

Najčešće statističke pogreške

doc. dr. sc. Ana-Maria Šimundić, specijalist medicinske biokemije

Klinički zavod za kemiju, Odjel za molekularnu dijagnostiku
Klinička bolnica "Sestre milosrdnice", Zagreb



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

1. POGRJEŠKA UZORKOVANJA

- ✓ slučajni uzorak
- ✓ reprezentativan

Primjer:



cilj: ustanoviti prosječnu visinu
studenata 1. godine studija u Zagrebu

ispitivana skupina: 35 studenata 1. godine DIF-a



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

2. POGRJEŠKA pri pripremi podataka

- ✓ provjeriti jesmo li ispravno unijeli podatke u tablice
- ✓ ispitati značajke raspodjele i načiniti deskriptivnu analizu

Primjer:



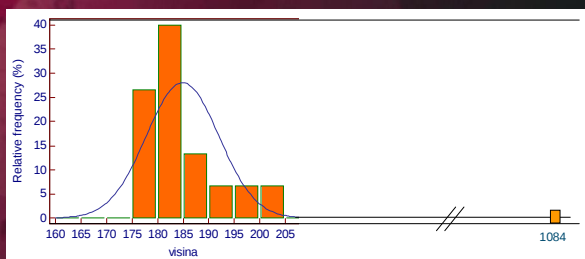
prikupili ste podatke (visina studenata 1.
godine DIF-a) i unijeli u MedCalc



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

	A	B	C	D	E	F	G	H	I	J	K
1	visina										
2	178										
3	179										
4	180										
5	181										
6	182										
7	183										
8	184										
9	185										
10	186										
11	187										
12	188										
13	189										
14	190										
15	191										
16	192										
17	193										
18	194										
19	195										
20	196										
21	197										
22	198										
23	199										
24	200										
25	201										
26	202										
27	203										
28	204										
29	205										
30	206										
31	207										
32	208										
33	209										
34	210										
35	211										

kako izgleda raspodjela?



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

deskriptivna analiza ?

Variable	visina
Sample size	30
Lowest value	178.0000
Highest value	211.0000
Arithmetic mean	184.6667
95% CI for the mean	182.2141 to 187.5193
Median	183.0000
95% CI for the median	180.0000 to 186.0000
Variance	50.4644
Standard deviation	7.1038
Relative standard deviation	0.03843 (3.84%)
Standard error of the mean	1.2370
Coefficient of Skewness	1.2657 (P=0.0063)
Coefficient of Kurtosis	0.5918 (P=0.3725)
Kolmogorov-Smirnov test for Normal distribution	accept Normality (P=0.108)

Percentiles	95% Confidence Interval
2.5	178.0000
5	178.0000
10	178.0000
25	178.0000 to 182.2312
75	187.0000 to 193.7688
90	199.0000
95	201.0000
97.5	201.0000



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

3. POGRJEŠKA u prikazu rezultata

- ☑ podatke valja prikazati s razumnom preciznošću



Primjer:

prosječna visina studenata iznosi ~~184,8667~~ cm



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

- ☑ izbjegavati postotke ako je uzorak mali

33 % umrlo

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

4. POGRJEŠKA pri prikazu rezultata

- ☑ ispravan izbor mjera središnjice i rasapa !

Primjer:

prosječna visina studenata iznosi:

~~$184 \pm 1,3$ cm~~
aritmetička sredina $\pm$ SEM

184 ± 7 cm
aritmetička sredina $\pm$ SD



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

5. POGRJEŠKA pri izboru testa

- ☑ osnovne pretpostavke za izvođenje testa moraju biti zadovoljene

Primjer:



cilj: ustanoviti razliku u prosječnoj visini studenata 1. godine studija u Zagrebu i Splitu

~~parni t-test~~

t-test



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

IZBOR STATISTIČKOG TESTA OVISI O:

- ☑ pitanju na koje želimo odgovoriti
- ☑ mornoj ljestvici
- ☑ vrsti obilježja
- ☑ značajkama raspodjele
- ☑ broju skupina
- ☑ broju mjerenja



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

OSNOVNE VRSTE TESTOVA



Pitanje	Statistička analiza
Što možemo reći o populaciji temeljem podataka iz uzorka?	Opisna analiza
Razlikuje li se neko obilježje između dvije ili više skupina?	Značajnost razlike
Jesu li neka dva obilježja povezana?	Korelacija
Možemo li predvidjeti jedno obilježje s pomoću drugih?	Regresija
Postoji li razlika u preživljavanju između dvije skupine?	Analiza preživljenja



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

ZNAČAJNOST RAZLIKE			
Mjerna ljestvica	intervalna	ordinalna	nominalna
Raspodjela	normalna	ne-normalna	ne-normalna
NEZAVISNE SKUPINE			
2	t-test	Mann-Whitney	Fisherov exact
≥ 3	ANOVA	Kruskal-Wallis	χ^2 test
ZAVISNE SKUPINE			
2	parni t-test	Wilcoxon test	McNemar test
≥ 3	RM ANOVA	Friedman ANOVA	Cochrane Q

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

6. POGRJEŠNO tumačenje P vrijednosti

☒ nulta hipoteza: nema razlike
☒ P vrijednost – vjerojatnost da opažena razlika postoji uz pretpostavku da je nulta hipoteza točna

→ statističko zaključivanje

	H_0 točna	H_1 točna
prihvaćamo H_0	ispravna odluka	pogrješka tipa 2 (β)
odbacujemo H_0	pogrješka tipa 1 (α)	ispravna odluka

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

7. POGRJEŠAN prikaz P vrijednosti

~~Visina studenata u Splitu i Zagrebu se statistički značajno razlikovala ($P < 0,05$).~~
~~Visina studenata u Splitu i Zagrebu se statistički značajno razlikovala ($P = NS$).~~
~~Visina studenata u Splitu i Zagrebu se statistički značajno razlikovala ($P < 0,000$).~~
~~Visina studenata u Splitu i Zagrebu se statistički značajno razlikovala ($P = 0,000001$).~~

ispravno → ($P < 0,001$).

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

8. POGRJEŠKA: smislenost rezultata

☒ osim statističke značajnosti razlike moraju biti i smislene, odnosno klinički značajne.

Primjer:
 Izmjerili smo prosječne visine studenata 1. godine studija u Zagrebu i Splitu.

Zagreb (N=10 000)	183 ± 7	$P < 0.001$ 
Split (N=10 000)	184 ± 7	

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Primjer:

Izmjerili smo prosječnu koncentraciju glukoze natašte u mladića i djevojaka na 1. godini studija.

mladići (N=7)	6,4 mmol/L	$P < 0.085$ 
djevojke (N=6)	4,6 mmol/L	

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Male razlike mogu biti statistički značajne, a potpuno besmislene.
 ako je uzorak veliki

Velike razlike mogu biti klinički značajne, a ne moraju ujedno biti statistički značajne.
 ako je uzorak mali

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

9. POGRJEŠNO tumačenje korelacije

☑ r ne označava uzročno-posljedičnu vezu već samo povezanost

Primjer:

Ustanovili smo da je koncentracija glukoze natašte u pozitivnoj korelaciji s visinom studenata ($r = 0,72$ i $P = 0,003$).

Visina uzrokuje porast ~~koncentracije~~ glukoze natašte.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

10. VIŠESTRUKO TESTIRANJE

Ako testiramo niz hipoteza, u konačnici ćemo negdje pronaći statistički značajnu razliku.



Je li ona stvarna?
Ili slučajna?

Ako izvedemo 20 testova na istom nizu podataka možemo očekivati barem jednu pogrešku tipa 1 (α).

ako je $\alpha = 0.05$



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Seminar



- 3 grupe
- svaka grupa dobije članak
- treba tražiti pogreške
- grupa koja nađe najviše pogrešaka dobija dodatne pluseve na konačnu ocjenu



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*