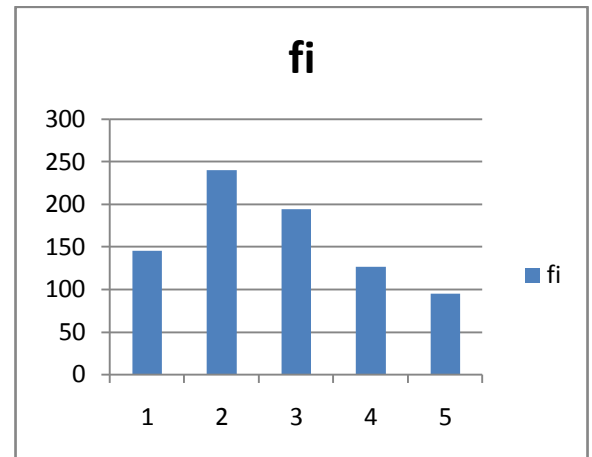


Deskriptivna statistika

1. U okviru istraživanja životnog ciklusa jedne vrste sisara ispitivao se broj mladunaca iz jednog legla. Na uzorku od 35 legala zabeležen je sledeći broj mladunaca po leglu: 1, 1, 2, 4, 3, 2, 2, 2, 2, 3, 4, 5, 5, 1, 2, 3, 2, 3, 2, 3, 2, 2, 1, 1, 1, 4, 3, 1, 5, 4, 3, 3, 2, 1.
- Bez grupisanja podataka izračunati mere centralne tendencije (srednju vrednost, medijanu, mod) i mere odstupanja (interval varijacije, standardnu devijaciju i koeficijent varijacije) koristeći Excel-ove funkcije.
 - Napraviti novu tabelu u kojoj će podaci biti grupisani. Podatke prikazati grafički (histogramom) i tabelom sa relativnim frekvencijama.
 - Izračunati mere centralne tendencije i mere odstupanja.

REŠENJE:

x_i	f_i	$x_i f_i$	$x_i^2 f_i$	F_i	p_i
1	8	8	8	8	22.86%
2	12	24	48	20	34.29%
3	8	24	72		22.86%
4	4	16	64		11.43%
5	3	15	75		8.57%
Σ	35	87	267		100.00%



n	35
Mod	2
Med	2
sr vred	2.485714

int var	4
varijansa	1.449796
st dev	1.204075
koef var	48%

Mod je vrednost obeležja sa najvećom frekvencijom. U ovom primeru to je vrednost 2, jer je njena frekvencija (12) veća od svih ostalih.

U tabeli sa kumulativnim frekvencijama, medijana se čita iz one vrste, u kojoj je kumulativna frekvencija F_i veća od polovine obima uzorka (u ovom primeru polovina od 35, dakle, 17,5).

Najzad, sredina obeležja se računa na sledeći način:

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{87}{35} = 2,485714 .$$

Procentualno odstupanje od sredine izražava se koeficijentom varijacije, V , koji se izračunava na sledeći način. Najpre se izračuna varijansa:

$$s^2 = \frac{\sum x_i^2 f_i}{n} - \bar{x}^2 = \frac{267}{35} - 2,485714^2 = 1,449796$$

zatim standardna devijacija

$$s = \sqrt{s^2} = \sqrt{1,449796} = 1,204075$$

i, najzad, koeficijent varijacije

$$V = \frac{s}{\bar{x}} \cdot 100\% = \frac{1,204075}{2,485714} \cdot 100\% = 48,44\%$$

2. Radi ispitivanja broja izostanaka sa nastave u toku jedne nedelje u jednoj školi sa 700 učenika, izabran je uzorak od 100 učenika i dobijeni su podaci dati u vidu sledeće tabele :

<i>Broj izostanaka</i>	0	1 - 4	5 - 9	10 - 14	15 - 30
<i>Broj učenika</i>	30	20	7	15	28

Izračunati mere centralne tendencije (srednju vrednost, medijanu, mod) i mere odstupanja (standardnu devijaciju i koeficijent varijacije).

REŠENJE:

Interval	x_i	f_i	$x_i f_i$	$x_i^2 f_i$	F_i
[0 , 1)	0.5	30	15	7.5	30
[1 , 5)	3	20	60	180	50
[5 , 10)	7.5	7	52.5	393.75	
[10 , 15)	12.5	15	187.5	2343.75	
[15 , 30]	22.5	28	630	14175	
Σ		100	945	17100	

n	100
sr vred	9.45
mod	0.5
med	5.25

var	81.698
st dev	9.0387
koef var	95.65%

Prateći dogovor u udžbeniku, računanje sa neprekidnim podacima, odnosno sa podacima grupisanim u intervale, zamenjujemo računanjima u kojima se za vrednosti x_i uzimaju sredine intervala. Stoga je mod sredina intervala sa najvećom frekvencijom, dakle sredina intervala $[0,1)$, odnosno 0.5.

Kod određivanja medijane može se desiti, baš kao u ovom primeru, da se tačno na sredini (u uređenom nizu podataka), nalaze dva različita podatka. U ovom primeru, u kome posmatramo 100 podataka, tačno na sredini su pedeseti i pedesetprvi podatak. Iz kumulativnih frekvencija čitamo da je pedeseti podatak 3, a pedesetprvi 7.5. Dakle, medijana je njihova aritmetička sredina

$$\text{medijana} = \frac{3 + 7,5}{2} = 5,25.$$

3. U okviru istraživanja životnog ciklusa jedne vrste sisara ispitivao se broj mladunaca iz jednog legla. Prikupljeni podaci o 800 slučajno izabranih legala svrstani su u tabelu:

Broj mladunaca	1	2	3	4	5
Broj legala	145	240	194	126	95

Odrediti mod, medijanu, aritmetičku sredinu obeležja i procentualno odstupanje od nje.

REŠENJE:

x_i	f_i	$x_i f_i$	$x_i^2 f_i$	F_i
1	145	145	145	145
2	240	480	960	385
3	194	582	1746	579
4	126	504	2016	
5	95	475	2375	
Σ	800	2186	7242	

n	800
Mod	2
Med	3
sr vred	2.7325

int var	4
var	1.585944
st dev	1.259343
koef var	46.09%

Mod je vrednost obeležja sa najvećom frekvencijom. U ovom primeru to je vrednost 2, jer je njena frekvencija (240) veća od svih ostalih.

U tabeli sa kumulativnim frekvencijama, medijana se čita iz one vrste, u kojoj je kumulativna frekvencija F_i veća od polovine obima uzorka (u ovom primeru polovina od 800, dakle, 400).

Najzad, sredina obeležja se računa na sledeći način:

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{2186}{800} = 2,7325 .$$

Procentualno odstupanje od sredine izražava se koeficijentom varijacije, V , koji se izračunava na sledeći način. Najpre se izračuna varijansa:

$$s^2 = \frac{\sum x_i^2 f_i}{n} - \bar{x}^2 = \frac{7242}{800} - 2,7325^2 = 1,5859$$

zatim standardna devijacija

$$s = \sqrt{s^2} = \sqrt{1,5859} = 1,2593$$

i, najzad, koeficijent varijacije

$$V = \frac{s}{\bar{x}} \cdot 100\% = \frac{1,2593}{2,7325} \cdot 100\% = 46,09\% .$$