

Osnove statistike u demografiji

Predavanje 3

Podsjetnik

- Populacija, statistički skup i uzorak
- Vitalni procesi
- Mjerne skale
- Varijable (kvantitativne i kvalitativne)
- Statistički nizovi

Mjere deskriptivne statističke analize

- Statističke metode:
 - Metode deskriptivne statistike
 - Metode inferencijalne statistike

Metode deskriptivne statistike

1. uređivanje
 2. prikazivanje statističkih podataka
 3. izračunavanje deskriptivnih brojčanih pokazatelja.
- Najčešće se raspolaže s velikim brojem podataka i njihovo uređivanje se svodi na
 - grupiranje,
 - tablično i
 - grafičko prikazivanje.
 - Pritom se svi dobiveni zaključci o karakteristikama analiziranih podataka odnose isključivo na analizirane podatke bez mogućnosti **poopćavanja**

Metode inferencijalne statistike

- polaze od podskupa ili uzorka podataka, a zaključci se uz primjenu teorije vjerojatnosti mogu poopćiti na cijeli skup.
- Polazeći od uzorka podataka metodama inferencijalne statistike donose se zaključci o cjelini (populaciji) i to:
 - **procjenjivanjem nepoznatih karakteristika** cjeline (populacije) uz određenu razinu pouzdanosti ili
 - **testiranjem pretpostavke** o nepoznatim karakteristikama cjeline (populacije) uz odabranu razinu značajnosti.

Metode inferencijalne statistike

- Zaključci o nepoznatim karakteristikama populacije doneseni na temelju uzorka podataka nemaju apsolutnu sigurnost
- Stoga se takvi zaključci mogu poopćiti na cijeli skup samo uz određenu razinu pouzdanosti i određenu razinu značajnosti.

Numerička analiza podataka

- Podatci se numerički opisuju
 1. srednjim vrijednostima (mjerama centralne tendencije),
 2. mjerama disperzije,
 3. mjerama rasporeda podataka (asimetrije i zaobljenosti),
 4. mjerama koncentracije

Srednje vrijednosti

- Mogu biti:
 - potpune (aritmetička, geometrijska i harmonijska sredina) i
 - Položajne (mod i medijan).
- Odabir pojedine srednje vrijednosti za primjenu na određenom skupu podataka ovisi o:
 - vrsti varijable,
 - prisutnosti netipičnih vrijednosti u nizu,
 - stupnju varijabilnosti podataka i sl.

Srednje vrijednosti

- Odabir ovisi o vrsti varijable:
 - za **analizu** nominalnog niza moguće koristiti samo mod,
 - za **redosljedni** se koriste mod i medijan,
 - za **numerički niz** se koriste sve navedene srednje vrijednosti

Srednje vrijednosti

- Sve srednje vrijednosti izražene su u mjernim jedinicama pojave (varijable) za koju se određuju.
- Srednje vrijednosti koje se određuju za redoslijednu ili numeričku varijablu, poprimaju vrijednost između najmanje i najveće vrijednosti varijable za koju se izračunavaju (određuju).
- Nije rijedak slučaj da srednja vrijednost, koja predstavlja niz varijabilnih podataka, nije jednaka niti jednoj vrijednosti u nizu za koji se izračunava

Aritmetička sredina

- \bar{x}
- potpuna srednja vrijednost
- Nastaje dijeljenjem zbroja svih vrijednosti numeričke varijable i opsega skupa N .
- Zbroj svih vrijednosti numeričke varijable naziva se total T .
- Stoga je aritmetička sredina omjer totala i opsega skupa
- $$\bar{x} = \frac{T}{N}$$

Aritmetička sredina

- Računa se na dva načina:
 1. Za negrupirane numeričke podatke se izračunava jednostavna aritmetička sredina, a
 2. za distribuciju frekvencija ponderirana (vagana) aritmetička sredina.
- Pritom su ponderi apsolutne ili relativne frekvencije ili njima proporcionalne veličine

Primjer: aritmetička sredina negrupiranih podataka

- Bahovec i Erjavec (2015) Statistika; str. 56

Primjer: aritmetička sredina negrupiranih podataka

Primjer: aritmetička sredina grupiranih podataka

- Bahovec i Erjavec (2015) Statistika; str. 58



Aritmetička sredina

- Posebni slučajevi:
 1. Aritmetička sredina aritmetičkih sredina
 2. Aritmetička sredina relativnih brojeva koordinacije
 3. Aritmetička sredina postotaka

Prosjek aritmetičkih sredina

Primjer: prosjek aritmetičkih sredina

- Bahovec i Erjavec (2015) Statistika; str. 65



Sredina relativnih brojeva koordinacije

- Relativni brojevi koordinacije R_i su omjeri vrijednosti dviju povezanih pojava:

$$R_i = \frac{V_i}{B_i}, \quad i = 1, 2, \dots, k$$

- Gustoća naseljenosti: omjer broja stanovnika i površine
- Često korišteni u demografiji

Relativni brojevi koordinacije

- *stopa nataliteta* = $\frac{\text{broj živorodjenih u godini}}{(\text{ukupna})\text{populacija sredinom godine}} 1000$
- *stopa fertiliteta* = $\frac{\text{broj živorodjenih u godini}}{\text{populacija žena dobi 15–44 sredinom godine}} 1000$

Mortalitet

- *opća stopa mortaliteta* = $\frac{\text{broj umrlih u godini}}{\text{populacija sredinom godine}} 1000$

- Specifične stope

dobno (spolno) specifična stopa mortaliteta = $\frac{\text{broj umrlih određene dobi (spola) u godini}}{\text{populacija određene dobi (spola) sredinom godine}} 1000$

- *mortalitet dojenčadi* = $\frac{\text{broj umrle djece u prvoj godini života}}{\text{broj živorođenih u godini}} 1000$

- *specifični mortalitet* = $\frac{\text{broj umrlih od određenog uzroka u godini}}{\text{populacija sredinom godine}} 100000$

Specifične stope mortaliteta

- *proporcionalni mortalitet* = $\frac{\text{broj umrlih od određenog uzroka u godini}}{\text{ukupan broj umrlih u godini}} 1000$
- *letalitet* = $\frac{\text{broj umrlih od određenog stanja/uzroka u godini}}{\text{broj oboljelih od istog stanja/uzroka u godini}} 100$

Odnos nataliteta i mortaliteta

- U statističkom smislu, dva relativna broja koordinacije se mogu oduzimati (zbrajati) ukoliko imaju istu bazu

- $$\textit{stopa nataliteta} = \frac{\textit{broj živorodenih u godini}}{\textit{(ukupna) populacija sredinom godine}} 1000$$

- $$\textit{opća stopa mortaliteta} = \frac{\textit{broj umrlih u godini}}{\textit{populacija sredinom godine}} 1000$$

Odnos nataliteta i mortaliteta

- Prirodni prirast stanovništva
- *prirodni prirast = broj živorođenih – broj umrlih*
- *stopa prirodnog prirasta = natalitet – mortalitet*
- *vitalni indeks = $\frac{\text{broj živorođenih}}{\text{broj umrlih}} 100$*

Primjer 1.

- U 2017 godini u Zadarskoj županiji je rođeno 1537 djece. Umrlo je 1116 muškaraca i 1043 žene. Ako je poznato da županija ima 82830 muškaraca i 85842 žene (izvor: Državni zavod za statistiku), odredite:
 - Stopu nataliteta
 - Stopu mortaliteta
 - Stopu mortaliteta žena
 - Prirodni prirast stanovništva
 - Stopa prirodnog prirasta stanovništva
 - Vitalni indeks

Primjer 1.

- $stopa\ nataliteta = \frac{broj\ živorođenih\ u\ godini}{(ukupna)populacija\ sredinom\ godine} 1000$

- $stopa\ nataliteta = \frac{1537}{82830+85842} 1000 = 9,11$

- $opća\ stopa\ mortaliteta = \frac{broj\ umrlih\ u\ godini}{populacija\ sredinom\ godine} 1000$

- $opća\ stopa\ mortaliteta = \frac{1116+1043}{82830+85842} 1000 = 12,8$

- Zaokružiti na dvije decimale

-

-

Primjer 1.

- *dobno (spolno) specifična stopa mortaliteta* = $\frac{\text{broj umrlih određene dobi (spola) u godini}}{\text{populacija određene dobi (spola) sredinom godine}} 1000$
- *stopa mortaliteta žena* = $\frac{\text{broj umrlih žena u godini}}{\text{populacija žena sredinom godine}} 1000$
- *stopa mortaliteta žena* = $\frac{1043}{85842} 1000 = 12,15$

Primjer 1.

- *prirodni prirast* = broj živorođenih – broj umrlih
- *prirodni prirast* = 1537 – 2159 = –622

- *stopa prirodnog prirasta* = natalitet – mortalitet
- *stopa prirodnog prirasta* = 9,11 – 12,8 = –3,69

- *vitalni indeks* = $\frac{\text{broj živorođenih}}{\text{broj umrlih}} 100$
- *vitalni indeks* = $\frac{1537}{2159} 100 = 71,19$

Sredina relativnih brojeva koordinacije

Sredina relativnih brojeva koordinacije

Primjer 2.

- Zadan je ukupan broj stanovnika i stopa nataliteta u tri županije:

Županija	Broj stanovnika	Stopa nataliteta
Osječko-baranjska	283035	8,43
Vukovarsko-srijemska	159213	8,23
Brodsko-posavska	143827	8,88

Izvor: Državni zavod za statistiku

- Odredite prosječnu stopu nataliteta u četiri županije zajedno.
- Koliki je ukupan broj rođene djece u Brodsko-Posavskoj županiji?

Primjer 2.

- $stopa nataliteta = \frac{broj\ živorođenih}{(ukupna)populacija} 1000$
- $broj\ živorođenih = \frac{stopa\ nataliteta \cdot (ukupna)populacija}{1000}$
- Za Osječko-baranjsku županiju:
- $broj\ živorođenih = \frac{8,43 \cdot 283035}{1000} = 2385,985 \approx 2386$

Primjer 2.

Županija	Broj stanovnika	Stopa nataliteta	Broj živorođenih
Osječko-baranjska	283035	8,43	2386
Vukovarsko-srijemska	159213	8,23	1310
Brodsko-posavska	143827	8,88	1277
Ukupno	586075	-	4973

- $\bar{R} = \frac{4973}{586075} 1000 = 8,48$

Sredina postotaka

Mod

- Položajna srednja vrijednost niza.
- Povezana je s učestalosti pojavljivanja vrijednosti kvantitativne varijable ili modaliteta kvalitativne varijable.
- Modalitet (ili vrijednost) obilježja koji se najčešće pojavljuje u nizu je mod
- Može se odrediti samo ako u nizu postoje barem dva jednaka podatka.
- Ako se u nizu više podataka pojavljuje jednako mnogo puta, niz ima više modalnih vrijednosti, odnosno govori se o **multimodalnom nizu**

Mod

- Mod je izražen u mjernim jedinicama varijable za koju se određuje (izračunava).
- Kao položajna srednja vrijednost nije osjetljiv na izrazito velike niti na izrazito male vrijednosti varijable
- Mod nominalnog i redosljednog niza (ako su modaliteti kodirani i grupirani) lako se određuje jer se lako uoči i nije ga potrebno izračunavati
- To vrijedi i za distribuciju frekvencija kod koje svaka vrijednost ima posebnu grupu. Tada je mod modalitet (ili vrijednost) obilježja s najvećom frekvencijom

Mod

- Za distribuciju frekvencija kontinuirane numeričke varijable, odnosno općenito za **distribuciju frekvencija s razredima, mod se izračunava**
- Da bi se izračunao mod distribucije frekvencija s razredima, potrebno je odrediti modalni razred. Modalni razred je razred s najvećom (korigiranom) frekvencijom. Mod se izračunava prema izrazu

$$Mo = L_1 + \frac{(b - a)}{(b - a) + (b - c)} \cdot i,$$

Primjer: mod

- Bahovec i Erjavec (2015) Statistika; str. 81

Primjer: mod

- Bahovec i Erjavec (2015) Statistika; str. 81

Primjer: mod

- Bahovec i Erjavec (2015) Statistika; str. 84



Medijan i kvantili

- Kvantili dijele niz uređenih podataka na j jednakobrojnih dijelova. U kvantile se ubrajaju:
 - medijan,
 - kvartili,
 - decili i
 - Percentili
- **Medijan** (Me) je položajna srednja vrijednost koja dijeli niz uređen po veličini na dva jednakobrojna dijela.
- **Kvartili** ga dijele na 4,
- **decili** na 10, a
- **percentili** na 100 dijelova

Medijan i kvantili

- Medijan i kvantili mogu se odrediti i za redosljedno obilježje jer se modaliteti redosljednog obilježja mogu poredati po intenzitetu mjerenog svojstva (npr. ocjene na ispitu, stupanj zadovoljstva proizvodom ili uslugom i sl.).
- Medijan je izražen u mjernim jedinicama varijable za koju se određuje (izračunava) i nije osjetljiv na izrazito velike niti na izrazito male vrijednosti obilježja.
- Poprima vrijednosti između najmanje i najveće vrijednosti obilježja i ne mora biti jednak niti jednoj vrijednosti u nizu.

Medijan i kvantili

- **Medijan i kvantili negrupiranih (pojedinačnih) podataka**, uređenih po veličini, određuju se ovisno o tome je li omjer cijeli broj (INT, engl. integer – cijeli dio broja) ili nije cijeli broj.

Medijan i kvantili

Medijan i kvantili

- Izračun medijana i ostalih kvantila za **distribuciju frekvencija s razredima** provodi se na temelju kumulativnog niza kako bi se odredio **medijalni (ili kvantilni) razred**. **Medijalni razred** je razred čija kumulativna frekvencija prvi put uključuje $N/2$ -tu jedinicu. Medijan distribucije frekvencija s razredima računa se prema izrazu:

Primjer: medijan

- Bahovec i Erjavec (2015) Statistika; str. 88



Primjer: medijan

- Bahovec i Erjavec (2015) Statistika; str. 91



Primjer: medijan

- Bahovec i Erjavec (2015) Statistika; str. 92



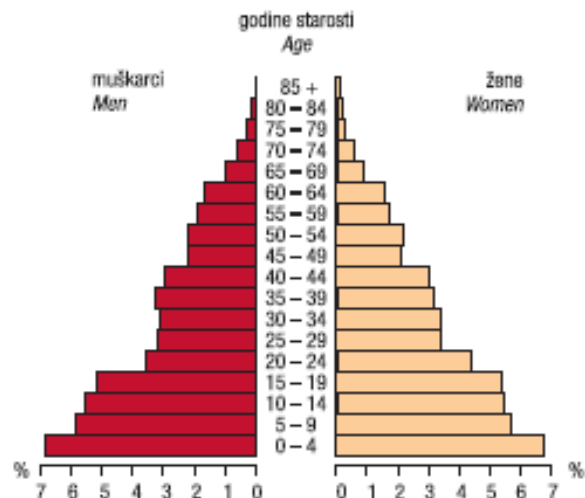


Grafički prikazi

- **dobno spolna piramida**
- Lexisov dijagram
- Grafikoni jednostavnih/višestrukih stupaca (za kvalitativne varijable)
- Histogram i poligon frekvencija (za numeričke varijable poput dobi)
- Grafički prikaz vremenskih nizova

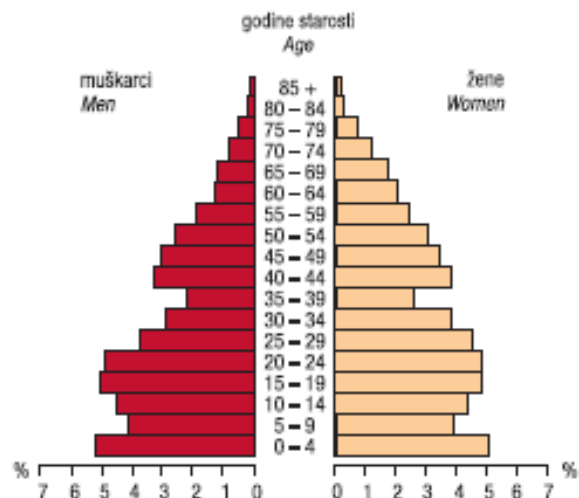
POPIS 1900.
1900 CENSUS

ukupno
Total 3 161 456



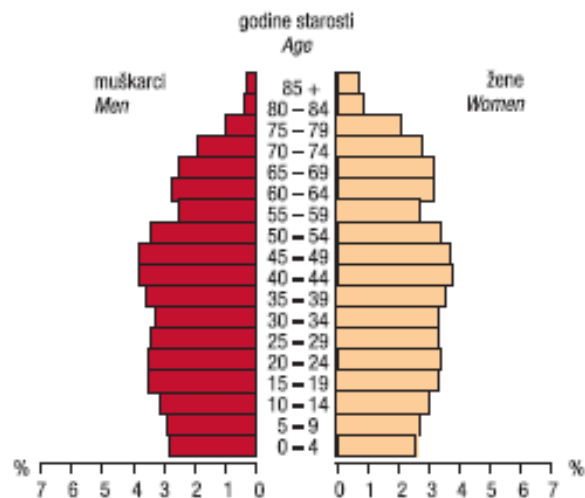
POPIS 1953.
1953 CENSUS

ukupno
Total 3 936 022



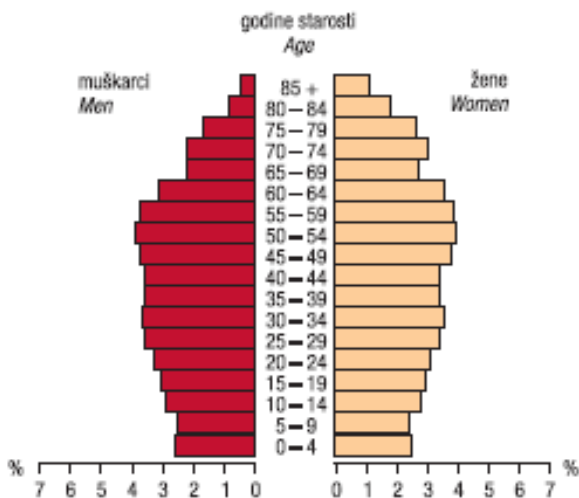
POPIS 2001.
2001 CENSUS

ukupno
Total 4 437 460



POPIS 2011.
2011 CENSUS

ukupno
Total 4 284 889



Grafički prikazi

- dobno spolna piramida
- **Lexisov dijagram**
- Grafikoni jednostavnih/višestrukih stupaca (za kvalitativne varijable)
- Histogram i poligon frekvencija (za numeričke varijable poput dobi)
- Grafički prikaz vremenskih nizova

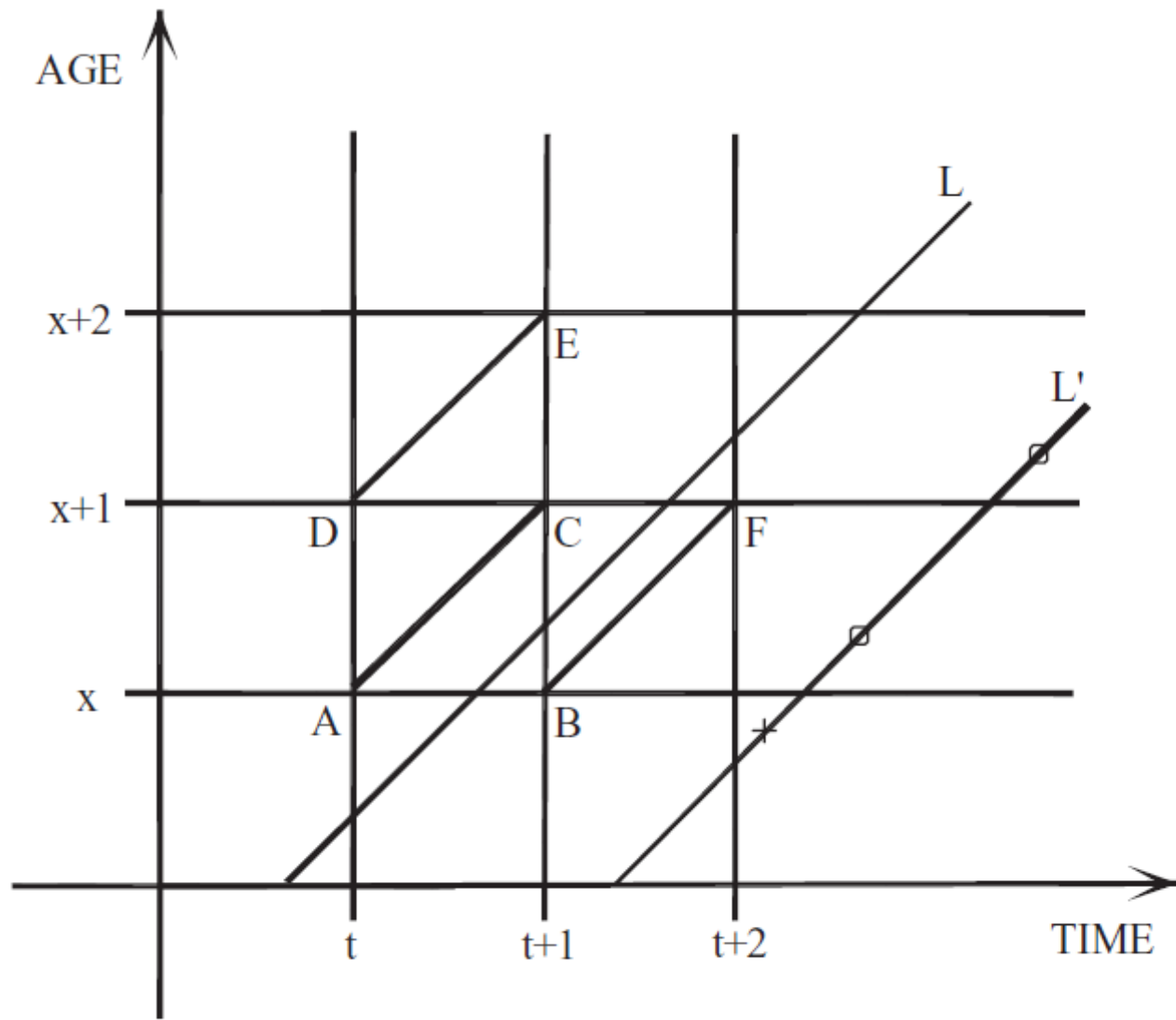


FIGURE 1. Lexis Diagram.