

Medicinska informatika i statistika vježba 5

Testiranje razlika numeričkih podataka
te povezanost podataka.

•1

zavisna/nezavisna istraživanja

- ▶ razlike mjerenja neke varijable na dvije ili više skupina ispitanika -> nezavisni uzorci
- ▶ razlike dva ili više mjerenja neke varijable na istoj skupini ispitanika -> zavisni uzorci

•2

Priprema podataka

1. jedinica promatranja (*ispitanik, preparat, pokusna životinja, organ*)
2. varijable:
 - vrsta varijable (numerička, kategorička)
 - ljestvica mjerenja (nominalna, ordinalna, intervalna, omjerna)
 - za numeričke varijable, broj decimalnih mjesta

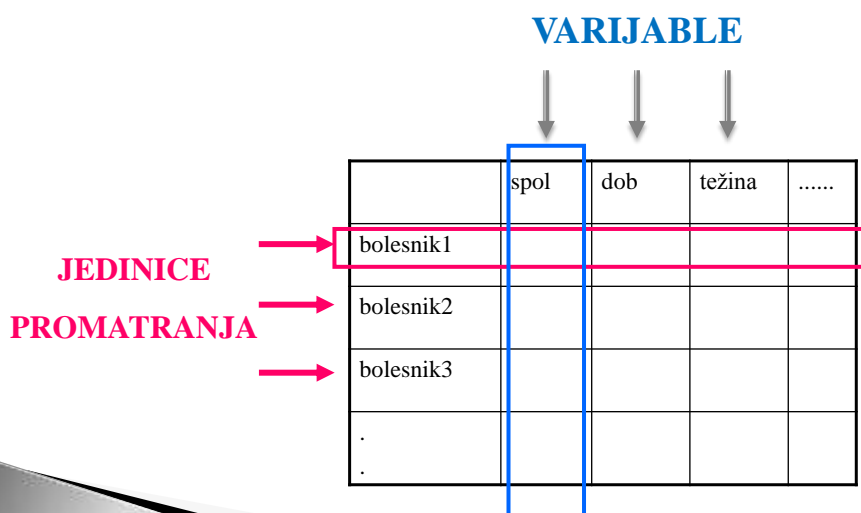
•3

Upis podataka

- ▶ numerički podatci
 - onako kako su izmjereni
- ▶ kategorički podatci
 - klasificirati u logičke, isključive skupine (prema problemu)

•4

Unos podataka



•5

Unos podataka o mjerenjima na nezavisnim skupinama

- nezavisne skupine = različiti ispitanici
(ispitanici koji pripadaju nekoj skupini ne pripadaju niti jednoj od preostalih skupina)
- za unos podataka o nekom mjerenju na nezavisnim skupinama ispitanika **UVIJEK** imamo 2 varijable (bez obzira koliko je skupina ispitanika):
 1. varijabla koja određuje pripadnost ispitanika pojedinoj skupini
 2. varijabla u koju unosimo vrijednost mjerenja za danog ispitanika

•6

Unos podataka o mjerenjima na nezavisnim skupinama

- npr. mjerenje dobi; skupine po spolu
- broj mogućih skupina: 2

varijabla koja sadrži vrijednost mjerenja

varijabla koja definira pripadnost skupini

	Dob	Spol
ispitanik1	35	M	
ispitanik2	37	M	
ispitanik3	32	M	
ispitanik4	33	Z	

•7

Unos podataka o mjerenjima na nezavisnim skupinama

- npr. mjerenje visine; skupine po razredu (osnovna škola)
- broj mogućih skupina: 8

varijabla koja sadrži vrijednost mjerenja

varijabla koja definira pripadnost skupini

	Visina	Razred
ispitanik1	110	2	
ispitanik2	140	2	
ispitanik3	100	1	
ispitanik4	176	7	

•8

Unos podataka o mjerenjima na zavisnim skupinama

- zavisne skupine = ponavljana mjerenja na ISTIM ispitanicima
- SVAKO mjerenje = JEDNA varijabla



koliko mjerenja toliko varijabli

•9

Unos podataka o mjerenjima na zavisnim skupinama

- npr. praćenje dnevnih varijacija sistoličkog tlaka; mjerenja u 6h, 10h, 14h, 18h, 22h

po jedna varijabla za svako mjerenje

	ST6	ST10	ST14	ST18	ST22
ispitanik1	120	135	140	180	160
ispitanik2	115	120	120	125	120
ispitanik3	140	145	150	150	180
ispitanik4	118	110	110	115	120

•10

STUDENTOV T-TEST

(t-test za nezavisne uzorke)

•11

ZADATAK 1

- ▶ Ispitivan je utjecaj sniženja tjelesne temperature na protrombinsko vrijeme. Izvršena su mjerenja PV na dvije skupine ispitanika. U jednoj skupini bilo je 7 ispitanika normalne temperature (kontrolna skupina). U drugoj skupini bilo je 9 ispitanika sa sniženom temperaturom (eksperimentalna skupina).

•12

Mjerenjem su dobiveni sljedeći rezultati (u sekundama):

kontrolna skupina (37°C)	eksperimentalna skupina (15°C)
7	12
8	9
9	11
6	10
8	8
10	9
8	7
	7
	8

• 13

▶ nezavisne skupine

2 varijable:

protrombinsko vrijeme
skupina

1 – kontrolna skupina

2 – eksperimentalna skupina

• 14

pvrijeme	skupina
7	1
8	1
9	1
6	1
8	1
10	1
8	1
12	2
9	2
11	2
10	2
8	2
9	2
7	2
7	2
8	2

• 15

Opis varijabli

protrombinsko vrijeme:

- numerička varijabla, omjerna ljestvica

skupina:

- kategorička varijabla, nominalna ljestvica

- ispitati normalnost raspodjele PV u svakoj skupini
- ispitati mjere sredine, raspršenja i oblika za PV u svakoj skupini

• 16

T-TEST DIFERENCIJA (t-test za zavisne uzorke)

•17

ZADATAK 2

Ispitivan je utjecaj alkohola na vrijeme reakcije vozača. Izvršeno je mjerenje vremena reakcije 10 vozača na standardne zadatke prije i nakon konzumacije određene količine alkohola.

•18

Mjerenjem su dobiveni slijedeći rezultati:

	prije	poslije
1	.68	.73
2	.64	.66
3	.68	.66
4	.82	.92
5	.58	.68
6	.80	.87
7	.72	.77
8	.65	.70
9	.84	.88
10	.73	.79

•19

- zavisne skupine

2 varijable:

prije

poslije

•20

T1	T2
,68	,73
,64	,66
,68	,66
,82	,92
,58	,68
,80	,87
,72	,77
,65	,70
,84	,88
,73	,79
,68	,73

•21

Opis varijabli

T1 (prije), T2 (poslije):

– numeričke, omjerna ljestvica

za obje varijable:

- ispitati normalnost raspodjele
- ispitati mjere sredine, raspršenja i oblika

•22

JEDNOSMJERNA ANALIZA VARIJANCE (One-way ANOVA)

•23

ZADATAK 3

Bolesnici s uznapredovalim stadijem raka želuca, bronhija, kolona, jajnika i dojke tretirani su novim lijekom. Svrha istraživanja je utvrditi je li preživljavanje bolesnika povezano sa zahvaćenim organom. Vrijeme preživljavanja (u mjesecima) dano je u tablici:

•24

želudac	bronhiji	kolon	jajnici	dojke
11	9	16	35	35
6	21	19	9	5
5	4	14	14	40
7	21	43	19	34
20	16	13	54	6
7	13	23	21	27
33	8	23		62
7	8	21		28
10	12	20		42
30	29	19		59
12	12	31		27
18	13	28		
20	6	19		
	15	13		
	12	10		
	8	4		
	16	17		

•25

● nezavisne skupine

5 skupina , ali **2 varijable:**

vrijeme

organ

1 - želudac

2 - bronhiji

3 - kolon

4 - jajnici

5 - dojke

•26

organ	vrijeme
1	11
:	:
1	20
2	21
:	:
2	16
3	16
:	:
3	17
4	35
:	:
4	21
5	35
:	:
5	27

•27

NEPARAMETRIJSKE METODE

•28

razlozi za primjenu:

- ▶ *distribucije varijabli očigledno nisu normalne*
- ▶ *uzorak je mali, a postoji sumnja na normalnost*
- ▶ *varijable su mjerene ordinalnom skalom*

•29

MODEL	PARAMETRIJSKI	NEPARAMETRIJSKI
dva nezavisna uzorka	<i>Studentov t - test</i>	<i>Mann-Whitney U test</i>
dva zavisna uzorka	<i>t-test diferencija</i>	<i>Wilcoxonov test sume rangova</i>
više nezavisnih uzoraka	<i>jednosmjerna analiza varijance</i>	<i>Kruskal - Wallis test</i>
više zavisnih uzoraka	<i>analiza varijance za ponavljana mjerenja</i>	<i>Friedmanov test (Friedmanova dvosmjerna analiza varijance rangova)</i>

•30

MANN-WHITNEY U TEST

(Wilcoxon-Mann-Whitney test)

- ▶ testira pripadaju li dva uzorka istoj populaciji (tj. imaju li iste medijane)

Slučajni uzorak $n_1=7$ iz populacije X dao je sljedeća opažanja:

3.7; -1.1; 2.6; 2.3; 4.1; 0.8; 3.9

a slučajni uzorak $n_2=5$ iz populacije Y opažanja:

4.6; 4; 5.3; 4.4; 3

Pripadaju li ovi uzorci istoj populaciji?

•31

WILCOXONOV TEST SUME RANGOVA

- ▶ testira hipotezu da uzorak pripada populaciji s određenom vrijednosti medijana, tj. da je medijan populacije razlika parova opažanja dvaju uzoraka jednak 0

Dvanaest parova jednojajčanih blizanaca podvrgnuto je psihološkom testiranju agresivnosti s ciljem istraživanja razlika u agresivnosti između prvorođenih i drugorođenih blizanaca. Dobiveni su sljedeći rezultati:

Par	1	2	3	4	5	6	7	8	9	10	11	12
Prv.	86	71	77	68	91	72	77	91	70	71	88	87
Drug.	88	77	76	64	96	72	65	90	65	80	81	72

•32

KRUSKAL-WALLIS TEST

- ▶ testira pripada li više uzoraka istoj populaciji

Ispitivane su 4 različite metode uzgoja kukuruza. Svaka je metoda probana na nekoliko oranica. Urodi po oranicama bili su:

Metoda 1	83	91	94	89	89	96	91	92	90	
Metoda 2	91	90	81	83	84	83	88	91	89	84
Metoda 3	101	100	91	93	96	95	94			
Metoda 4	78	82	81	77	79	81	80	81		

•33

FRIEDMANOV TEST

(Friedmanova dvosmjerna analiza varijance rangova)

- ▶ testira pripada li više zavisnih uzoraka istoj populaciji

Dvanaest slučajno izabranih studenata bilo je uključeno u pokus učenja. Svaki student je ispunjavao 4 testa znanja iz istog gradiva, jedan test za drugim. Bodovi na testu bili su (20 najviše moguće):

Student	1	2	3	4	5	6	7	8	9	10	11	12
Test 1	18	7	13	15	12	11	15	10	14	9	8	10
Test 2	4	6	14	10	11	9	16	8	12	9	6	11
Test 3	16	5	16	12	12	9	10	11	13	9	9	13
Test 4	20	10	17	14	18	16	14	16	15	10	14	16

•34

KORELACIJA

•35

Izmjerena je visina u centimetrima i vitalni kapacitet pluća (VC) u litrama 33 studentice prve godine. Dobiveni su sljedeći rezultati:

	180.6	4.74		155.0	2.20		174.2	4.27
	168.0	3.63		171.0	3.38		167.0	3.45
	163.0	3.40		171.5	3.82		162.0	2.88
	171.0	3.75		167.6	3.26		172.0	4.13
	177.0	4.23		160.2	2.63		161.0	2.90
	169.4	3.20		166.6	3.06		155.0	2.65
	161.0	2.90		167.0	3.52		162.0	3.12
	170.0	3.88		163.0	2.82		174.0	4.02
	158.0	2.40		172.0	3.41		161.0	2.80
	161.0	2.60		158.0	2.81		166.0	3.46
	163.0	2.72		165.0	3.07		166.0	3.26

Ocijenite postoji li povezanost visine i vitalnog kapaciteta pluća