

MJERILA KOJA OPISUJU DISTRIBUCIJU

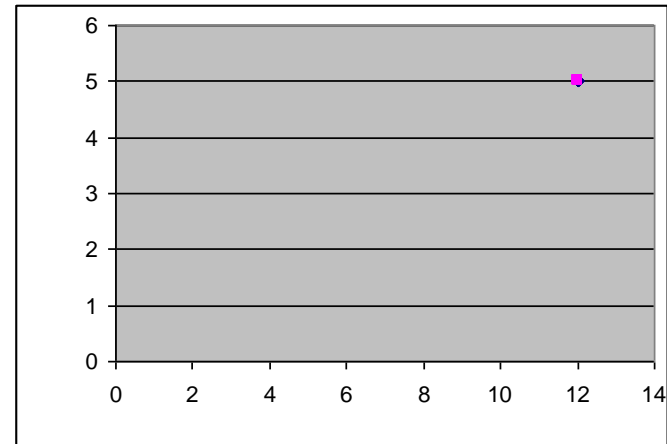


- ◆ Izgled **histograma ili poligona** daje sliku distribucije varijanata u uzorku
- ◆ Opis distribucije treba numeričke pokazatelje
- ◆ Procesi koji uvjetuju grupiranje varijanata oko sredine
- ◆ Procesi koji uvjetuju raspršivanje varijanata unutar varijacijske širine



Sve varijante uzorka su iste

x
12
12
12
12
12
12



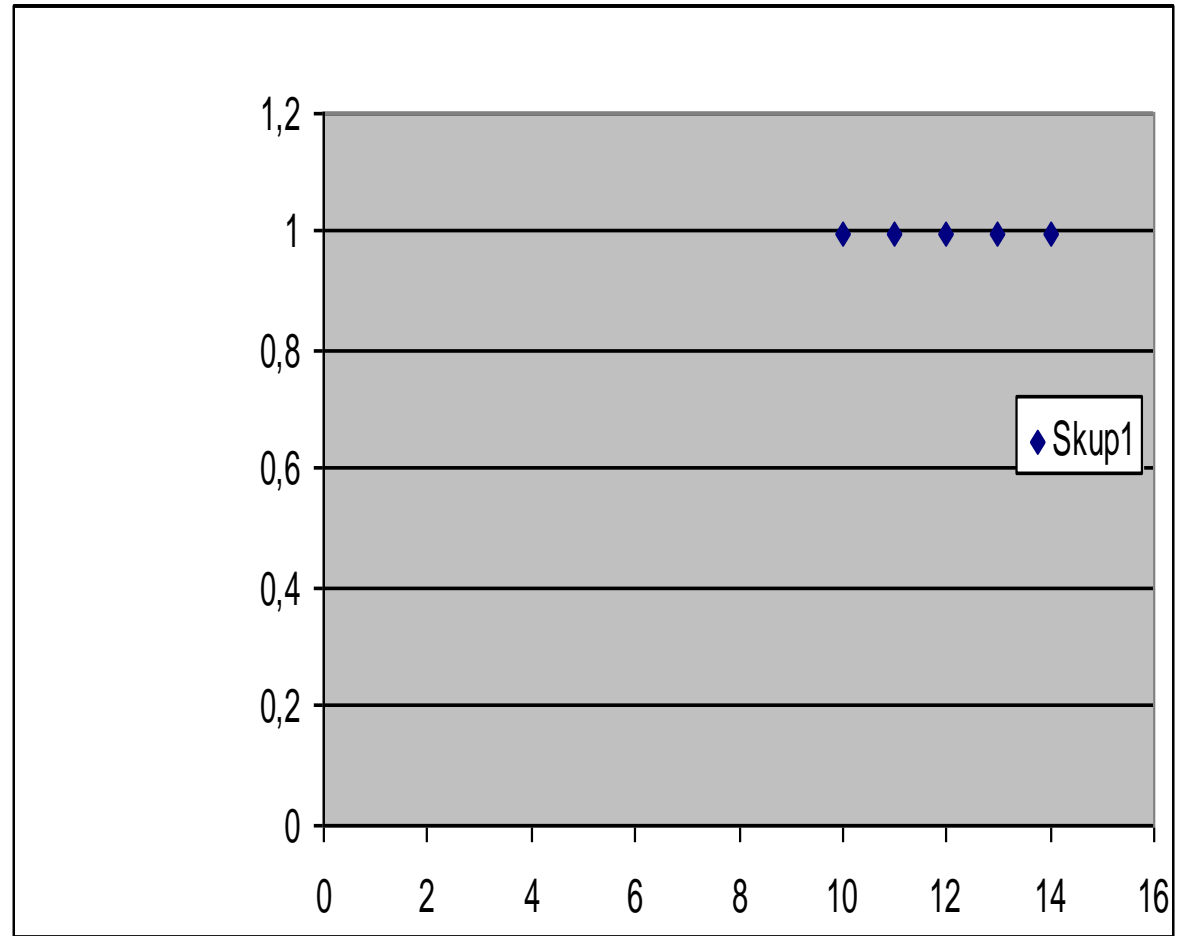
Distribucija vrijednosti varijable

- distribucija kada se pri svakom mjerenju dobije isti rezultat



Sve varijante uzorka su različite

x	f
10	1
11	1
12	1
13	1
14	1



◆ **DISPERZIJA** – VARIJABILNOST – RAZLIČITOST

◆ **CENTRALNA TENDENCIJA** – GRUPIRANJE OKO
SREDINE – SLIČNOST

◆ PROCESI SE MJERE: MJERILA SREDINE I
MJERILA DISPERZIJE



Mjerila centralne tendencije – Mjerila sredine

“Bolje je nešto uraditi temeljito, nego svašta površno.”(Non multa, sed multum.)

Latinska poslovice



Pokazatelji sredine distribucije

Prosječne ili srednje vrijednosti:

- Aritmetička sredina
- Modus
- Medijana
- Harmonijska sredina
- Geometrijska sredina



Aritmetička sredina (\bar{x})

- ◆ Prosječna vrijednost svih varijanata
- ◆ Računa se zbrajanjem svih podataka i dijeljenjem sume sa ukupnim brojem varijanata u uzorku



Pojedinačna opažanja u uzorku - x_i

$x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$

ili

x_1, x_2, \dots, x_n



$$\sum_{i=1}^{i=n} x_i$$

ili

$$\sum_{i=1}^{i=n} x_i = x_1 + x_2 + \dots + x_n$$



$$\sum_{i=3}^{i=8} x_i$$

$$\sum x$$



Aritmetička sredina

$$\bar{x} = \frac{\sum x}{n}$$



Karakteristike aritmetičke sredine

1. Uzima u obzir sve brojčane vrijednosti.

$$2. \sum (x - \bar{x}) = 0$$

$$3. \sum (x - \bar{x})^2 = \min$$

Medijana –medias-lat.-sredina

- ◆ Vrijednost one varijante koja se nalazi u sredini uzorka, ako su poredane rastućim redom



Medijana (Med)

- Medijana je centralna vrijednost u nizu podataka
 - 50% vrednosti je iznad, 50% ispod medijane
- Prije određivanja medijane podaci se urede po veličini
- Na Med ne utječu ekstremne vrednosti



medijana = 3



medijana = 3

Modus

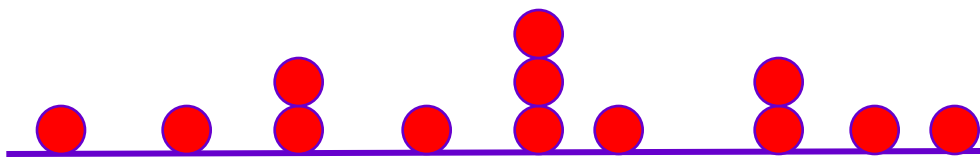
- ◆ Vrijednost one varijante koja se u uzorku najčešće pojavljuje
- ◆ Vrijednost varijante sa najvećom frekvencijom
- ◆ Za dovoljno velike uzorke

Unimodalni, bimodalni uzorci itd...



Modus (M_o)

- Na M_o ne utječu ekstremne vrijednosti
- U skupu može biti jedan ili više modusa
- Skup može biti bez modusa
- M_o se može odrediti i za numeričke i kategoričke podatke

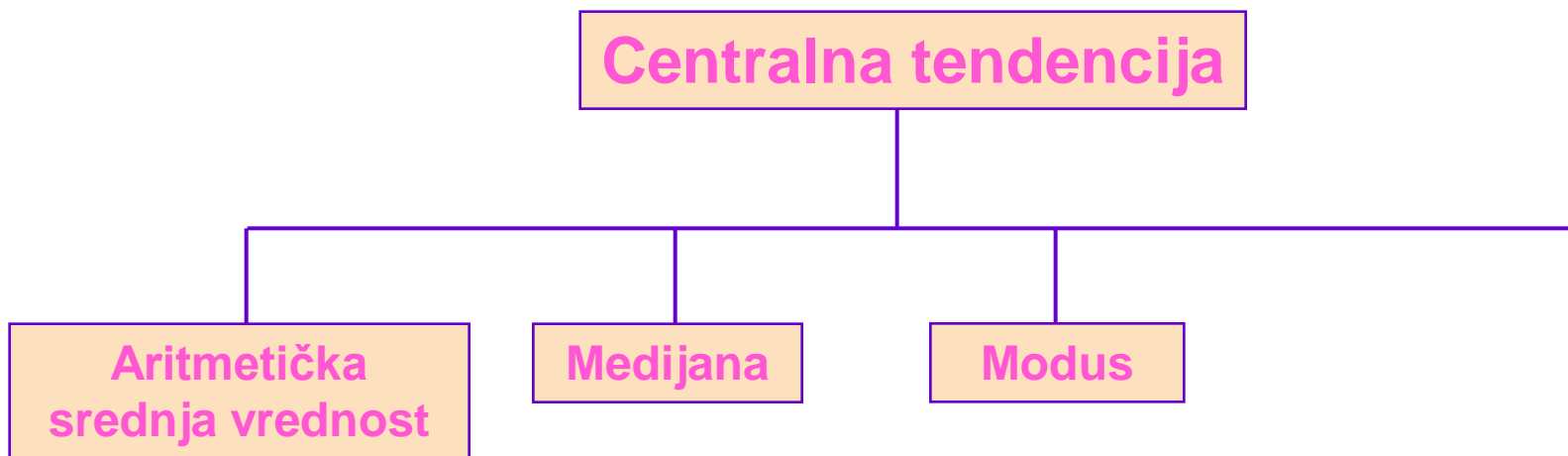


modus = 9

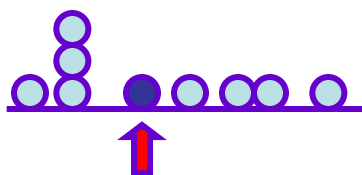


nema modusa

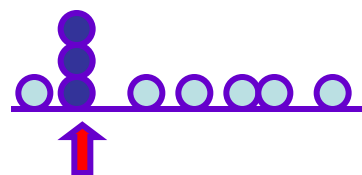
Mjerila centralne tendencije



$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$$



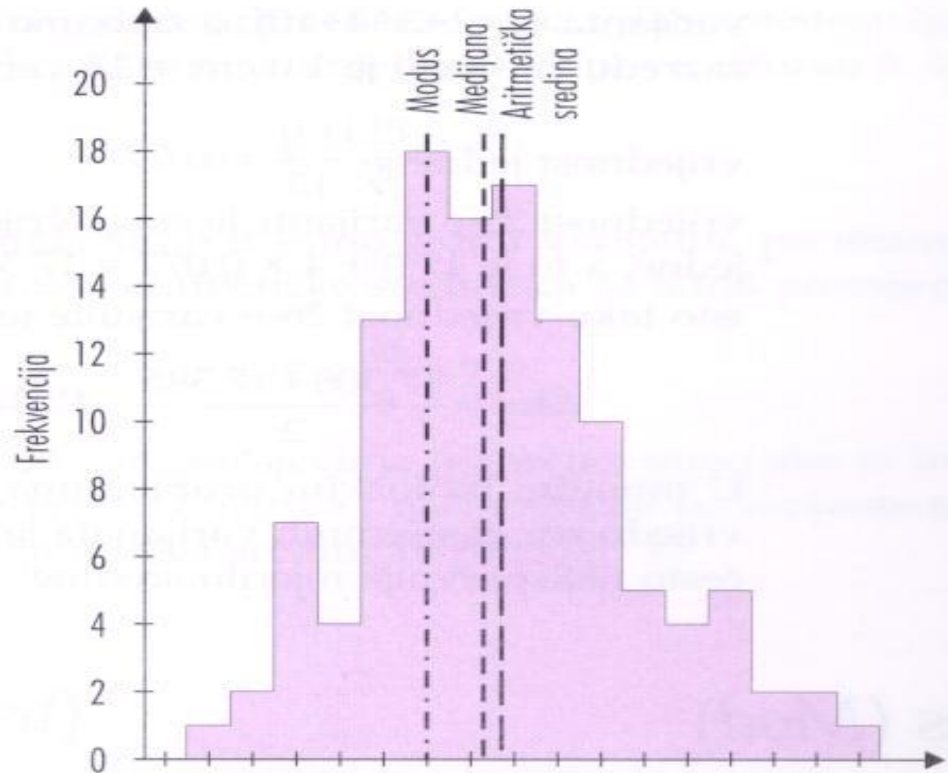
sredina rangiranih vrijednosti

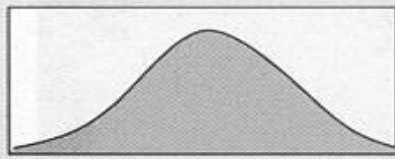


najfrekventnija vrijednost

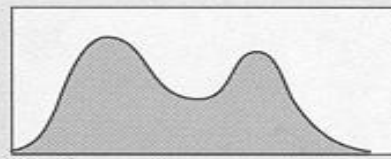
Nesimetrične distribucije

SLIKA 3. Aritmetička sredina (\bar{x}), medijana (Med) i modus (Mod) u nesimetričnoj distribuciji





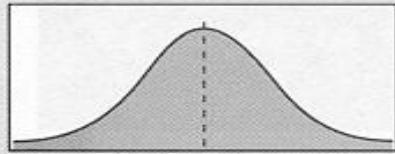
(a) Unimodal



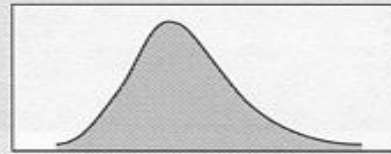
(b) Bimodal



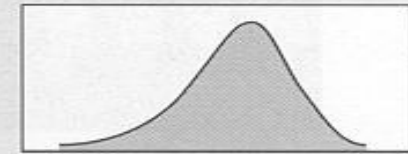
(c) Trimodal



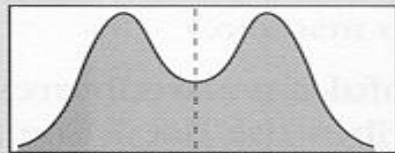
(d) Symmetric



(e) Positively skewed
(long upper tail)



(f) Negatively skewed
(long lower tail)



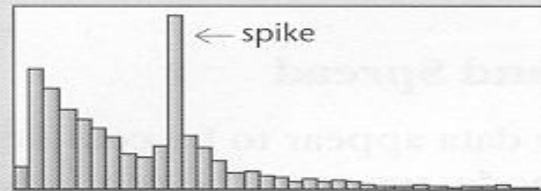
(g) Symmetric



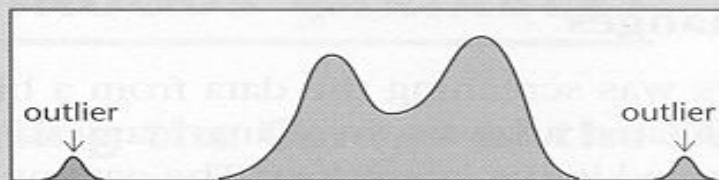
(h) Bimodal with gap



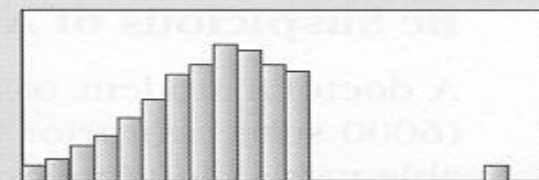
(i) Exponential shape



(j) Spike in pattern



(k) Outliers



(l) Truncation plus outlier