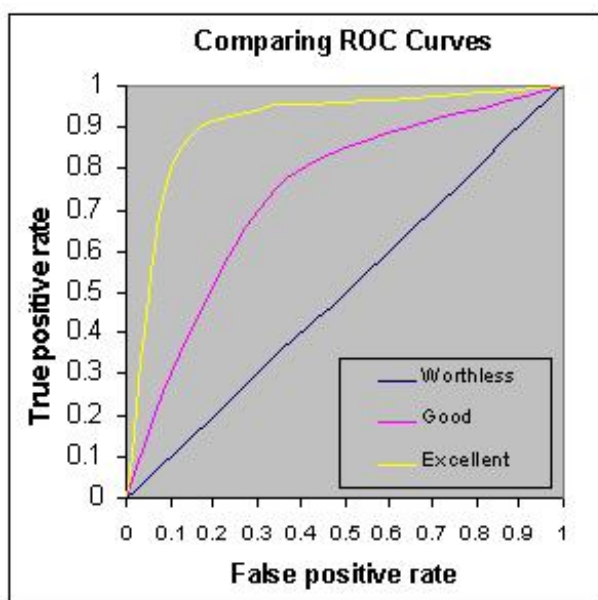
Kod kvantitativnih testova granica koja odvaja “test-pozitivne” od “test-negativnih” arbitrarno se određuje. Moguće je odrediti tu granicu na različitim razinama i tako definirati testove različite osjetljivosti i specifičnosti. To vrijedi i za testove čije su vrijednosti kvalitativne

Analiza osjetljivosti i specifičnosti testa ovisno o postavljanju granice koja odvaja “test-pozitivne” od “test negativnih” naziva se *ROC analiza* (engl. Receiver Operating Characteristic Curve).

ROC krivulja konstruira se pomicanjem granice normalnih vrijednosti. Određuje se senzitivnost i specifičnost na osnovu granica normalnih vrijednosti - prema strogim, prema manje strogim, manje blagim i blagim kriterijima.

Rezultati se prikazuju u koordinatnom sustavu:

- na apscisi : odnos lažno pozitivnih (1-specifičnost);
- na ordinati: odnos ispravno pozitivnih (osjetljivost)



Slika 1. ROC - krivulje

Izvor slike: <http://gim.unmc.edu/dxtests/roc3.htm>

Točnost testa (engl, accuracy) ovisi o tome kako dobro separira testirane grupe u one sa i one bez ispitivane bolesti.

Površina ispod ROC krivulje mjera je točnosti testa:

- 0,90-1 = izvrstan test
- 0,80-0,90 = dobar test
- 0,70-0,80 = osrednji test
- 0,60-0,70 = slabiji test
- 0,50-0,60 = test bez uspjeha.

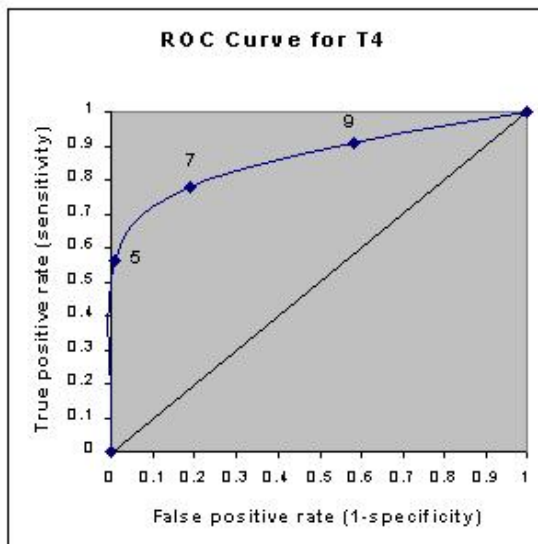
Primjer: T₄

Definirane su različite vrijednosti T₄ kao kriterij za razlikovanje hipotiroidnih od eutiroidnih bolesnika. Površina ispod T₄ ROC krivulje iznosi 0,86. Soga se T₄ može smatrati “dobrim” separatorom hipotiroidnih od eutiroidnih pacijenata.

T4 vrijednost	hipotiroidni	Eutiroidni
5 or less	18	1
5.1 - 7	7	17
7.1 - 9	4	36
9 or more	3	39
Totals:	32	93

Granična vrijednost	Osjetljivost	Specifičnost
5	0.56	0.99
7	0.78	0.81
9	0.91	0.42

Granična vrijednost	pozitivni	Lažno pozitivni
5	0.56	0.01
7	0.78	0.19
9	0.91	0.58



Slika 2. ROC – krivulje – T4

Izvor slike: <http://gim.unmc.edu/dxtests/roc3.htm>

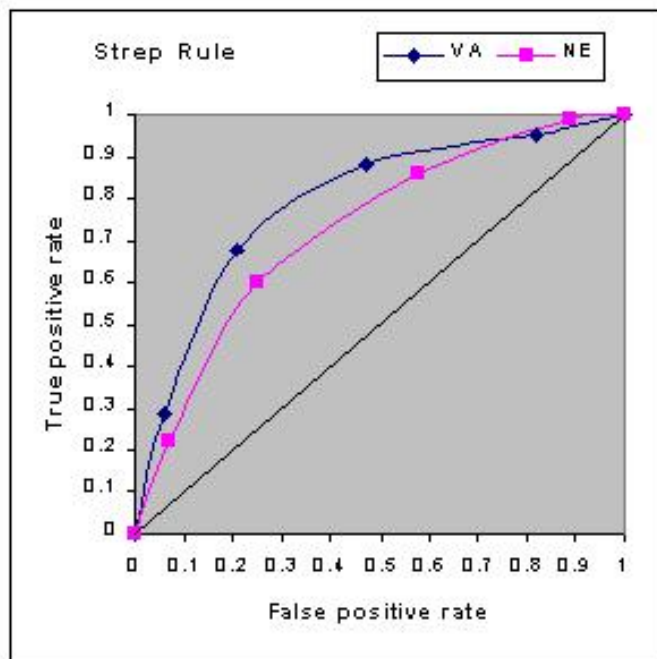
Primjer: klinička predikcija streptokoknog faringitisa

Klinički prediktori streptokoknog faringitisa:

- tonzilarni eksudat
- temperature
- adenopatija
- nedostatak kašlja

Krivulje su konstruirane računanjem osjetljivosti i specifičnosti za porast broja kliničkih znakova (od 0 do 4)

- Virginia (površina ispod krivulje = 0,78)
- Nebraska (površina ispod krivulje = 0,73)



Slika 3. ROC – krivulje- streptokokni faringitis

Izvor slike: <http://gim.unmc.edu/dxtests/roc3.htm>

Površina ispod krivulje mjeri diskriminaciju, odnosno sposobnost korektnog klasificiranje pacijenata u dvije grupe. Pacijenti su korektno klasificirani u dvije skupine: bolesni i zdravi. Slučajno se izabere jedan bolestan i jedan zdrav - bolestan bi trebao imati abnormalan, a zdrav normalan rezultat testa.

Površina ispod krivulje je postotak slučajno izabranih parova kod kojih je to točno. Pomicanjem granice može se podesiti specifičnost i osjetljivost tako da odgovaraju dijagnostičkom problemu.

Ako je lažno negativna dijagnoza ozbiljniji propust nego lažno pozitivna daje se prednost osjetljivijem i manje specifičnom testu. Ako je pak lažno pozitivna dijagnoza veći propust nego lažno negativna daje se prednost manje osjetljivom i više specifičnom testu.

Literatura:

1. *Ivanković D, i sur. Osnove statističke analize za medicinare. Zagreb: Medicinski fakultet Sveučilišta u Zagrebu, 1989.*
2. *Petrie A, Sabin C. Medical Statistics at a Glance (2nd Ed). Oxford: Blackwell Science Ltd, 2005.*
3. *Glantz. SA. Primer of Biostatistics (4th Ed). New York: McGraww-Hill: 1997.*
4. *Altman DG. Practical Statistics for Medical Research. London. Chapman & Hall, 1991.*
5. *Bland M. An Introduction to Medical Statistics (3rd Ed). Oxford: Oxford University Press, 2005.*
6. *Armitage P, Berry P. Statistical Methods in Medical Research. Oxford: Blackwell Science Ltd, 1994.*