

FAKULTET TEHNIČKIH NAUKA
GEODEZIJA I GEOINFORMATIKA

RAČUN IZRAVNANJA
VEŽBA 1

NOVI SAD, 2024.

Statistički skup

Statistički skup predstavlja skup jedinica na osnovu kojih se ispituje jedno ili više svojstava (varijabli, obeležja, osobina, karakteristika), koja su od jedinice do jedinice promenljiva. Prema obimu statistički skupovi se dele na konačne i beskonačne.

Osnovni skup (populacija) je skup podataka svih jedinica (elemenata) posmatranja čije karakteristike ispituujemo.

Uzorak je deo osnovnog skupa koji je izabran u svrhu izvođenja statističke analize.

Varijacioni niz čine elementi uzorka poredani od najmanjeg ka najvećem.

Za uzorak (X_1, X_2, \dots, X_n) varijacioni niz čini niz slučajnih promenljivih sačinjen od elemenata ovog uzorka u oznaci $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ za koji važi $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$

Distribucija frekvencija

Distribucija frekvencija predstavlja tabelarno prikazivanje podataka, gde podatke grupišemo u dve kolone tako da su u prvoj koloni navedene sve različite vrednosti obeležja, a u drugoj koloni broj jedinica navedene vrednosti obeležja. Različite vrednosti obeležja se nazivaju modaliteti obeležja.

Broj ponavljanja svake navedene vrednosti obeležja ili grupe (intervala) obeležja naziva se **apsolutna frekvencija** (f_i). Na ovaj način se dobija distribucija ili raspodela frekvencija.

Kumulativna apsolutna frekvencija (F_i) određene vrednosti obeležja dobija se sabiranjem apsolutnih frekvencija svih prethodnih obeležja i apsolutne frekvencije tog obeležja

$$F_i = \sum_{j=1}^i f_j, (i = 1, \dots, k)$$

Kompletno rešenje zadatka dostupno je u fajlu *Račun izravnjanja - Vežba 1.xlsx*.

Distribucija frekvencija

Relativna frekvencija (p_i) se dobija kao količnik apsolutne frekvencije svake vrednosti obeležja i ukupnog broja jedinica posmatranja.

$$p_i = \frac{f_i}{n}, (i = 1, \dots, k)$$

Kumulativna relativna frekvencija (F_i^r) određene vrednosti obeležja dobija se sabiranjem relativnih frekvencija svih prethodnih obeležja i relativne frekvencije tog obeležja

$$F_i^r = \sum_{j=1}^i p_j, (i = 1, \dots, k)$$

NUMERIČKA PREZENTACIJA REZULTATA MARENJA

Kategorije numeričkih ocenjivača

- Mere centralne tendencije i pozicione mere
- Mere disperzije

Mere centralne tendencije

Mere centralne tendencije su statističke veličine koje ukazuju na vrednost iz skupa rezultata merenja koja ima tendenciju približenja sredini skupa.

- **Aritmetička sredina** ili empirijsko matematičko očekivanje

(Predstavlja broj koji se izračunava kada se zbir svih vrednosti skupa podeli sa ukupnim brojem tih vrednosti u skupu podataka)

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$$

Excel: AVERAGE

Mere centralne tendencije (pozicione mere)

• Medijana

Excel: MEDIAN

Ukoliko skup sadrži neparan broj rezultata, medijanu predstavlja rezultat koji se nalazi u sredini rastućeg niza. Ako je broj rezultata paran, tada sredina dva susedna rezultata u sredini niza formiraju medijanu.

$$M_e = X_{\frac{n+1}{2}}$$

Primer: (42, 46, 48, 56, 63) ili

$$M_e = \frac{1}{2} \left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right)$$

Primer: 42, 46, 48, 52, 56, 63 => (48 + 52) / 2 = 50

• Moda

Vrednost koja se najčešće pojavljuje u datom skupu merenja neke fizičke veličine. Retko se u praksi koristi jer je često skup merenja mali. Nije jednoznačna jer se često u nekom skupu isti broj puta pojavi nekoliko rezultata.

11, 14, 14, 16, 18, 19 ili 11, 14, 14, 16, 18, 18, 19

Napomena: Primenljiva za vertikalno uređen skup rezultata merenja.

Za vertikalno
uređen niz:

Excel: MODE.MULT

Za horizontalno
uređen niz:

Excel: TRANSPOSE(MODE.MULT())

Kompletno rešenje zadatka dostupno je u fajlu *Račun izravnjanja - Vežba 1.xlsx*.

Mere disperzije (raspršenosti ili rasturanja podataka)

Mere disperzije ukazuju na raspršenost raspodele.

Disperzija populacije
$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{n} = \frac{\sum_{i=1}^n \varepsilon_i^2}{n},$$

gde je: $\varepsilon_i = X_i - \mu$ istinita greška; n broj merenja.

Disperzija uzorka
$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n V_i^2}{n-1},$$

gde je: $V_i = X_i - \bar{X}$ rezidual (razlika između pojedinog rezultata merenja i najverovatnije vrednosti merene veličine); $n - 1$ broj stepeni slobode.

Mere disperzije (raspršenosti ili rasturanja podataka)

Raspon merenja

$$\omega_n = |X_{\max} - X_{\min}|$$

Standardna greška

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{n}} = \sqrt{\frac{\sum_{i=1}^n \varepsilon_i^2}{n}}$$

Standardno odstupanje

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^n V_i^2}{n-1}}$$

Excel: STDEV

Standardno odstupanje je ocena standardne greške populacije. Ona predstavlja praktičan izraz ocene preciznosti uzorka merenja. Teorijski, oko 68.3% rezultata merenja nalazi se u intervalu = **najverovatnija vrednost $\pm S$** .

Standardno odstupanje sredine

$$S_{\bar{X}} = \frac{S}{\sqrt{n}}$$

Kada $n \rightarrow \infty$, tada $S \rightarrow 0$. To znači da što se više veličina uzorka približava populaciji sračunata vrednost sredine približava se sredini kao parametru populacije (μ).

Kompletno rešenje zadatka dostupno je u fajlu *Račun izravnjanja - Vežba 1.xlsx*.

Grupisanje podataka

Kada obeležje ima veliki broj različitih vrednosti one se grupišu u unapred određene intervale.

Broj i veličina (širina) intervala zavise od broja podataka (n) i od prirode samog obeležja.

Broj grupnih intervala (k) može se izračunati na osnovu izraza $k \leq 5 * \log n$

Na osnovu broja intervala izračunava se **širina intervala (i)**:

$$i = \frac{X_{\max} - X_{\min}}{k}$$

gde su X_{\max} i X_{\min} najveća i najmanja vrednost obeležja u seriji.

Distribucija frekvencija

Gustina raspodele

$$G(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Funkcija raspodele

$$F(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Excel: NORMDIST

Uputstvo za računanje u Excel-u:

- **X** – uzima se vrednost mediane za zadati interval
- **Mean** – srednja vrednost rezultata merenja
- **Standard_dev** – standardno odstupanje
- **Cumulative** – za funkciju raspodele upisuje se **True**, a za gustinu raspodele upisuje se **False**

Kompletno rešenje zadatka dostupno je u fajlu *Račun izravnjanja - Vežba 1.xlsx*.

Grafičko prikazivanje statističkih podataka

Za grafičko prikazivanje numeričkih statističkih serija, odnosno distribucija frekvencija najčešće se koristi **histogram frekvencija**.

Histogram frekvencija čine pravougaonici čija je osnovica jednaka veličini grupnog intervala, a visina odgovara frekvenciji grupnog intervala.

Mere disperzije (pri grupisanom rasporedu)

**Aritmetička sredina
pri grupisanom rasporedu**

$$\tilde{x} \approx \bar{x} = \sum (x_j * p_j)$$

**Standardno odsupanje
pri grupisanom rasporedu**

$$s_{\tilde{x}}^2 = \frac{1}{n-1} \left(\sum_{i=1}^n v_i^2 - n\tilde{x}^2 \right)$$