



# STATISTIKA

## UVODNO PREDAVANJE

# POJAM I PREDMET STATISTIKE

- Začeci statistike sreću se još u najstarijoj civilizaciji. U vreme kada su se narodi počeli udruživati u veće zajednice i stvarati države, cilj statistike je bio da ustanovi ratnu i ekonomsku snagu nacije i države.
- Sa stvaranjem političke vlasti razvijala se i ideja evidentiranja, tj. prebrojavanja stanovnika, utvrđivanja obima bogatstva i naseljene teritorije.
- Statistika je uglavnom zadovoljava potrebe administracije, vojske i privrede.
- Najnoviju etapu razvoja statistike kao metoda karakteriše njena primena u svim drugim naukama. Posebno je česta njena primena u poljoprivredi.

# POJAM I PREDMET STATISTIKE

- Reč statistika potiče od latinske reči **STATUS** koja ima dva značenja:
- **stanje** (u početnoj fazi razvoja statistika se bavila prikupljanjem numeričkih podataka pomoću kojih se moglo opisati stanje neke pojave)
- **država** (statistika je obezbeđivala podatke koje je državna administracija koristila u poslovima upravljanja državom).
- Statistika bi se danas mogla definisati kao **nauka o metodima istraživanja masovnih pojava**.
- **Masovnost neke pojave ispoljava se kroz njeno prisustvo kod velikog broja individualnih objekata** (stanovništvo, proizvodnja, uvoz robe, izvoz robe i sl).

# POJAM I PREDMET STATISTIKE

- Jedan od zadataka statistike je da ukaže na *postupke prikupljanja, sređivanja i prezentacije podataka*, kao i *metode određivanja parametara skupova*. Ovom problematikom se bavi *deskriptivna statistika*.
- Posle prikupljanja, sređivanja i prikazivanja podataka vrši se njihova analiza prema izabranom modelu i izvode zaključci. Ovom problematikom se bavi *analitička statistika*.

# STATISTIČKI SKUP

- Skup elemenata na kojima se izvesna pojava manifestuje zove se **statistički skup** ili **statistička masa** tj. **populacija**.
- Pojedinačni elementi iz kojih se skup sastoji, nazivaju se **elementi, objekti ili jedinice skupa**.
- ***Statistički skup treba definisati prostorno, vremenski i pojmovno.***
- Prostorno definisati statistički skup znači odrediti ga u geografskom smislu.
- Vremenski definisati statistički skup znači utvrditi momenat ili period u kojem će se posmatrati jedinice.
- Prema vremenskoj komponenti statistički skupovi se dele na **skupove stanja** i **skupove kretanja**.
- Pojmovno definisati statistički skup znači odrediti obeležja koja mora imati element da bi bio uključen u određeni statistički skup.

# OBELEŽJA

- *Odlike po kojima se posmatrane jedinice razlikuju nazivaju se OBELEŽJIMA.*
- U statistici, obeležja se označavaju velikim slovima A, B, C itd. ili X, Y, Z itd., a njihove pojedinačne vrednosti malim slovima sa indeksima (X:  $x_1, x_2, \dots, x_n$ ).
- Sa gledišta statističke obrade podataka najvažnija je klasifikacija obeležja na **numerička (brojčana, kvantitativna) i atributivna (opisna, kvalitativna)**.
- Numerička obeležja se iskazuje brojčano dok se atributivna obeležja iskazuju rečima (pol, školska sprema, boja kose, očiju, zanimanje, vrsta robe,...).
- Numerička obeležja se dele na prekidna i neprekidna.
- Ako **numeričko obeležje** uzima izolovane vrednosti, naziva se **prekidno**, a ako može uzeti ma koju vrednost iz nekog intervala naziva se **neprekidno**.

# MERNE SKALE

- Statistička obeležja se iskazuju na različitim vrstama skala i u različitim jedinicama mere.
- U praksi se koriste najčešće četiri merne skale: nominalna, ordinalna, intervalna i skala odnosa (proporcionalna).
- Nominalna i ordinalna skala se koriste kod atributivnih obeležja.
- Intervalna skala i skala odnosa se primenjuju kod numeričkih obeležja pa se nazivaju **kardinalne skale**.

# MERNE SKALE

- Nominalna skala se koristi kod atributivnog obeležja kada između kategorija ne postoji uređenost.
- Posmatrani elementi se sortiraju po kategorijama kojima se dodeljuju brojevi, kao kodovi za razgraničenje, a ne kao pokazatelji kvaliteta modaliteta.
- Kategorije se mogu navesti bilo kojim redom, pa prema tome može im se pridružiti i bilo koji broj. (obeležje pol)



# MERNE SKALE

- U slučaju ordinalne skale svaka kategorija ima tačno određen položaj u odnosu na ostale kategorije, pa se kategorije moraju rangirati prema značaju u utvrđenom kriterijumu.
- Brojevi označavaju rang, ali ne i jačinu razlikovanja, tj. ne može se meriti rastojanje između kategorija.
- Ovakva vrsta merenja omogućava ne samo grupisanje posmatranih jedinica u klase već i uređenje samih kategorija. (obeležje školska sprema)

# MERNE SKALE

- Za intervalnu skalu neophodno je da postoji fizička jedinica mere koja je prihvaćena kao standard.
- U ovom slučaju posmatrane jedinice se klasifikuju u odnosu na stepen posedovanja posmatrane karakteristike (kao kod ordinalne skale), ali je moguće i precizno odrediti rastojanje između klasa. (merenje temperature vazduha)

# MERNE SKALE

- Skalu odnosa pored jedinice mere karakteriše i postojanje apsolutne nulte tačke.
- Kod ove skale može se iskazati proporcionalni odnos vrednosti obeležja, pa se naziva i proporcionalna skala.
- Ova skala je **najpreciznija i najčešće korišćena** merna skala za kvantitativne promenljive.

# DISTRIBUCIJA FREKVENCIJA

- *Frekvencija je broj koji pokazuje koliko se puta jedna ista vrednost obeležja ili isti modalitet javlja u seriji.*
- *Distribucija frekvencija pokazuje kako su raspoređene jedinice posmatranja po pojedinim vrednostima ili modalitetima obeležja.*
- *Frekvencije mogu biti apsolutne, relativne i kumulativne.*
- *Vrednosti frekvencija koje iskazuju broj jedinica sa određenom vrednošću tj. modalitetom obeležja nazivaju se apsolutnim i označavaju se sa  $f$ .*
- *Frekvencije se mogu iskazati i u vidu svojih **relativnih vrednosti**, koje se dobijaju deljenjem pojedinačnih apsolutnih frekvencija sa zbirom svih apsolutnih frekvencija.*

$$p_i = \frac{f_i}{\sum f_i}$$

# DISTRIBUCIJA FREKVENCIJA

- *Kumulativne frekvencije* se dobijaju sukcesivnim sabiranjem naniže ili naviše apsolutnih ili relativnih frekvencija.
- Kod *kumuliranja "ispod" (naniže)*, vrednost kumulante ( $k_i$ ) pokazuje koliko je jedinica raspoređeno od najniže do posmatrane vrednosti obeležja.
- Kod kumuliranja *"iznad" (naviše)*, polazi se od poslednjeg člana u nizu frekvencija i ide prema prvom. U tom slučaju vrednost kumulante pokazuje koliko ima jedinica od najviše do posmatrane vrednosti obeležja.

# DISTRIBUCIJA FREKVENCIJA

- Tabela 1: Formiranje kumulanti od apsolutnih frekvencija

$x_i$	$f_i$	$k_i \downarrow$	$k_i \uparrow$
$x_1$	$f_1$	$k_1 = f_1$	$k_1 = f_5 + f_4 + f_3 + f_2 + f_1$
$x_2$	$f_2$	$k_2 = f_1 + f_2$	$k_2 = f_5 + f_4 + f_3 + f_2$
$x_3$	$f_3$	$k_3 = f_1 + f_2 + f_3$	$k_3 = f_5 + f_4 + f_3$
$x_4$	$f_4$	$k_4 = f_1 + f_2 + f_3 + f_4$	$k_4 = f_5 + f_4$
$x_5$	$f_5$	$k_5 = f_1 + f_2 + f_3 + f_4 + f_5$	$k_5 = f_5$