

MERE SREDNJE VREDNOSTI IN RAZPRŠENOSTI

doc. dr. Nino RODE

Univerza v Ljubljani, Fakulteta za socialno delo
nino.ode@fsd.uni-lj.si

1

NEKAJ TEMELJNIH POJMOV

- **Enota:** primer stvari, pojava ali koncepta, ki ga proučujemo
- **Spremenljivka:** (postvarjena) lastnost določene vrste enot, ki se med enotami spreminja
- **Opažanje:** določanje/ugotavljanje vrednosti (meritev) spremenljivke pri konkretni enoti v danem trenutku in kontekstu
- **Podatek:** rezultat opažanja
- **Sveženj podatkov:** več podatkov, ki se nanašajo na isto vrsto opažanj (enot in spremenljivk)

2

MERE SREDNJE VREDNOSTI temeljne definicije



3

POLOŽAJ IN RAZPRŠENOST PODATKOV

- Vsak podatek ima svojo **vrednost**, ki **opisuje** lastnost (spremenljivko) enote, ki ji pripada
- **Vrsto enot** lahko določimo samo tako, da ugotovimo določeno **pomembno podobnost** med enotami => => enote **določene vrste** imajo **podobne lastnosti** => => podatki o njih imajo **podobno vrednost**
- To **podobnost vrednosti** podatkov poimenujemo **položaj (lokacijo)** podatkov
- **Razlike med vrednostmi** podatkov (ki so posledica dejavnikov nepovezanih z določanjem vrste enot) poimenujemo **razpršenost** podatkov

4

MERE POLOŽAJA: MERE SREDNJE VREDNOSTI

- **Modus**: točka, kjer so podatki v svežnju **najbolj gosti**, pri diskretnih podatkih najpogostejši podatek
==> mera **horde**: zmagajo tisti, ki jih je največ
- **Mediana**: točka, od katere je natančno **polovica** podatkov **večja** in **polovica** podatkov **manjša**
==> **birokratska** mera: glavno je, da so podatki urejeni po vrsti, kakšen je posamezen podatek, ni važno
- **Aritmetična sredina**: točka, kjer so **odkloni** (razlike) podatkov od nje **v ravnotežju**, težišče podatkov
==> **demokratska** mera: vsak podatek prispeva enak delež k njenemu določanju

5

MERE RAZPRŠENOSTI

- Na podlagi **intervala** potrebnega, da zajamemo določen del podatkov:
 - **Razpon** podatkov: razdalja med največjim in najmanjšim podatkom, interval, v katerem so vsi podatki
 - **Kvartilni razmik**: razdalja med prvim in tretjim kvartilom, interval, v katerem je srednja polovica podatkov

6

MERE RAZPRŠENOSTI

- Na podlagi **odklonov** od srednje vrednosti:
 - **Povprečni absolutni odklon**: povprečna vrednost (aritmetična sredina) absolutnih odklonov (razlik) podatkov od aritmetične sredine
 - **Varianca in standardni odklon**: povprečna vrednost (aritmetična sredina) kvadratov odklonov (razlik) podatkov od aritmetične sredine (varianca) in njen kvadratnikoren (standardni odklon)

7

DOLOČANJE MER SREDNJE VREDNOSTI

Primer: Za 20 študentk smo dobili naslednje podatke o številu prijateljev/ic, s katerimi se srečujejo vsaj enkrat na teden:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	11	3	4	8	4	7	4	5	6	0	6	1	2	7	5	9	4	3

Frekvenčna porazdelitev z razredi širokimi tri ljudi je takšna:

Št. prijateljev	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		

Podatke bomo uporabili za izračun mer srednje vrednosti in razpršenosti

8

DOLOČANJE MER SREDNJE VREDNOSTI

Modus (oznaka M_o):

- Modus določamo, ne računamo. Najprej podatke uredimo v skupine po vrednostih:

12	1	14	2	15	4	20	5	7	9	19	10	17	11	13	16	8	6	18	3
0	1	1	2	2	3	3	4	4	4	4	5	5	6	6	7	7	8	9	11

Nato najdemo vrednost, ki se največkrat pojavlja; v našem primeru:

- vrednost 4, ki jo imajo štirje študentke: 5., 7., 9. in 19. študentka,

$$M_o = 4$$

9

DOLOČANJE MER SREDNJE VREDNOSTI

Mediana (oznaka **Me** ali **Md**):

- Mediano določamo, ne računamo. Najprej podatke uredimo v ranžirno vrsto (uredimo po velikosti – po rangju):

12	1	14	2	15	4	20	5	7	9	19	10	17	11	13	16	8	6	18	3
0	1	1	2	2	3	3	4	4	4	4	5	5	6	6	7	7	8	9	11

Nato najdemo vrednost, od katere je pol podatkov (tu 10) večjih, pol (tu 10) pa manjših; v našem primeru:

- med 10. in 11. podatkom po velikosti,
- med podatoma za 9. in 19. študentko,
- med vrednostima 4 in 4, torej $(4 + 4)/2 = 4$

$$\mathbf{Me = 4}$$

10

DOLOČANJE MER SREDNJE VREDNOSTI

Aritmetična sredina (oznaka **M** ali \bar{x}):

- Aritmetično sredino izračunamo tako, da seštejemo vse vrednosti podatkov in jih delimo s številom podatkov:

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

- V našem primeru imamo 20 podatkov:

$$\bar{x} = \frac{(1+2+11+3+4+8+4+7+4+5+6+0+6+1+2+7+5+9+4+3)}{20} = \frac{92}{20} = 4,6$$

$$\mathbf{M = 4,6}$$

11

DOLOČANJE MER RAZPRŠENOSTI

- Razpon: najdemo največji in najmanjši podatek in izračunamo razliko med njima

12	1	14	2	15	4	20	5	7	9	19	10	17	11	13	16	8	6	18	3
0	1	1	2	2	3	3	4	4	4	4	5	5	6	6	7	7	8	9	11

$$\mathbf{MAX - min = 11 - 0 = 11}$$

12

DOLOČANJE MER RAZPRŠENOSTI

- Kvartilni razmik (Q): najdemo kvartile in izračunamo razliko med prvim in tretjim

12	1	14	2	15	4	20	5	7	9	19	10	17	11	13	16	8	6	18	3
0	1	1	2	2	3	3	4	4	4	4	5	5	6	6	7	7	8	9	11

Q_1

Q_2

Q_3

- **Kvartili:** točke, ki podatke razdelijo na četrtine:

- pod 1. kvartilom ($Q_1 = 2,5$) je 1/4 podatkov
- med 1. in 2. ($Q_2 = Me = 4$) kvartilom je 1/4 podatkov
- med 2. in 3. ($Q_3 = 6,5$) kvartilom je 1/4 podatkov
- nad 3. kvartilom je 1/4 podatkov

$$Q = Q_3 - Q_1 = 6,5 - 2,5 = 4$$

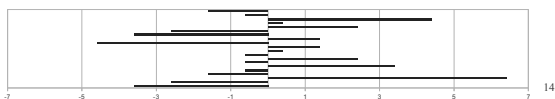
13

DOLOČANJE MER RAZPRŠENOSTI

- Odkloni: vsak podatek je bolj ali manj različen od srednje vrednosti podatkov. to razliko imenujemo odklon (d).

$$d = x - \bar{x}$$

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	11	3	4	8	4	7	4	5	6	0	6	1	2	7	5	9	4	3
-3,6	-2,6	6,4	-1,6	-0,6	3,4	-0,6	2,4	-0,6	0,4	1,4	-4,6	1,4	-3,6	-2,6	2,4	0,4	4,4	-0,6	-1,6



14

DOLOČANJE MER RAZPRŠENOSTI

- Večji odkloni pomenijo večjo razpršenost

Toda

- Če odklone vseh podatkov od aritmetične sredine seštejemo, **se med sabo uničijo** (aritmetična sredina je težišče podatkov)

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

15

DOLOČANJE MER RAZPRŠENOSTI

- Rešitev je v tem, da se **znebimo predznakov** odklonov.
- Lahko jih kar zanemarimo in upoštevamo le **absolutne vrednosti** odklonov. Iz njih izračunamo

Povprečni absolutni odklon (AD):

$$AD = \frac{\sum |x - \bar{x}|}{n}$$

16

DOLOČANJE MER RAZPRŠENOSTI

Povprečni absolutni odklon (AD):

- seštejemo absolutne vrednosti vseh podatkov (vrednosti brez predznakov) in jih delimo s številom podatkov. V našem primeru:

$$AD = \frac{\sum |x - \bar{x}|}{n} =$$

$$= \frac{(3,6+2,6+6,4+1,6+0,6+3,4+0,6+2,4+0,6+0,4+1,4+4,6+1,4+3,6+2,6+2,4+0,4+4,4+0,6+1,6)}{20} =$$

$$= \frac{45,2}{20} = 2,26$$

$$\mathbf{AD = 2,26}$$

17

DOLOČANJE MER RAZPRŠENOSTI

- Drugi način, da se **znebimo predznakov** odklonov je, da izračunano povprečje **kvadratov vrednosti** odklonov. Tako dobimo

Varianco (VAR, σ^2):

$$VAR = \sigma^2 = \frac{\sum (x - \bar{x})^2}{n}$$

18

DOLOČANJE MER RAZPRŠENOSTI

- Problem variance je, da je podana v **kvadratih enot**. Zato varianco korenimo in tako dobimo

Standardni odklon (SD, σ):

$$SD = \sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

19

DOLOČANJE MER RAZPRŠENOSTI

Varianco in standardni odklon (VAR ali σ^2 in SD ali σ) izračunamo tako:

- seštejemo kvadrate vrednosti vseh podatkov in jih delimo s številom podatkov. V našem primeru:

$$VAR = \sigma^2 = \frac{\sum (x - \bar{x})^2}{n} =$$

$$= \frac{(12,96+6,76+40,96+2,56+0,36+11,56+0,36+5,76+0,36+0,16+1,96+21,16+1,96+12,96+6,76+5,76+0,16+19,36+0,36+2,56)}{20}$$

$$= \frac{154,8}{20} = 7,74 \quad SD = \sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{7,74} = 2,78$$

VAR = σ^2 = 7,74

SD = σ = 2,78

20
