

Normalna raspodjela

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

PONOVIMO...

iz Uzorak, populacija i nasumični odabir (M. Petrovečki)

1. Što u statističkome nazivlju razlikuje entitete promatranja, attribute (varijable) i podatke?
2. Što je inferencijalna statistika? Što je biostatistika?
3. Osobine populacije i uzorka u znanstvenom istraživanju.
4. Pojmovna, vremenska i prostorna odrednica populacije.
5. Biološka, vremenska i analitička varijabilnost.
6. Uzorak, uzorkovanje i pogreška uzorkovanja. Zašto je reprezentativnost važno svojstvo uzorka?
7. Vrste i osobine probabilističkih uzoraka.
8. Kako opisujemo slučajno biranje jedinka populacije u uzorak i kako se to praktično izvodi?
9. Veličina i povezanost uzoraka.
10. Kako tumačimo pojam vjerojatnosti i kako se ona matematički iskazuje?



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

PONOVIMO...

iz Varijable, mjerenje, podatci (L. Bilić-Zulle)

1. Što su stalne, a što promjenjive varijable? Navedite primjere.
2. Što su zavisne, a što nezavisne varijable? Navedite primjere.
3. Koji je ispravan način (kriterij) odabira ispitanika i razvrstavanja u skupine?
4. Koje su značajke kontrolne skupine?
5. Što je mjerenje? Što su mjerni instrumenti?
6. Koja su počela ispravnog mjerenja?
7. Kako nastaju i kako se izbjegavaju sustavne, a kako slučajne pogreške mjerenja?
8. Nabrojite i opišite značajke kvalitativnih mjernih ljestvica. Navedite primjere.
9. Nabrojite i opišite značajke kvantitativnih mjernih ljestvica. Navedite primjere.
10. Opišite razlike između kvalitativnih i kvantitativnih, diskretnih i kontinuiranih te jednostavnih i složenih podataka. Navedite primjere.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

STATISTIČKI PROGRAMI

- jednostavni za korištenje
- nije potrebno poznavati formule
- pružaju pomoć pri izboru testa (*Advisor Wizard*)
- daju izvještaje (*test report*) s tumačenjem
- imaju dobre priručnike (*Manual*)
- većina pod Windows okruženjem
- SigmaStat, MedCalc, Analyse-it, Statistika, SAS, SPSS, NCSS i dr.



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Advisor Wizard

What do you need to do?

- ☐ 1. Describe your data with basic statistics
- ☒ 2. Compare groups or treatments for significant differences
- ☐ 3. Predict a trend, find a correlation, or fit a curve
- ☐ 4. Determine the sample size for an experimental design
- ☐ 5. Determine the sensitivity of an experimental design

< Back

Next >

Cancel

Advisor Wizard

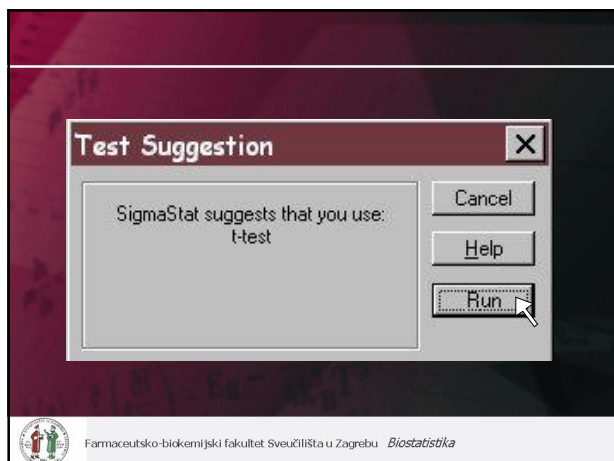
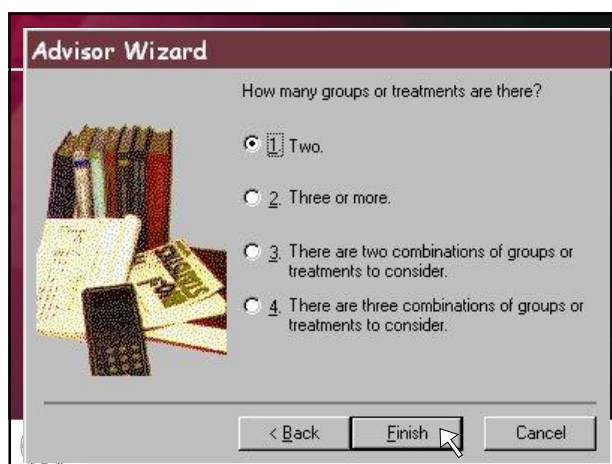
How are the data measured?

- ☒ 1. By numeric values (e.g. meters, degrees, etc.)
- ☐ 2. By order or rank (e.g., poor, fair, good, excellent)
- ☐ 3. By proportion or number of observations in categories (e.g., male vs. female)

< Back

Next >

Cancel



ŠTO MORAMO ZNATI?

- što želimo analizirati?
(osnovne značajke skupine, razlika, korelacija ?)
- način mjerenja nekog obilježja?
(vrste obilježja, mjerna ljestvica ?)
- vrsta raspodjele?
(varijanca, značajke raspodjele)
- kako izabrati odgovarajući test?
(nužni uvjeti)
- kako ispravno protumačiti rezultat testa?

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

OSNOVNE ZNAČAJKE SKUPINE

- **Mjere središnjice**
 - aritmetička sredina
 - medijan (centralna vrijednost)
 - mod (najučestalija vrijednost)
- **Mjere rasapa**
 - raspon
 - standardna devijacija

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

SIMBOLI

	Populacija	Uzorak
aritmetička sredina	μ	$\bar{X}$
standardna devijacija	σ	SD, s



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

ARITMETIČKA SREDINA

aritmetička sredina = $\frac{\text{suma svih rezultata}}{\text{broj rezultata}}$

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

- "prosjeak", težište rezultata
- zajednička aritmetička sredina !



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Koliko kava dnevno popije student 1. godine?

student →	1	2	3	4	5	6	7	8	9
broj kava →	1	2	2	3	3	3	4	4	5



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

MEDIJAN (centralna vrijednost)

- vrijednost koja se u nizu rezultata nalazi točno u **sredini**; položajna mjera (ne računamo ju)
- dijeli skup podataka na **dva jednaka dijela**
- točka od koje je **najmanja** suma svih odstupanja

1	2	3	4	5	6	7	8	9
1	2	2	3	3	3	4	4	5

N paran broj ⇒ medijan je aritmetička sredina dva srednja rezultata



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

MOD

- **najučestalija** vrijednost u nekom nizu podataka
- ako su rezultati grupirani u razrede ⇒ mod je aritmetička sredina razreda s najvećom frekvencijom

1	2	3	4	5	6	7	8	9
1	2	2	3	3	3	4	4	5

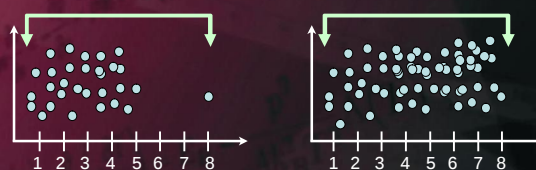


Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

RASPON

- razlika između maksimuma i minimuma
- najmanje informativna mjera varijabilnosti rezultata

1	2	3	4	5	6	7	8	9
1	2	2	3	3	3	4	4	5



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika



KVANTILI

- **SUSTAV KVANTILA:**
 - 3 kvartila dijele grupu na 4 jednaka dijela
 - 9 decila dijeli grupu na 10 jednakih dijelova
 - 99 centila dijeli grupu na 100 jednakih dijelova

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

STANDARDNA DEVIJACIJA

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{N - 1}}$$

- varijanca = s^2
- označava **prosječno odstupanje** od aritmetičke sredine
- smije se računati **samo uz aritmetičku sredinu** !

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Broj kava?

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{N - 1}}$$

student →	1	2	3	4	5	6	7	8	9
broj kava →	1	2	2	3	3	3	4	4	5

$$SD = \sqrt{\frac{(1-3)^2 + (2-3)^2 + (2-3)^2 + (4-3)^2 + (3-3)^2 + (3-3)^2 + (5-3)^2}{9-1}}$$

$$SD = \sqrt{\frac{4+1+1+1+1+1+4}{8}} = 1,2$$

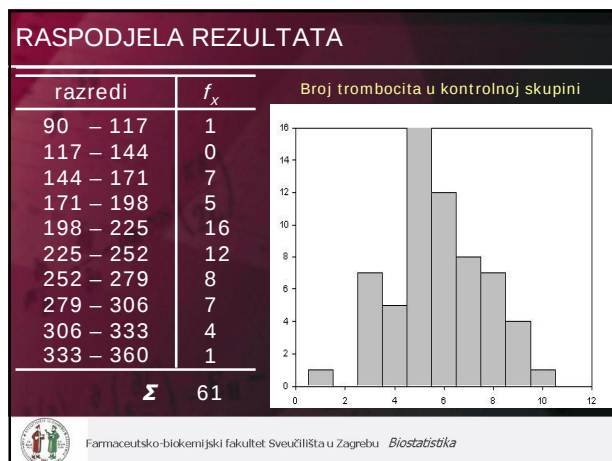
Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

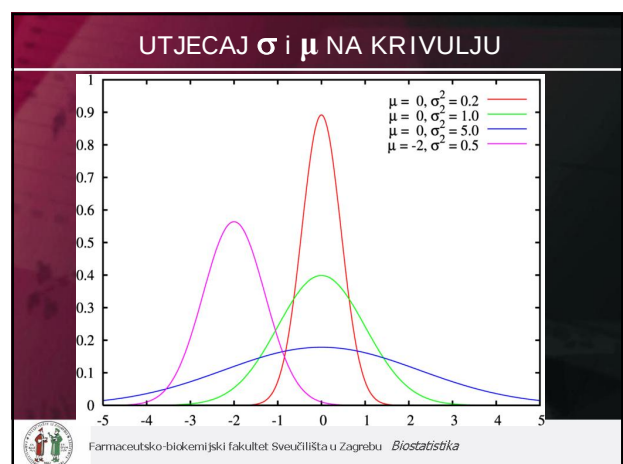
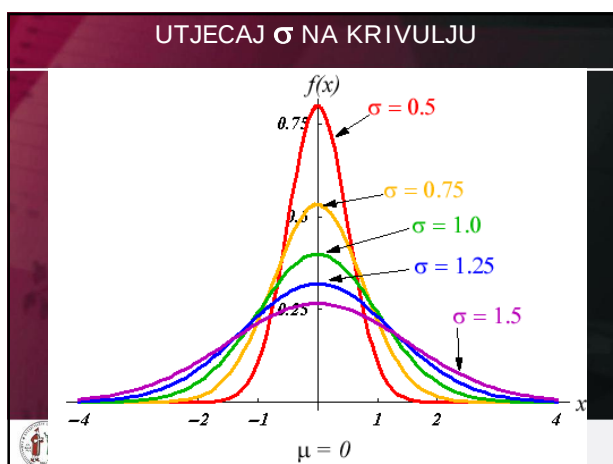
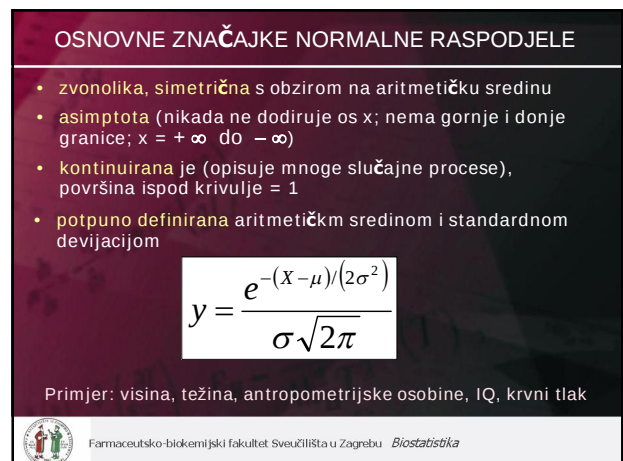
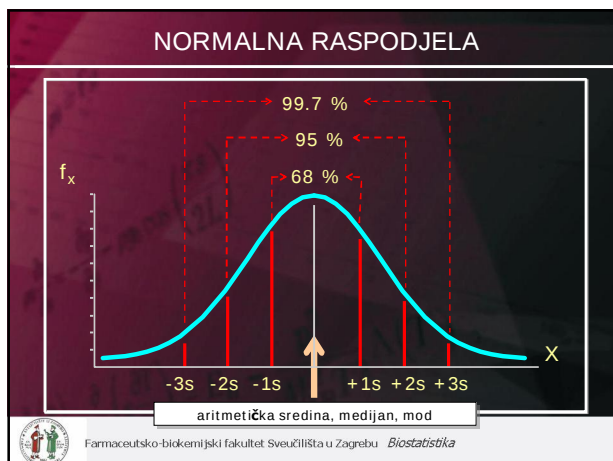
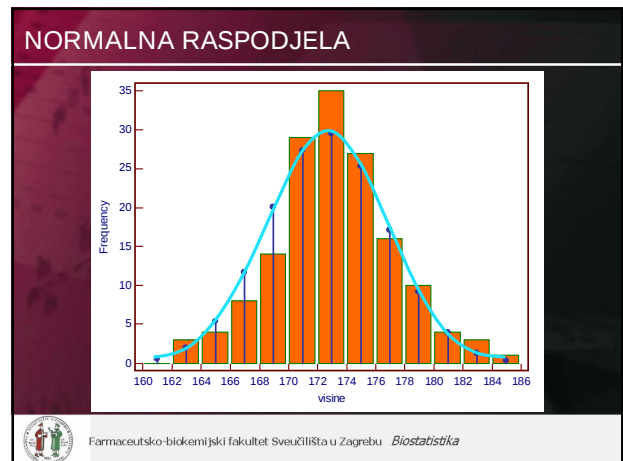
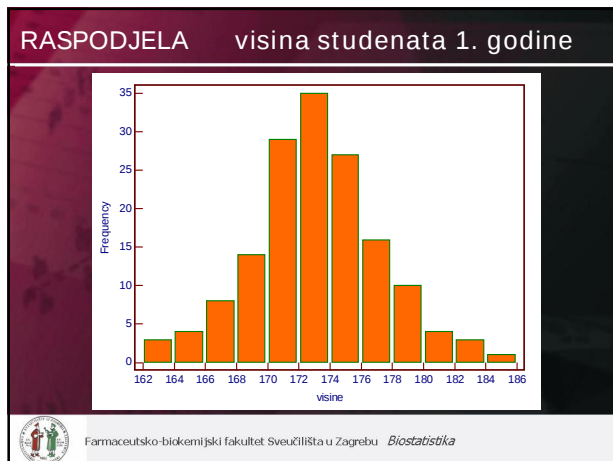
Prosječna visina studenata 1. godine?

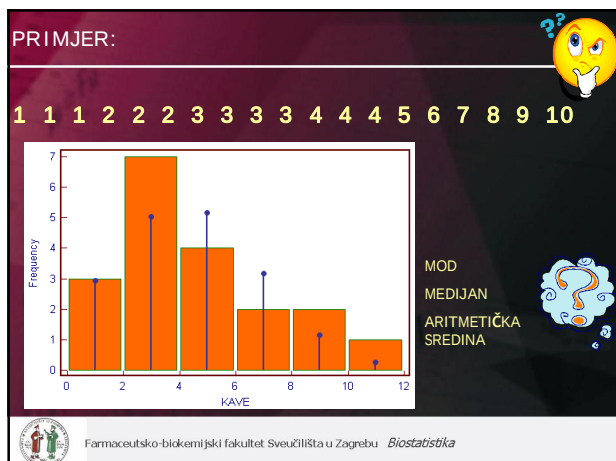
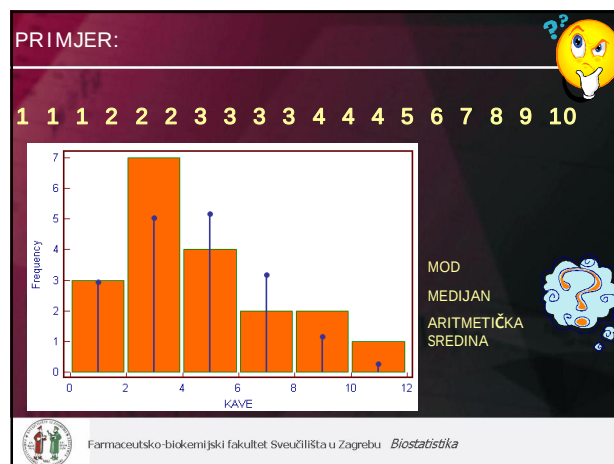
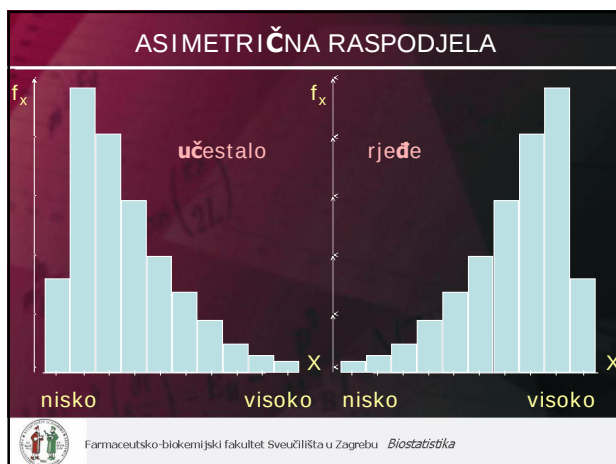
veličina uzorka	155
minimum	162
maksimum	184
aritmetička sredina	172
medijan	172
standardna devijacija	4,1
mod	172

raspodjela

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika







KAKO TESTIRATI RASPODJELU NA NORMALNOST ?

- ručno
- Kolmogorov-Smirnov, Shapiro-Wilk i dr...
- testira značajnost razlike teoretske i empirijske raspodjele
- Nul-hipoteza = nema razlike
- $P < 0.05 \Rightarrow$ empirijska raspodjela **nije normalna**
- $P > 0.05 \Rightarrow$ empirijska raspodjela **je normalna**

VRSTA RASPODJELE I VARIJANCA
UVJETUJU IZBOR STATISTIČKOG TESTA !!!

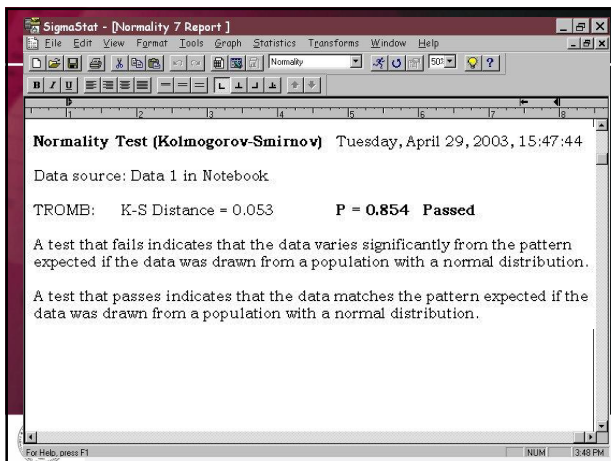
Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

KOLMOGOROV- SMIRNOVLJEV TEST

Test data column for normality

For Help, press F1

For Help, press F1



ZAŠTO TESTIRATI RASPODJELU NA NORMALNOST ?

- bolji **uvid** u podatke
- uvjetuje izbor **mjera središnjice i rasapa**
Normalna raspodjela $\Rightarrow$ aritmetička sredina $\pm$ SD
Ne-normalne raspodjela $\Rightarrow$ medijan, raspon
- uvjetuje izbor statističkog **testa**
Normalna raspodjela $\Rightarrow$ parametrijski testovi
Ne-normalne raspodjela $\Rightarrow$ neparametrijski testovi
(i **varijanca** uvjetuje izbor testa !!!)
- osnova je mnogih statističkih **testova**



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Kako nastaje STANDARDNA NORMALNA RASPODJELA?

NORMALNA RASPODJELA



srednja vrijednost populacije (μ) = 0
standardna devijacija populacije (σ) = 1

STANDARDNA NORMALNA RASPODJELA

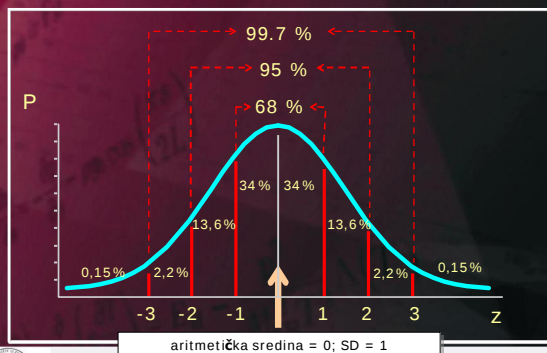
$$Z = (X - \mu) / \sigma$$

Z nam govori koliko standardnih devijacija je udaljena neka vrijednost (X) od aritmetičke sredine



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

STANDARDNA NORMALNA RASPODJELA



raspodjela frekvencija $\triangleright$ raspodjela vjerojatnosti

Primjer:



Pretpostavke:

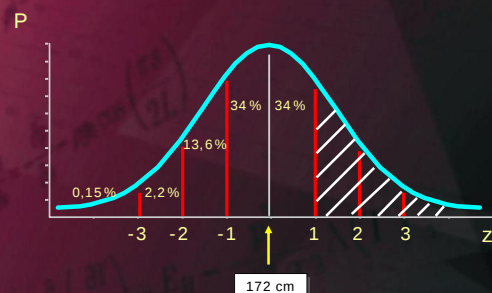
- visina studenata 1. godine u Hrvatskoj slijedi normalnu raspodjelu
- prosječna visina je 172 cm, a standardna devijacija je 4 cm.

Kolika je vjerojatnost da potpuno slučajnim izborom odaberemo studenta koji je viši od 176 cm?



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Vjerojatnost da potpuno slučajnim izborom odaberemo studenta koji je viši od 176 cm je:



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

PITANJA ZA PROVJERU ZNANJA



1. Što je deskriptivna analiza?
2. Što su mjere središnjice i rasapa?
3. Definirajte aritmetičku sredinu, mod i medijan.
4. Definirajte standardnu devijaciju.
5. Što je raspon? Što je interkvartilni raspon?
6. Koje su osobine normalne raspodjele?
7. Koliko je vrijednosti nekog skupa podataka obuhvaćeno s dvije standardne devijacije?
8. U kojem su odnosu aritmetička sredina, mod i medijan u normalnoj raspodjeli?
9. U kojem su odnosu aritmetička sredina, mod i medijan u asimetričnoj raspodjeli?
10. Što je standardna normalna raspodjela?



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*