

# 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. POGREŠKA 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*

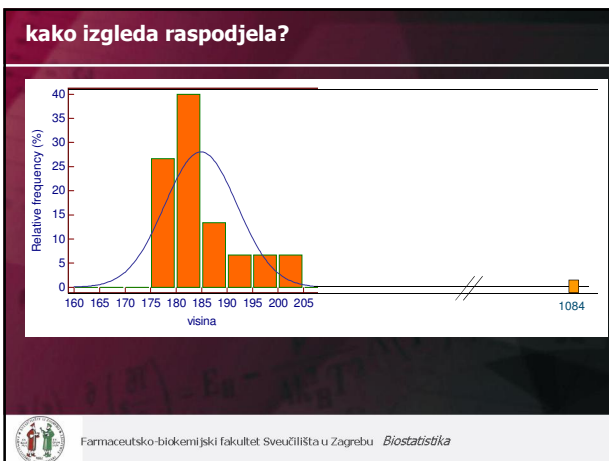
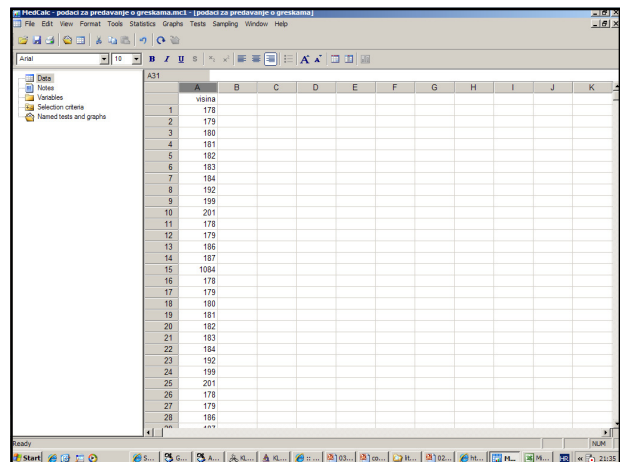
## 2. POGREŠKA pri pripremi podataka

- 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*



### deskriptivna analiza ?

Variable: visina		Variable: visina	
Sample size	30	Sample size	30
Lowest value	178.0000	Lowest value	178.0000
Highest value	201.0000	Highest value	201.0000
Arithmetic mean	184.6667	Arithmetic mean	184.6667
95% CI for the mean	182.2141 to 187.5193	95% CI for the mean	182.2141 to 187.5193
Median	183.0000	Median	183.0000
95% CI for the median	180.0000 to 186.0000	95% CI for the median	180.0000 to 186.0000
Variance	29996.6713	Variance	50.4644
Standard deviation	171.7193	Standard deviation	7.1038
Relative standard deviation	0.7647 (76.47%)	Relative standard deviation	0.03843 (3.843%)
Standard error of the mean	29.9967	Standard error of the mean	1.2970
Coefficient of Skewness	5.4809 (P<0.0001)	Coefficient of Skewness	1.2857 (P=0.0083)
Coefficient of Kurtosis	29.8759 (P<0.0001)	Coefficient of Kurtosis	0.5818 (P=0.3725)
Kolmogorov-Smirnov test for Normal distribution	reject Normality (P<0.001)	Kolmogorov-Smirnov test for Normal distribution	accept Normality (P=0.108)
<b>Percentiles</b>		<b>Percentiles</b>	
2.5	178.0000	2.5	178.0000
5	178.0000	5	178.0000
10	178.0000	10	178.0000
25	178.0000	25	179.0000
75	187.0000	75	187.0000
90	200.0000	90	189.0000
95	201.0000	95	201.0000
97.5	863.2500	97.5	201.0000

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

### 3. POGRJEŠKA u prikazu rezultata

- ☑ podatke valja prikazati s razumnom preciznosti



Primjer:

prosječna visina studenata iznosi ~~184,5667~~ cm



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

### 5. POGRJEŠKA pri izboru testa

- ☑ 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 ± 1,3 cm~~  
aritmetička sredina ± SEM

184 ± 7 cm  
aritmetička sredina ± 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 $\chi^2$ test
$\geq 3$	ANOVA	Kruskal-Wallis	$\chi^2$ test

### ZAVISNE SKUPINE

2	parni t-test	Wilcoxon test	McNemar test
$\geq 3$	RM ANOVA	Friedman ANOVA	Cochrane Q

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

### 6. POGRJESNO 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	<b>pogrješka tipa 2 (<math>\beta</math>)</b>
odbacujemo $H_0$	<b>pogrješka tipa 1 (<math>\alpha</math>)</b>	ispravna odluka

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

### Primjer:

**cilj: ustanoviti razliku u prosječne visine studenata 1. godine studija u Zagrebu i Splitu**

Variable	Split	Zagreb
Sample size	30	30
Arithmetic mean	175.3200	184.8667
95% CI for the mean	172.6170 to 178.0230	182.2141 to 187.5193
Variance	52.3884	50.4644
Standard deviation	7.2387	7.1038
Standard error of the mean	1.3216	1.2970

F-test for equal variances P = 0.920

T-test (assuming equal variances)

Difference	9.5467
Standard Error	1.8517
95% CI of difference	5.8401 to 13.2532
Test statistic T	5.156
Degrees of Freedom (DF)	58
Two-tailed probability	P < 0.0001

mala je vjerojatnost da je opažena razlika slučajna **snaga?**

nema dokaza da postoji razlika

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

### 7. POGRJESAN 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. POGRJESKA: smislenost rezultata

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

Primjer:

Izmjerali smo prosječne visine studenata 1. godine studija u Zagrebu i Splitu.

Zagreb (N=10 000)	183 ± 7
Split (N=10 000)	184 ± 7

$P < 0.001$

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

### Primjer:

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

mladići (N=7)	6,4 mmol/L
djevojke (N=6)	4,6 mmol/L

$P < 0.085$


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 ( $\alpha$ ),  
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*