

Mere oblika rasporeda

Deskriptivna statistika

Aleksandar Tomašević

April 2020.

1 Urađen primer

Zadatak 1. Varijablom X je označen broj pacijenata u stotinama ($1 = 100$ pacijenata) kojima je u toku jednog dana laboratorijski potvrđen COVID-19 na nivou jedne države. Podaci su beleženi tokom 42 dana.

U tabeli su prikazane vrednosti varijable X kao i njene frekvencije, odnosno broj dana u kojima je zabeleženo X potvrđenih slučajeva.

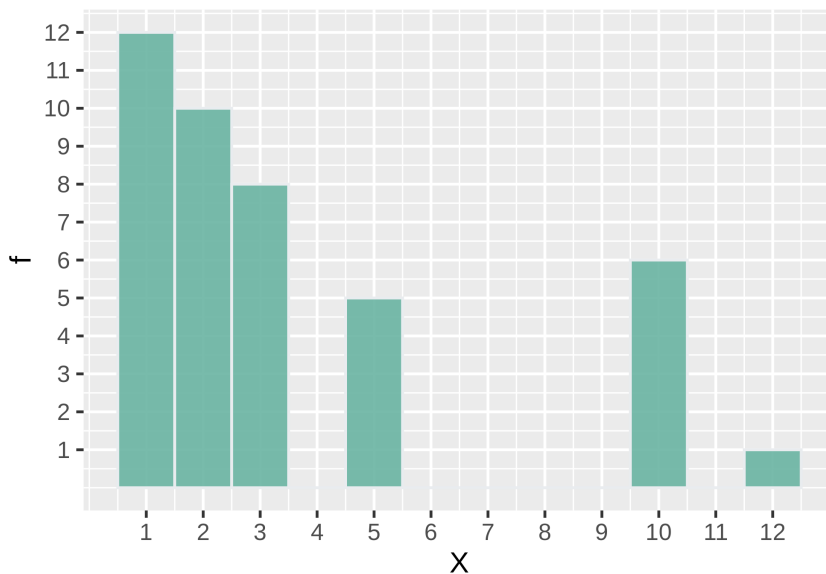
Tabela 1: Varijabla X i njene frekvencije

X	f
1	12
2	10
3	8
5	5
10	6
12	1
Σ	42

- A) Izračunati treći centralni moment varijable X
- B) Izračunati i prokomentarisati koeficijent asimetrije α_3
- C) Izračunati četvrti centralni moment varijable X
- D) Izračunati i prokomentarisati koeficijent spljoštenosti α_4

1.1 Formule i objašnjenja

Pre nego što prikazemo formulu za koeficijente asimetrije i spljoštenosti prikazaćemo rezultate grafički.

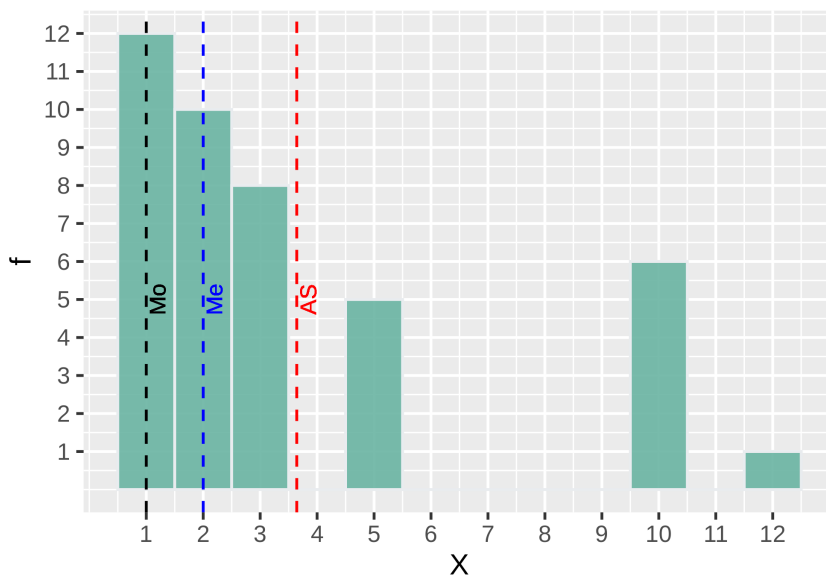


Slika 1: Histogram varijable X

Ono što možemo videti na slici 1 jeste da se najveći broj stubaca histograma nalazi sa leve strance grafikona, odnosno na vrednostima 1, 2 i 3.

Sa histograma se jasno vidi da je serija unimodalna i da je modus $M_o = 1$.

Ako na histogramu obeležimo vrednosti modusa, medijane i aritmetičke sredine, njihov međusobni odnos ukazaće nam na simetričnost ili asimetričnost ovih podataka.



Slika 2: Histogram sa merama centralne tendencije

U situaciju kada je $M_o < M_e < \bar{X}$ imamo pozitivno asimetričan raspored, odnosno asimetriju u desno. Detaljnije objašnjenje možete naći u udžbeniku na stranici 123.

Asimetrija u desno podrazumeva da se netipične vrednosti nalaze sa desne strane mera centralne tendencije. Zbog toga je aritmetička sredina pomerenjena u desno u odnosu na

modus i medijanu. U ovoj situaciji, možemo očekivati pozitivnu vrednosti koeficijenta asimetrije α_3 (vidi grafikon 4.1 u udžbeniku).

Formula 1. *Formula koeficijenta asimetrije ima sledeći oblik.*

$$\alpha_3 = \frac{M_3}{\sigma^3} \quad (1)$$

Pri čemu su:

M_3 — treći centralni moment aritmetičke sredine
 σ^4 — standardna devijacija na četvrti

Da bismo mogli da izračunamo vrednost ovog koeficijenta potreban nam je treći centralni moment aritmetičke sredine.

Centralne momente smo već koristili samo ih nismo tako imenovali.

Prvi centralni moment je suma razlika između vrednosti varijable X i njene aritmetičke sredine: $X - \bar{X}$. On je uvek jednak 0.

Drugi centralni moment je suma kvadriranih razlika između vrednosti varijable X i njene aritmetičke sredine. Koristi se prilikom izračunavanja varijanse.

Formula 2. *Treći centralni moment aritmetičke sredine.*

$$M_3 = \frac{\sum f(X - \bar{X})^3}{\sum f} \quad (2)$$

Dakle, da bismo izračunali vrednost ovog koeficijenta potrebno nam je da izračunamo:

- standardnu devijaciju
- M_3

Kada izračunamo α_3 , interpretacija je vrlo jednostavna. Ukoliko je $\alpha_3 > 0$ imamo pozitivnu asimetriju, a ukoliko je obrnuto imamo negativnu asimetriju. Ukoliko je $\alpha_3 = 0$ podaci su simetrični.

Dalje, imamo sledeće okvirne smernice za tumačenje intenziteta asimetrije.

- $|\alpha_3| < 0.25$ — slabo izražena asimetrija
- $0.25 \leq |\alpha_3| < 0.5$ — srednje izražena asimetrija
- $0.5 \leq |\alpha_3|$ — jako izražena asimetrija

Kada je u pitanju koeficijent spljoštenosti, njegova forma je simetrična prethodnm koeficijentu, s tim što se koristi četvrti centralni moment.

Formula 3. *Koeficijent spljoštenosti.*

$$\alpha_4 = \frac{M_4}{\sigma^4} \quad (3)$$

Formula 4. *Čevrti centralni moment aritmetičke sredine.*

$$M_4 = \frac{\sum f(X - \bar{X})^4}{\sum f} \quad (4)$$

Interpretacija koeficijenta spljoštenosti je rukovodi prema sledećim parametrima.

- $\alpha_4 < 3$ — Spljoštenost veća od normalne (spljošten raspored)
- $\alpha_4 = 3$ — Spljoštenost normalnog rasporeda
- $\alpha_4 > 3$ — Spljoštenost manja od normalne (izdužen raspored)

Videti stranu 133. u udžbeniku za detaljna pojašnjenja koncepta spljoštenosti.

1.2 A

Da bismo krenuli sa računanjem trećeg centralnog momenta potrebno je izračunati aritmetičku sredinu i varijansu, odnosno standardnu devijaciju.

Potrebne su nam sledeće nove kolone u tabeli: fX , $(X - \bar{X})$, $f(X - \bar{X})$ i $f(X - \bar{X})^2$.

Tabela 2: Dodate nove kolone

f	fX	$X - \bar{X}$	$f(X - \bar{X})$	$f(X - \bar{X})^2$
12	12	-2.64	-31.71	83.82
10	20	-1.64	-16.43	26.99
8	24	-0.64	-5.14	3.31
5	25	1.36	6.79	9.21
6	60	6.36	38.14	242.48
1	12	8.36	8.36	69.84
42	153			435.64

Aritmetička sredina iznosi $\bar{X} = \frac{153}{42} = 3.64$. Što znači da je prosečan broj pozivinih pacijenata u toku jednog dana bio 364.

Varijansa iznosi $\sigma^2 = \frac{435.64}{42} = 10.37$, dok je standardna devijacija $\sigma = \sqrt{\sigma^2} = 3.22$.

Da bismo izračunali treći centralni momenat potrebna nam je kolona $f(X - \bar{X})^3$. Nju možemo brzo izračunati tako što ćemo pomnožiti vrednosti kolone $f(X - \bar{X})^2$ i $X - \bar{X}$.

Tabela 3: Dodata nova kolona $f(X - \bar{X})^3$

X	f	fX	$X - \bar{X}$	$f(X - \bar{X})$	$f(X - \bar{X})^2$	$f(X - \bar{X})^3$
1	12	12	-2.64	-31.71	83.82	-221.51
2	10	20	-1.64	-16.43	26.99	-44.34
3	8	24	-0.64	-5.14	3.31	-2.13
5	5	25	1.36	6.79	9.21	12.50
10	6	60	6.36	38.14	242.48	1541.48
12	1	12	8.36	8.36	69.84	583.68
Σ	42	153			435.64	1869.67

Na osnovu sume dodate kolone možemo izračunati treći centralni moment.

$$M_3 = \frac{1869.67}{3.22^3} = 44.52 \quad (5)$$

Napomena. Izračunavanje trećeg centralnog momenta iziskuje dosta vremena. Dešava se na testovima da studenti kada izračunaju ovu vrednost, napišu da je to vrednost koeficijenta asimetrije. Drugim rečima, pomešaju dve formule. Vodite računa da uvek posle izračunavanja trećeg centralnog momenta nastavite dalje sa izračunavanjem koeficijenta asimetrije.

1.3 B

A potom možemo izračunati i koeficijent asimetrije.

$$\alpha_3 = \frac{M_3}{\sigma^3} = \frac{44.52}{3.22^3} = 1.33 \quad (6)$$

Kao što smo mogli da pretpostavimo gledajući histogram, reč je o pozitivnoj asimetriji (asimetriji u desno) i ona je snažna ili snažno izražena.

Napomena. Takođe, često se dešava da studenti uzmu sumu kolone i direktno je primene u formuli za koeficijent, bez izračunavanja trećeg centralnog momenta. Vodite računa i pratite da li ste u svakoj formuli zapisali pravu vrednost.

1.4 C

Da bismo izračunali četvrti centralni moment potrebno je dodati još jednu kolonu u tabelu: $f(X - \bar{X})^4$. Ovu kolonu ćemo izračunati tako što ćemo pomnožiti vrednosti kolona $f(X - \bar{X})^3$ i $X - \bar{X}$.

Tabela 4: Dodata nova kolona

X	f	fX	$X - \bar{X}$	$f(X - \bar{X})$	$f(X - \bar{X})^2$	$f(X - \bar{X})^3$	$f(X - \bar{X})^4$
1	12	12	-2.64	-31.71	83.82	-221.51	585.43
2	10	20	-1.64	-16.43	26.99	-44.34	72.84
3	8	24	-0.64	-5.14	3.31	-2.13	1.37
5	5	25	1.36	6.79	9.21	12.50	16.96
10	6	60	6.36	38.14	242.48	1541.48	9799.39
12	1	12	8.36	8.36	69.84	583.68	4877.88
Σ	42	153			435.64	1869.67	15353.88

Sumu ove kolone iskoristićemo da izračunamo četvrti centralni moment.

$$M_4 = \frac{15354}{42} = 365.67$$

1.5 D

Na kraju, koeficijent spljoštenosti možemo izračunati na sledeći način.

$$\alpha_4 = \frac{365.67}{3.22^4} = 3.39$$

Spljoštenost je u slučaju ovih podataka manja od normalne, odnosno reč je o izduženoj seriji podataka. To se može videti i na histogramu, po razlici u visini najmanjih i najvećih stubaca.