

Kolegij: ROLP

Statistička terminologija I. - raspodjela i prikaz podataka

Doc. Dr. sc. Ksenija Baždarić

ožujak 2017.

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Neki temeljni statistički postupci u znanstvenom istraživanju

- odabir uzorka
- prikupljanje podataka
određivanje mjerne ljestvice
- obrada podataka
raspodjela podataka
opis podataka (prikazi, ...)

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

VRSTE PODATAKA (mjerne ljestvice)

KVANTITATIVNI (BROJČANI)	
OMJERNA BMI Krvni tlak Razina glukoze u krvi	INTERVALNA Broj djece Broj napadaja tjedno Dob (god)
KVALITATIVNI (SKUPNI, KATEGORIČKI)	
ORDINALNA Stupanj karcinoma Bolje, jednako, lošije Ne slažem se, neopredjeljen/a, slažem se (Likertova ljestvica (stavovi))	NOMINALNA Spol Pozitivan i negativan nalaz Krvna grupa Pripadnost rasi

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Kvalitativni podatci - nominalna ljestvica

- frekvencije i relativne frekvencije - N(%)
- grafički prikaz - stupčasti grafikoni

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Kvalitativni podatci - ordinalna ljestvica

- sve kao kod nominalne ljestvice ali i mjere središnjice i raspršenja - medijan, mod

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Medcalc - kategorijski podatci

Codes X	spol		
1	79	48.5%	
2	114	69.1%	
Total	193	100.0%	

Chi-square = 5.999
DF = 1
Significance level = P = 0.0144

Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

Kada se ispituje vrsta raspodjele?

... SAMO kada podaci vaših istraživanja pripadaju
INTERVALNOJ ili
OMJERNOJ mjerne ljestvici

- kontinuirani brojevi: vrijednosti hemoglobina, broj leukocita u krvi, tjelesna težina, stav prema radu...



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Zašto je važno znati kakva je raspodjela podataka?

- za odabir mjere središnjice i mjere varijabilnosti
- za odabir statističkog testa
- u slikovnom prikazu daje jasan uvid u podatke



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Mjere središnjice i odstupanja

Najčešće se upotrebljavaju:

- aritmetička sredina i standardna devijacija (normalna raspodjela, velik uzorak)
- medijan i raspon (percentili, interkvartili)
- interval pouzdanosti (aritmetička sredina, medijan)

Rjeđe se upotrebljavaju:

- varijanca, raspon, standardna pogreška, mod



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Mjere središnjice

Prikazuju prosječnu vrijednost neke skupine rezultata

- Aritmetička sredina
- Medijan (središnja vrijednost)
- Mod (najčešća vrijednost)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Aritmetička sredina (M, \bar{X})

- srednja ili prosječna vrijednost

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$$

- Prosječna visina studenata
- Srednja ocjena svih ispitanika na testu znanja



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Izračunavanje aritmetičke sredine

1	2	3	4	5
7	8	7	5	8

$$\bar{X} = \frac{5+7+7+8+8}{5} = \frac{35}{5} = 7$$



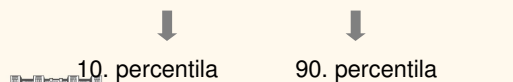
Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



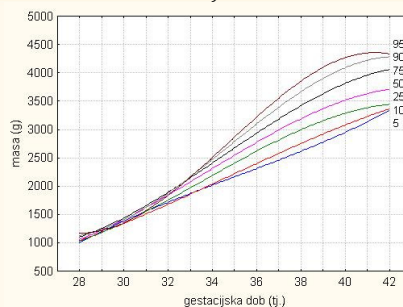
Percentil

- vrijednost odstupanja od medijana
- određena mjestom gdje se nalazi u nizu rezultata

1	2	...	5	...	26	...	47	...	50	51
10	18	...	34	...	52	...	71	...	81	89



Percentilna krivulja



Vrsta raspodjele određuje odabir mjere središnjice i mjere raspršenja

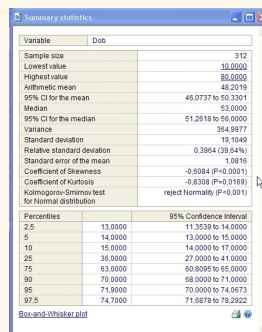
Normalna raspodjela: M i SD

Srednja vrijednost membranskog biopotencijala iznosila je $(70 \pm 3,5)$ mV.

Odstupajuća raspodjela: medijan i percentil

Prosječna životna dob ispitanika iznosila je 52 (34 - 71) godine.

Prikaz rezultata deskriptivne statistike u programu MedCalc



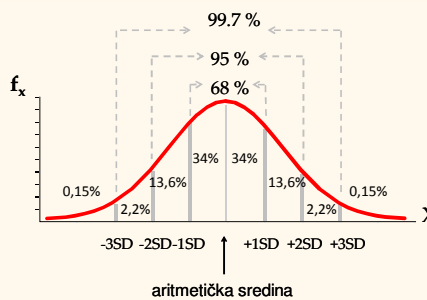
Krivulja normalne raspodjele ili Gaussova krivulja

- simetrična u odnosu na aritmetičku sredinu
- zvonolika
- oblika koji ovisi o aritmetičkoj sredini i standardnoj devijaciji

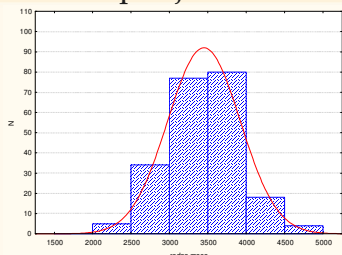
Svojstva normalne raspodjele:

- $(x \pm 1 SD) \rightarrow 68,26\%$
- $(x \pm 2 SD) \rightarrow 95,44\%$
- $(x \pm 3 SD) \rightarrow 99,73\%$

Normalna raspodjela



Normalna raspodjela



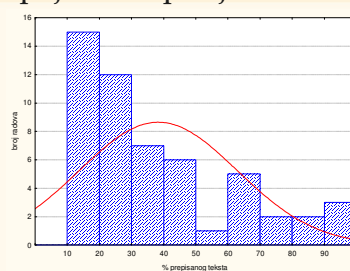
- podaci se opisuju s pomoću aritmetičke sredine i standardne devijacije



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Odstupajuća raspodjela



- podaci se opisuju s pomoću medijana i percentila (kvartila, raspona)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Kako ispitati normalnost raspodjele?

- Kolmogorov-Smirnovljevi test



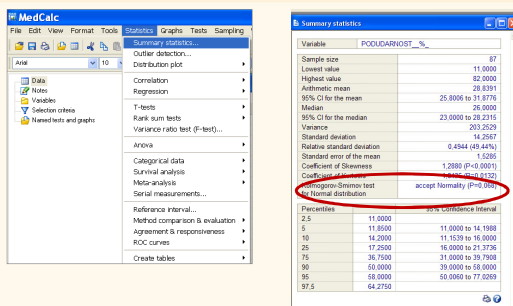
- $P \geq 0,05$ empirijska raspodjela je normalna
- $P < 0,05$ empirijska raspodjela značajno odstupa od normalne raspodjele



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Testiranje normalnosti raspodjele



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Vrsta raspodjele određuje odabir statističkog testa

	podatci	broj skupina	test
Normalna raspodjela	Nezavisni	2	t-test
	Zavisni		parni t-test
Odstupajuća raspodjela	Nezavisni	2	Mann-Whitney
	Zavisni		Wilcoxon



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Koji prikaz podataka odabrati?

Ovisi o

- podacima
- populaciji kojoj se prezentiraju
- našem znanju i vještinama

Najčešće se koriste

- tablice
- slike (grafikoni)



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Tablični prikaz podataka

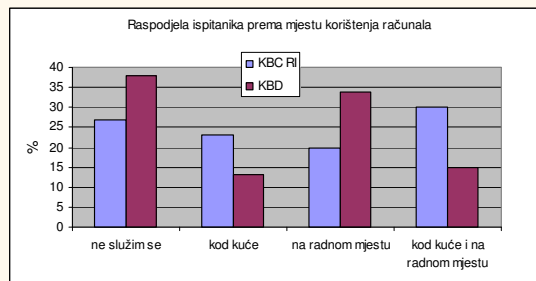
RB	skupina	dob	spol	...
1	kontrola	56	M	...
...
1	ispitanici	45	Ž	...
...



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Vertikalni stupčasti grafikon (2 skupine)

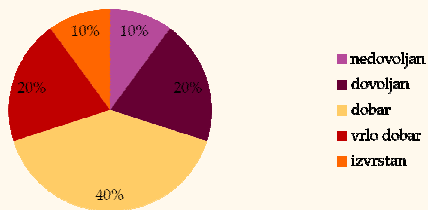


Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



Kružni dijagram

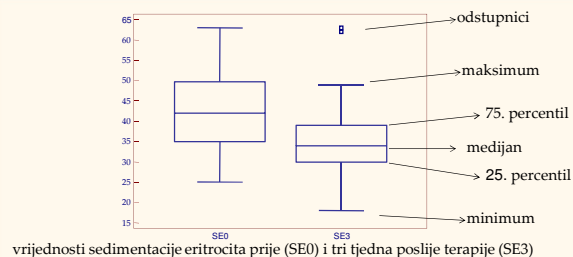
Raspodjela ocjena iz medicinske informatike



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku



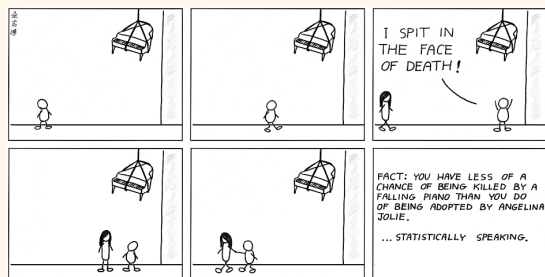
Grafikon okvira s ručicama



vrijednosti sedimentacije eritrocita prije (SEO) i tri tjedna poslije terapije (SES)

	Medijan	Minimum	Maximum	25. - 75. perc.
SEO	42	25	63	35 - 50
SES	34	18	63	30 - 63

I na kraju ...



Medicinski fakultet Sveučilišta u Rijeci Katedra za medicinsku informatiku

