

MJERE ASIMETRIJE I MJERA ZAobljenosti

Mjere asimetrije

Mjerama asimetrije mjeri se način rasporeda podataka prema aritmetičkoj sredini ili nekoj drugoj vrijednosti. Najvažnije mjere asimetrije jesu: *koeficijent asimetrije* α_3 , *Pearsonova* (Pearsonova) i *Bowleyjeva* (Baulijeva) mjera. Koeficijent asimetrije potpuna je mjera asimetrije jer se temelji na uporabi svih vrijednosti niza, što nije slučaj s Pearsonovom i Bowleyjevom mjerom.

Koeficijent asimetrije α_3 najvažnija je mjera asimetrije za numeričke nizove.

KOEFICIJENT ASIMETRIJE α_3

Koeficijent asimetrije numeričkog niza jest omjer trećeg momenta oko sredine μ_3 i standardne devijacije σ podignute na treću potenciju.

$$(1) \alpha_3 = \frac{\mu_3}{\sigma^3}, \mu_3 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3, \mu_3 = \frac{1}{N} \sum_{i=1}^k f_i (x_i - \bar{x})^3$$

U navedenom izrazu¹ μ_3 je treći moment oko sredine. Treći moment oko sredine jest aritmetička sredina odstupanja vrijednosti numeričke varijable od njezine aritmetičke sredine podignutih na treću potenciju.

Ako je numerički niz dan kao distribucija frekvencija u navedenom izrazu za μ_3 , odstupanja varijable od prosjeka ponderiraju se frekvencijama.

Kada je raspored podataka simetričan, tada postoji jednako odstupanja $(x - \bar{x})^3$ s pozitivnim i negativnim predznakom. Budući da neparna potencija ne dokida predznak, u zbroju i prosjeku ta će odstupanja biti jednaka 0. I koeficijent asimetrije za simetričan raspored jednak je 0.

U pozitivno asimetričnom rasporedu (distribuciji) vrijednosti navedenog odstupanja s pozitivnim predznakom prevladavaju one s negativnim pa je prosjek tih odstupanja (treći moment oko sredine) i koeficijent asimetrije pozitivan (veći od nule). U negativno asimetričnom rasporedu podataka navedena odstupanja s negativnim predznakom prevladavaju u zbroju i prosjeku pa je koeficijent asimetrije negativan (manji od nule).

Treći moment oko sredine dijeli se sa standardnom devijacijom na treću potenciju kako bi se dobio pokazatelj neovisan o mjernim jedinicama varijable. Koeficijent asimetrije uobičajeno poprima vrijednost ± 2 , a može poprimiti i veću vrijednost.

PRIMJER 1.

Za distribuciju turista prema izvanpansionskoj potrošnji izračunana je prosječna potrošnja (aritmetička sredina) i prosječno odstupanje od prosječne potrošnje (standardna devijacija) te su te vrijednosti: $\bar{x} = 361.38$, $\sigma = 117.87831$. Određuje se koeficijent asimetrije α_3 .

¹ Oznaka μ_3 čita se mi tri.

Koeficijent asimetrije za distribuciju frekvencija jest:

$$\mu_3 = 2977918.848, \alpha_3 = \frac{2977918.848}{117.87831^3}, \alpha_3 = 1.82.$$

Treći moment oko sredine i koeficijent asimetrije jesu:

$$\mu_3 = 2977918.848, \alpha_3 = \frac{2977918.848}{117.87831^3}, \alpha_3 = 1.82.$$

Distribucija turista prema potrošnji jest pozitivno asimetrična, to jest zakošena prema većim vrijednostima. Drugim riječima, koncentracija turista veća je na manjoj prosječnoj potrošnji. U takvoj je distribuciji najveća aritmetička sredina, slijedi medijan, a najmanju vrijednost poprima mod.

Na način rasporeda podataka upućuju i odnosi srednjih vrijednosti. Prosudba simetričnosti odnosno asimetričnosti distribucije *frekvencija kontinuirane numeričke varijable* provodi se na temelju sljedećih odnosa među srednjim vrijednostima. U simetričnim distribucijama sve su tri srednje vrijednosti međusobno jednake. U pozitivno asimetričnim distribucijama najveća je aritmetička sredina, nju slijedi medijan te mod. U negativno asimetričnim distribucijama najmanja je aritmetička sredina, a nju slijede prema većim vrijednostima medijan i mod.

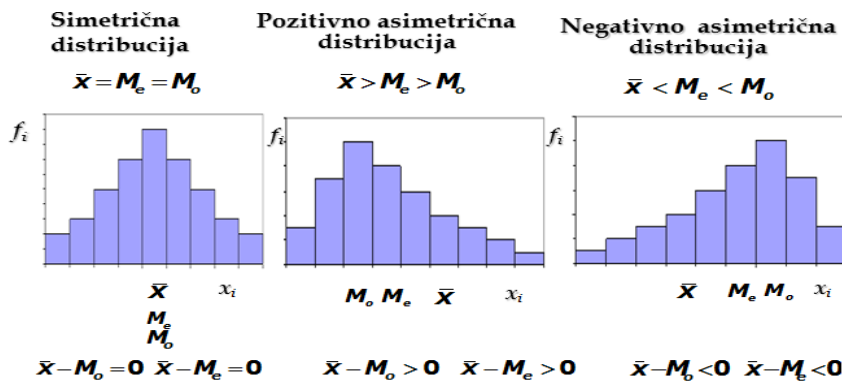
Pearsonova mjera S_k

Pearsonova mjera temelji se na odnosima aritmetičke sredine medijana, odnosno moda distribucije kontinuirane varijable. Mjera je definirana sljedećim izrazom.

PEARSONOVA MJERA dana je izrazom:

$$(2 \text{ a}) S_k = \frac{3(\bar{x} - M_e)}{\sigma}, (2 \text{ b}) S_k = \frac{(\bar{x} - M_o)}{\sigma}.$$

Pearsonova mjera jest standardizirano odstupanje vrijednosti medijana ili moda od aritmetičke sredine. U simetričnim distribucijama kontinuirane varijable sve su tri vrijednosti jednake pa je razlika moda ili medijana i aritmetičke sredine jednaka nuli. Stoga je i ta mjera jednaka 0. U pozitivno asimetričnim distribucijama razlika sredine i medijana je pozitivna, a u negativno asimetričnim razlika je negativna. Razlika medijana i sredine (moda i sredine) dijeli se standardnom devijacijom radi uklanjanja utjecaja mjernih jedinica varijable na vrijednost pokazatelja.



Slika 1. Odnosi srednjih vrijednosti u distribucijama kontinuirane numeričke varijable

Pearsonova mjera uobičajeno poprma vrijednosti iz intervala ± 3 , ali može biti i izvan tog intervala. Vrijednosti dobivene izrazima (2 a) i (2 b) po pravilu se u manjoj mjeri razlikuju. Predznak mjere uvijek je jednak.

PRIMJER 2.

Za distribuciju u primjeru III-10.b koja se odnosi na zaposlene u djelatnosti *trgovina na veliko i malo; popravak motornih vozila i motocikala* prema starosti u pravnim osobama, stanje 31. ožujka 2010. određeni su aritmetička sredina, medijan i mod te standardna devijacija. Računa se Pearsonova mjera asimetrije.

Vrijednosti za računanje su: $\bar{x} = 38.2$, $M_e = 36.8$, $M_o = 32.0$, $\sigma = 10.5$.

Pearsonova mjera (2 a) jest:

$$S_k = \frac{3(\bar{x} - M_e)}{\sigma} = \frac{3(38.2 - 36.8)}{10.5}, S_k = +0.40.$$

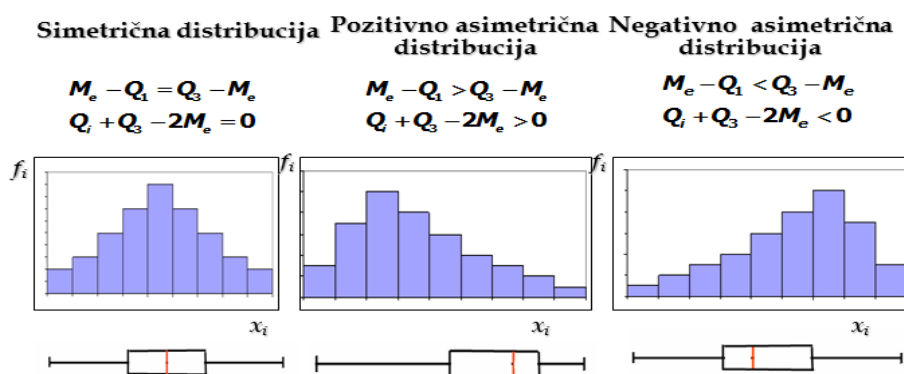
Na temelju odnosa sredine i moda, vrijednost koeficijenta jest:

$$S_k = \frac{3(\bar{x} - M_o)}{\sigma} = \frac{3(38.2 - 32.0)}{10.5}, S_k = +0.59.$$

Distribucija zaposlenih prema starosti umjereno je pozitivno asimetrična. Pearsonova mjera asimetrije je +0.40, odnosno +0.59. Razlika u vrijednosti mjere posljedica je činjenice što navedeni odnosi srednjih vrijednosti nisu egzaktni, nego aproksimativni.

Bowleyjeva mjera S_{kQ}

Bowleyjeva mjera asimetrije temelji se na odnosima kvartila i medijana. U simetričnom rasporedu vrijednosti razlika gornjega kvartila i medijana jednaka je razlici medijana i donjega kvartila, to jest $Q_1 + Q_3 - 2M_e = 0$. U pozitivno asimetričnom rasporedu razlika gornjega kvartila i medijana veća je od razlike medijana i donjega kvartila, a u negativno asimetričnom razlika gornjega kvartila i medijana manja je od razlike medijana i donjega kvartila. Vrijednost $Q_1 + Q_3 - 2M_e$ dijeli se interkvartilom radi normiranja pokazatelja.



Slika 2. Odnosi medijana i kvartila

BOWLEYJEVA MJERA S_{kQ} dana je izrazom:

$$(3) S_{kQ} = \frac{Q_1 + Q_3 - 2M_e}{Q_3 - Q_1}, \quad -1 \leq S_{kQ} \leq 1.$$

Bowleyjeva mjera mjera poprima vrijednosti iz intervala ± 1 .

PRIMJER 3.

Za distribuciju rezultata na završnom ispitu pristupnika iste vrste srednje škole i iste razine ispita iz predmeta matematika i engleski jezik izračunani su donji i gornji kvartil te medijan. Varijabla je *postotak riješenih zadataka*. Određuje se mjera asimetrije. Vrijednosti navedenih pokazatelja dani su tablici.

	Matematika	Engleski
donji kvartil	58.82	54.27
medijan	62.74	75.83
gornji kvartil	69.90	82.76
Bowleyjeva mjera	0.29242	-0.51351

Bowleyjeva mjera distribucije rezultata iz matematike jest:

$$S_{kQ} = \frac{Q_1 + Q_3 - 2M_e}{Q_3 - Q_1} = \frac{58.82 + 69.90 - 2 \cdot 62.74}{69.90 - 58.82} = 0.29242, S_{kQ} = 0.29242.$$

Bowleyjeva mjera distribucije rezultata iz engleskog jezika jest:

$$S_{kQ} = \frac{Q_1 + Q_3 - 2M_e}{Q_3 - Q_1} = \frac{54.27 + 82.76 - 2 \cdot 75.83}{82.76 - 54.27} = -0.51351, S_{kQ} = -0.51351.$$

Rezultati pokazuju da je medijalni postotak riješenih zadataka iz matematike 62.74 %, a iz engleskog jezika je veći i iznosi 75.83 %. Drugim riječima, uspjeh na ispitu iz engleskog jezika predložen medijanom bolji je od uspjeha na ispitu iz matematike. Asimetrija rezultata iz matematike mjerena Bowleyjevom mjerom umjereno je pozitivna i upućuje na koncentraciju rezultata na manjim vrijednostima (slabijim ispitnim rezultatima). Asimetrija rezultata iz engleskog jezika mjerena Bowleyjevom mjerom umjereno je negativna i upućuje na veću koncentraciju rezultata na većim vrijednostima varijable.

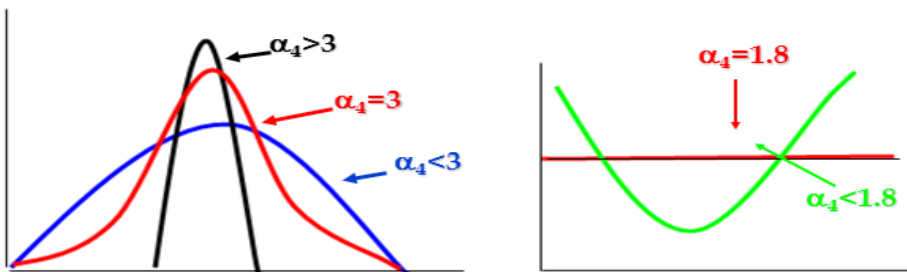
Mjera zaobljenosti (kurtosis)

Mjerom zaobljenosti mjeri se zaobljenost modalnog vrha distribucije. Osnova zaključivanja o zaobljenosti jest zaobljenost normalne distribucije. Zaobljenost se mjeri koeficijentom zaobljenosti.

KOEFICIJENT ZAobljenosti dan je izrazom:

$$(4) \alpha_4 = \frac{\mu_4}{\sigma^4}, \quad \kappa = \alpha_4 - 3.$$

Koeficijent zaobljenosti α_4 jest omjer četvrtog momenta oko sredine i standardne devijacije na četvrtu potenciju. Četvrti moment oko sredine jest prosječno odstupanje vrijednosti varijable od aritmetičke sredine podignuto na četvrtu potenciju. Koeficijent zaobljenosti normalne distribucije je 3. Ako je izračunana vrijednost koeficijenta jednaka 3, kaže se da je jednako zaobljena kao i normalna distribucija. Kada je koeficijent zaobljenosti veći od 3, govori se o „šiljatijoj” distribuciji u usporedbi s normalnom distribucijom. Ako je koeficijent manji od 3, distribucija je „plosnatija” u usporedbi s normalnom distribucijom. Koeficijent zaobljenosti pravokutne distribucije jednak je 1.8, a U-distribucije manji od 1.8.



Slika 3. Zaobljenost distribucija

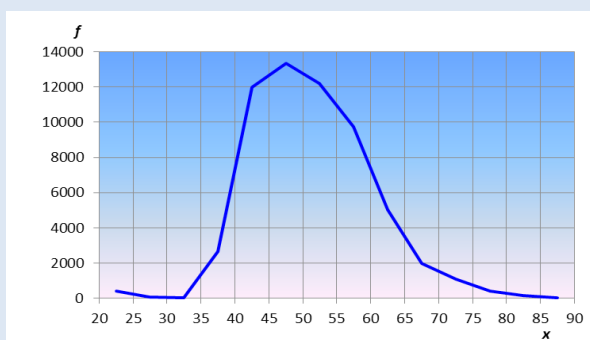
Kadšto se mjera definira izrazom $\kappa = \alpha_4 - 3$ i naziva se *ekscesom*. Kada je κ (kapa) jednak 0, riječ o zaobljenosti jednakoj zaobljenosti normalne distribucije. Pozitivna vrijednost mjere upućuje na „šiljatiju”, a negativna vrijednost na „plosnatiju” zaobljenost distribucije prema zaobljenosti normalne distribucije.

PRIMJER 4.

Distribucija frekvencija numeričke varijable s razredima prikazana je poligonom frekvencija. Za distribuciju su izračunane ove veličine:

$$\bar{x} = 51.3, Me = 50.4, Q_1 = 44.8, Q_3 = 56.9, \sigma = 8.8, \mu_3 = 273.3, \mu_4 = 23384.1.$$

Određuje se koeficijent zaobljenosti te pokazatelji asimetrije (koeficijent asimetrije, Pearsonova i Bowleyjeva mjera).



Zaobljenost distribucije (kurtosis) mjeri se mjerom zaobljenosti α_4 . Mjera je:

$$\alpha_4 = \frac{\mu_4}{\sigma^4} = \frac{23384.1}{8.8^4} = 3.9, \alpha_4 = 3.9, \kappa = \alpha_4 - 3 = 3.9 - 3, \kappa = 0.9.$$

Distribucija je „šiljatija“ u usporedbi sa zaobljenošću normalne distribucije jer je $\alpha_4 > 3$, odnosno $\kappa > 0$.

$$\text{Koeficijent asimetrije je } \alpha_3 \text{ i poprima vrijednost: } \alpha_3 = \frac{\mu_3}{\sigma^3} = \frac{273.3}{8.8^3} = 0.4, \alpha_3 = +0.4.$$

$$\text{Pearsonova mjera asimetrije jest: } S_k = \frac{3(\bar{x} - M_e)}{\sigma} = \frac{(51.3 - 50.4)}{8.8}, S_k = +0.3.$$

$$\text{Bowleyjeva je mjera: } S_{kQ} = \frac{Q_1 + Q_3 - 2M_e}{Q_3 - Q_1} = \frac{44.8 + 56.9 - 2 \cdot 50.4}{56.9 - 44.8} = +0.07.$$

Pokazatelji asimetrije upućuju na zaključak da je distribucija blago pozitivno asimetrična.