

Zadaci za vježbu: oblik razdiobe i standardizacija

Vježbe 3

1 Zadaci za rješavanje olovkom i papirom

Zadatak 1. Promjena mjerne jedinice i pomak nule.

Četiri izmjerene temperature su 18, 20, 20, 22 °C. Temperatura u kelvinima dobiva se dodavanjem 273.15.

- (a) Izračunajte sredinu i empirijsku varijancu u stupnjevima Celzija.
- (b) Bez ponovnog zbrajanja odstupanja odredite sredinu i varijancu u kelvinima.
- (c) Obrazložite mijenjaju li se standardizirane vrijednosti promjenom jedinice prema formuli $y_i = 1.8x_i + 32$ (stupnjevi Fahrenheita).

Zadatak 2. Usporedba položaja unutar skupine.

Duljina lista A iznosi 54 mm; njegova skupina ima sredinu 50 mm i empirijsku standardnu devijaciju 2 mm. Duljina lista B iznosi 72 mm; njegova skupina ima sredinu 60 mm i standardnu devijaciju 8 mm.

- (a) Odredite obje z -vrijednosti i protumačite ih.
- (b) Koji je list dulji, a koji je relativno više iznad sredine svoje skupine?
- (c) Koja bi duljina u skupini B imala istu z -vrijednost kao list A? Možete li samo iz z -vrijednosti odrediti postotak kraćih listova?

Zadatak 3. Simetrija ne određuje broj modova.

Brojevi izdanaka u dvama uzorcima su

$$A : 0, 0, 2, 2, 2, 4, 4, \quad B : 0, 0, 0, 2, 4, 4, 4.$$

- (a) Nacrtajte stupčaste dijagrame frekvencija. Odredite modove, medijane i sredine.
- (b) Bez računanja svih standardiziranih vrijednosti pokažite da je zakaošenost obaju uzoraka nula.
- (c) Objasnite zašto nulta zakaošenost sama po sebi ne jamči jedan mod.

Zadatak 4. Zrcaljenje podataka.

Udio oštećene površine četiriju listova iznosi 0, 0, 0, 4 posto. Udio neoštećene površine dobiva se kao $y_i = 100 - x_i$.

- (a) Izračunajte sredinu, varijancu i zakošenost udjela oštećene površine.
- (b) Odredite sredinu i varijancu neoštećenog udjela pomoću transformacije.
- (c) Pokažite kako se mijenjaju z -vrijednosti, zakošenost i spljoštenost pri zrcaljenju. Usporedite položaje repova dviju razdioba.

Zadatak 5. Ista varijanca, različiti repovi.

Odstupanja očitavanja senzora od referentne vrijednosti (mg/L) su

$$A : -1, -1, -1, -1, 1, 1, 1, 1, \quad B : -2, 0, 0, 0, 0, 0, 0, 2.$$

- (a) Provjerite da uzorci imaju jednaku sredinu i varijancu.
- (b) Izračunajte zakošenost i spljoštenost obaju uzoraka.
- (c) Za svaki uzorak odredite udio opažanja strogo udaljenih više od 1.5s od sredine. Povežite rezultat s četvrtim potencijama u spljoštenosti.

2 Zadaci za rad u R-u

Zadatak 1. Provjera standardizacije.

Upotrijebite stvarne ugrađene podatke `iris$Sepal.Width` (cm).

- (a) Izračunajte empirijsku sredinu i standardnu devijaciju te sami standardizirajte podatke. Provjerite sredinu 0 i empirijsku varijancu 1.
- (b) Usporedite rezultat s `as.numeric(scale(x))`. Objasnite razliku zbog nazivnika n odnosno $n - 1$.
- (c) Ispišite pet najmanjih i pet najvećih z -vrijednosti te ih povežite s izvornim mjerenjima. Jesu li se rangovi opažanja promijenili?

Zadatak 2. Svojstva linearnih transformacija.

Vremena obrade (s) su 5, 6, 6, 7, 8, 8, 9, 10, 12, 19, 23, 31.

- (a) Napišite kratku funkciju koja vraća sredinu, empirijsku varijancu, zakošenost i spljoštenost.
- (b) Primijenite je na $\mathbf{x}$, $\mathbf{x}/60$, $\mathbf{x} + 10$ i $-\mathbf{x}$. Posljednji vektor služi samo kao matematička transformacija.
- (c) Prikažite izvorni i zrcaljeni uzorak histogramima s preslikanim granicama razreda. Povežite grafove s promjenom predznaka zakošenosti.

Zadatak 3. Dva vrha nastala spajanjem skupina.

Duljine sadnica (cm) iz dviju razvojnih skupina su:

A	<-	c(8, 9, 9, 9, 10, 10, 10, 10, 11, 11, 11, 12)
B	<-	c(18, 19, 19, 19, 20, 20, 20, 20, 21, 21, 21, 22)

- (a) Nacrtajte histograme obje skupine i svih podataka zajedno, s granicama `seq(7.5, 22.5, by = 1)` i jednakom skalom osi frekvencija.
- (b) Izračunajte sredinu, medijan i zakošenost spojenog uzorka.
- (c) Je li ukupna sredina blizu mnogim opažanjima? Objasnite zašto simetrija spojenog histograma ne isključuje dvije različite skupine.

Zadatak 4. Utjecaj udaljenosti jednoga opažanja.

Početne koncentracije (mg/L) su 4, 5, 5, 6, 6, 6, 7, 7, 8, 8, 9.

- (a) Dodajte im po jedno opažanje u , za $u = 10, 15, 20, 30, 50$. Za svaki novi uzorak izračunajte sredinu, medijan, zakošenost i spljoštenost.
- (b) Prikažite promjenu svake mjere u ovisnosti o u na zasebnom grafu.
- (c) Objasnite zašto treća i četvrta potencija naglašavaju udaljena opažanja, ali se pritom u tim koeficijentima mijenjaju i sredina i standardna devijacija.

Zadatak 5. Što čini uparivanje opažanja?

Promjene mase (mg) na šest uzoraka u prvom koraku pokusa su $x = (-3, -2, -1, 1, 2, 3)$. Za drugi korak razmotrite $y1 \leftarrow x$, $y2 \leftarrow \text{rev}(x)$ i $y3 \leftarrow c(-1, 3, -2, 2, -3, 1)$.

- (a) Provjerite da tri vektora drugog koraka imaju jednake sredine i varijance.
- (b) Izračunajte sredine i varijance ukupnih promjena $x + y1$, $x + y2$ i $x + y3$ te ih prikazite stupčastim dijagramima.
- (c) Objasnite zašto se sredine mogu zbrojiti, a varijance zbroja nisu određene samo varijancama dvaju koraka. Ne koristite još koeficijent korelacije.