

**OPIS PODATKOV
ENE
SPREMENLJIVKE
frekvenčne porazdelitve in
mere srednje vrednosti**

as. dr. Nino RODE

Uni-Lj. Fakulteta za socialno delo

O ČEM BOMO GOVORILI

- **NAMEN OPISNE STATISTIKE**

- Kako opisati podatke

- **OPIS ENE SPREMENLJIVKE**

- Pregled porazdelitve podatkov:

- Frekvenčne porazdelitve

- Kako ugotoviti, kje najdemo podatke:

- Mere srednje vrednosti

NAMEN OPISNE STATISTIKE

- **Opisna Statistika** se ukvarja z opisovanjem **velike količine** podatkov
- Njen **namen** je **redukcija** podatkov:
 - s **čim manj** števil
 - **čim bolje** opisati podatke
- Z drugimi besedami: doseči **varčnost** – **parsimoničnost** prikazov podatkov

OPIS PODATKOV ENE SPREMENLJIVKE

- Pri eni spremenljivki dosegamo varčnost s:
 - štetjem vrednosti: **frekvenčna porazdelitev**
 - izračunom parametrov: mere **srednje vrednosti** in mere **razpršenosti** podatkov

FREKVENČNA PORAZDELITEV

- Primer: Povprečne ocene študentk in študentov FSD v prvem letniku od leta 1997 do leta 2008:

Veliko podatkov => 1009 ocen

Ocene	f
Od 6 do 6,9	135
Od 7 do 7,9	507
Od 8 do 8,9	328
Od 9 do 9,9	39
SKUPAJ	1009

FREKVENČNA PORAZDELITEV

- Primer: Povprečne ocene študentk in študentov FSD v prvem letniku od leta 1997 do leta 2008:

Struktura => odstotki: koliko enot bi bilo v celici, če bi bilo

Ocene	f	f%
Od 6 do 6,9	135	13,4
Od 7 do 7,9	507	50,2
Od 8 do 8,9	328	32,5
Od 9 do 9,9	39	3,9
SKUPAJ	1009	100,0

FREKVENČNA PORAZDELITEV

- Primer: Povprečne ocene študentk in študentov FSD v prvem letniku od leta 1997 do leta 2008:

Nalaganje => kumulativna frekvenca in odstotek

Ocene	f	f%	F	F%
Od 6 do 6,9	135	13,4	135	13,4
Od 7 do 7,9	507	50,2	642	63,6
Od 8 do 8,9	328	32,5	970	96,1
Od 9 do 9,9	39	3,9	1009	100,0
SKUPAJ	1009	100,0		

MERE SREDNJE VREDNOSTI

temeljne definicije

- **Modus**

- *Točka, okoli katere so vrednosti **najgostejše** (največja gostota porazdelitve) – pri nominalni spremenljivki modaliteta, v kateri je **največ** enot*

- **Mediana**

- *Točka, od katere je NATNČNO POL vrednosti **večjih** in NATNČNO POL vrednosti **manjših***

- **Aritmetična sredina**

- *Točka, v kateri se **oddaljenosti** vseh točk od nje **izenačijo**, težišče vrednosti, (matematično upanje, če ima vsaka vrednost enako verjetnost pojavljanja) vsaka enota **prispeva enak delež** k njeni vrednosti*

MERE SREDNJE VREDNOSTI

Kako deluje aritmetična sredina



MERE SREDNJE VREDNOSTI

Imamo podatke:

1, 2, 11, 3, 4, 8, 4, 7, 4, 5, 6, 0, 6, 1, 2, 7, 5, 9, 4, 3

Za njih določmo mere srednje vrednosti

MODUS

Neurejeni podatki:

- Najpametneje je podatke urediti v ranžirno vrsto (po velikosti) in najti najpogostejši podatek.

0, 1, 1, 2, 2, 3, 3, 4, 4, 4, 4, 5, 5, 6, 6, 7, 7, 8, 9, 11

- V našem primeru je to 4.

MODUS

Iz frekvenčne porazdelitve:

	f	f%	F	F%
0 - 2	5	25	5	25
3 - 5	8	40	13	65
6 - 8	5	25	18	90
9 - 11	2	10	20	100
Skupaj	20	100		

- Najprej najdemo modalni razred (razred v katerem je največ enot)
- Nato izračunamo modus po formuli:

$$M_o = x_{M_o} + \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})} \cdot i$$

MEDIANA

Neurejeni podatki:

- Iz neurejenih podatkov mediane ne računamo, temveč jo določamo. Podatke uredimo v ranžirno vrsto (uredimo po velikosti ali učenno po rangu):

0, 1, 1, 2, 2, 3, 3, 4, 4, 4, 4, 5, 5, 6, 6, 7, 7, 8, 9, 11

- Nato določimo točko, ki razdeli podatke točno na polovico:

tule



0, 1, 1, 2, 2, 3, 3, 4, 4, 4 4, 5, 5, 6, 6, 7, 7, 8, 9, 11

- Če imamo liho število podatkov, je to realni podatek, če je število podatkov sodo, pa izračunamo sredino med podatkoma, v našem primeru $(4 + 4) / 2 = 4$.

MEDIANA

Iz frekvenčne porazdelitve:

	f	f%	F	F%
0 - 2	5	25	5	25
3 - 5	8	40	13	65
6 - 8	5	25	18	90
9 - 11	2	10	20	100
Skupaj	20	100		

- Najprej najdemo **medianski razred** (razred v katerem leži mediana) To je razred, kjer kumulativni procent F% prvič preseže 50%
- Mediano izračunamo po formuli:

$$Me = x_{Me;min} + \frac{\frac{n}{2} - F_{Me-1}}{f_{Me}}$$

ARITMETIČNA SREDINA

Neurejeni podatki:

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

- x_i je vsak podatek (i je zaporedna številka podatka in je od 1 do n), n je število podatkov.

ARITMETIČNA SREDINA iz frekvenčne porazdelitve:

	x_k	f	f%	F	F%
0 - 2	1	5	25	5	25
3 - 5	4	8	40	13	65
6 - 8	7	5	25	18	90
9 - 11	10	2	10	20	100
Skupaj		20	100		

$$M = \bar{x} = \frac{\sum_{i=1}^k (x_k \cdot f_k)}{n} = \frac{(1 \cdot 5) + (4 \cdot 8) + (7 \cdot 5) + (10 \cdot 2)}{20} = \frac{5 + 32 + 35 + 20}{20} = \frac{92}{20} = 4,6$$

- x_k so sredine razredov (k je zaporedna številka razreda od 1 do r), n je število podatkov

O ČEM SMO GOVORILI

- ***NAMEN OPISNE STATISTIKE***
 - Parsimonično (varčno) opisati veliko podatkov
- ***OPIS ENE SPREMENLJIVKE***
 - **Frekvenčna porazdelitev**
 - Frekvence
 - Odstotki: struktura
 - Kumulativne frekvence
 - **Mere srednje vrednosti** (mere položaja – lokacije)
 - Modus
 - Mediana
 - Aritmetična sredina