

I predavanje : ZADATAK STATISTIKE

Statistika je sistematizovani skup znanja o statističkim metodama koje se bave kvantitativnim istraživanjem masovnih pojava.

Zadatak statistike je da upozna studente sa statističkom metodologijom i primenom statističkih modela na brojne fenomene koji utiču na život i zdravlje ljudi. Mnogo važniji cilj je usvajanje statističkog načina mišljenja – specijalne tehnike mišljenja i rada bez koje nema naučnog mišljenja. Statistika je osnov naučne metodologije, t.j. primenljiva je u svim disciplinama.

Masovnost u medicini treba shvatiti uslovno.
Kvantitativnost znači da je sve izraženo brojevano.

U medicini postoji više vrsta istraživanja ili eksperimenata:

1. laboratorijski
2. klinički
3. terenski

Osnovna statistička metodologija ili faze ispitivanja:

1. proučavanje
 - a) bazične oblasti
 - b) statistike i informatike
 2. planiranje
 - a) kadrova
 - b) minimalnog potrebnog uzorka (minimalnog broja jedinica posmatranja)
 - c) utvrđivanje graške prve vrste (α)
 3. prikupljanje materijala (podataka)
 4. obrada podataka (statistička obrada u užem smislu)
 - a) sredjivanje podataka
 - b) opisivanje (deskripcija)
 - c) analiza unutargrupnih i medjugrupnih odnosa
 - d) generalizacija zaključaka
- }deskriptivna statistika
}inferencijalna statistika

Obrada podataka

- a) Sredjivanje podataka je prva faza obrade podataka koja podrazumeva:
 - grupisanje ili distribuciju frekvencija (naječementarniji način obrade podataka) i
 - prikazivanje rezultata: tabelarno i grafički.
- b) Opisivanje podataka: obavezno odrediti parametre statističkog skupa koje čine 3 karakteristike svakog skupa: njegova veličina (N), jedna mera centralne tendencije (\bar{x}), jedna mera varijabiliteta (SD).
- c) Analiza – procena unutargrupnih i medjugrupnih odnosa, gde se zaključci donose imajući u vidu maksimalno dozvoljenu grešku.
- d) Generalizacija zaključaka.

Definicija: *statistika je nauka koja se bavi kvantitativnim istraživanjem pojava u cilju njihove deskripcije, analize i generalizacije zaključaka.*

Osnovni pojmovi

1. Statistički skup je predmet proučavanja statistike. On predstavlja celinu sastavljenu od istovrsnih elemenata sa zajedničkom promenljivom karakteristikom. Elementi statističkog skupa su istovrsni ali ne i istovetni. Osnovni statistički skup je skup svih istovrsnih jedinica posmatranja čija je veličina N . Uzorak je reprezentativni deo osnovnog statističkog skupa i njegova veličina je n .
2. Jedinice posmatranja su istovrsni ali nikada istovetni elementi statističkog skupa. One su nosioci karakteristika statističkog skupa.
3. Obeležje posmatranja je karakteristika jedinica posmatranja, bez obzira da li su kvantitativnog ili kvalitativnog tipa.

Po svojoj prirodi obeležja mogu biti

kvantitativna (numerička) : inteligencija, starost	kontinuirana (neprekidna): godine
kvalitativna (atributivna) : pol, boja očiju, kose...	diskontinuirana (prekidna): broj trombocita, broj otkucaja srca

Po svojim karakteristikama obeležja se dele na:

- a) atributivna (kvalitativna) diskontinuirana koja variraju svojom kategorijom,
 - b) numerička (kvantitativna) diskontinuirana, i
 - c) numerička (kvantitativna) kontinuirana
4. Varijabilitet (varijabla) promenljivost obeležja posmatranja od jedinice do jedinice posmatranja.
Numerička obeležja variraju kvantitetom, a atributivna svojim kvalitetom ili klasom. Varijabla je jedan iznos ili kvantitet numeričkog obeležja, odnosno jedan modalitet (klasa, gradacija, vid, kategorija) atributivnog obeležja.
 5. Grupisanje je razvrstavanje jedinica posmatranja po varijablama obeležja posmatranja. Grupisanje se vrši u grupe ili grupne intervale. Atributivna u numerička diskontinuirana obeležja grupišu se u grupe čije izvorne varijable imaju mali interval variranja. Kontinuirana numerička obeležja grupišu se u grupne intervale. Broj jedinica posmatranja koji odgovara jednoj gradaciji obeležja posmatranja naziva se frekvencija.
 6. Frekvencije mogu biti
 - apsolutne odnosno - parcijalne
 - relativne - kumulativne

Apsolutna frekvencija je empirijska učestalost. odnosno direktan rezultat razvrstavanja jedinica posmatranjapo grupama ili grupnim intervalima.

Relativna frekvencija je odnos apsolutne učestalosti i ukupnog broja jedinica posmatranja osnovnog skupa.

Parcijalna frekvencija odgovara jednom broju učestalosti.

Kumulativna frekvencija je sukcesivni niz parcijalnih parcijalnih frekvenci i dobija se sukcesivnim sabiranjem parcijalnih frekvencija.

Distribucija frekvencija tj. raspodela učestalosti je kompletan prikaz rasporedjivanja jedinica posmatranja statističkog skupa po grupama ili grupnim intervalima.

II predavanje : DESKRIPCIJA - ODREĐIVANJE PARAMETARA STATISTIČKOG SKUPA

I	Relativni brojevi	Pravi deskriptivni statistički metodi koji opisuju statistički skup koristeći proste matematičke operacije.
II	Mere centralne tendencije	
III	Mere varijabiliteta	
IV	Mere oblika raspodele	
V	Mere za procenu parametara osnovnog skupa	Služe za opisivanje osnovnog skupa, dakle deskriptivni su samo po svom cilju. Za izračunavanje je potrebno poznavanje složenih matematičkih operacija. Pomoću ovih metoda zaključak se sa uzorka proširuje na osnovni statistički skup.
VI	Metodi za ocenu oblika empirijskih raspodela	

I Relativni brojevi su pravi ali ne i obavezni deskriptivni statistički metodi. Koriste se za ispitivanje strukture i obima pojave (odnosa dela prema celini). Omogućavaju poredjenje istovrsnih ali i raznovrsnih pojava bez obzira na njihov intenzitet tj. nivo. Relativni broj predstavlja odnos dva apsolutna broja.

V_r = brojilac koji se upoređuje

V_b = imenilac (bazna vrednost) sa kojom se upoređuje.

$$R_b = \frac{V_r}{V_b} \times 100$$

R_b = mali decimalni broj koji se množi dekadnim brojem (100) i izražava se u %

Vrste relativnih brojeva

1. KVOTA – pokazatelj strukture , predstavlja odnos dela prema celini i uvek se izražava u %.

2. INDEKSI – relativni brojevi kojima se upoređuje obim ili nivo dve istoimene pojave. Zavisno od prirode pojave mogu biti:

a) vremensko- dinamski

b) teritorijalno geografski

Prema načinu poredjenja indeksi su: bazni i lančani.

3. STATISTIČKI KOEFICIJENTI – koriste se za poredjenje obima ili nivoa dve ili više raznoimenih pojava od kojih bar jedna mora biti masovna. Karakteristični su za terenska ispitivanja.

II Mere centralne tendencije broj (iznos, vrednost) koji na najbolji mogući način prezentuje sve ostale vrednosti obeležja posmatranja statističkog skupa. Njihova reprezentativnost posledica je činjenice da sve vrednosti obeležja posmatranja pokazuju tendenciju da se grupišu

oko one vrednosti koja se nalazi u sredini prirodnog intervala variranja (raspon od najmanje do najveće vrednosti uočen u empirijskoj sredini).

Prema načinu određivanja dele se na:

1. pozicione, koje se određuju na osnovu položaja, mesta koje zauzimaju u distribuciji frekvencija – medijana i mod
2. računске, koje se određuju na osnovu specifično primenjenih računskih operacija koje obavezno obuhvataju sve vrednosti obeležja posmatranja statističkog skupa. To su aritmetička, geometrijski i harmonijska sredina. **SVE MERE CENTRALNE TENDENCIJE SU APSOLUTNOG KARAKTERA** tj. izražavaju se u istim mernim jedinicama kao i posmatrano empirijsko obeležje.

Aritmetička sredina

\bar{x} (x bar, x srednje) predstavlja količnik zbira vrednosti obeležja posmatranja svake jedinice posmatranja i veličine statističkog skupa.

Izračunavanje:

a) za *negrupisane podatke* $\bar{x} = \frac{\sum x}{n}$

b) za *grupisane podatke* $\bar{x} = \frac{\sum fx}{\sum f}$

Aritmetička sredina može se računati i za više skupova i to je aritmetička sredina aritmetičkih sredina.

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{n_1 + n_2 + n_3}$$

Aritmetička sredina može se računati i za proporciju i procenete

$$p = \frac{\sum np}{\sum n}$$

Aritmetička sredina predstavlja najosetljiviju i najprecizniju srednju vrednost. Njena preciznost rezultat je činjenice da je suma odstupanja svake pojedinačne vrednosti od aritmetičke sredine jednaka nuli:

$$\Sigma(\bar{x} - x_n) = 0$$

a zbir kvadrata odstupanja jednak je minimumu: $\Sigma(x - x_n)^2 = \min$

Aritmetička sredina je reprezentativna samo ako se primenjuje na homogene skupove (koji su raspoređeni po tipu normalne raspodele), što znači da njihov varijabilitet ne prelazi 30%. Razlog tome je što na aritmetičku sredinu utiču vrednosti obeležja posmatranja merene na svakoj jedinici posmatranja statističkog skupa, pa samim tim i ekstremne vrednosti koje vuku aritmetičku sredinu ka sebi.

Medijana

Centralna srednja vrednost je vrednost obeležja posmatranja one jedinice posmatranja koja se u uredjenom nizu podataka nalazi u sredini . Medijana se određuje u tri faze :

1. uredjivanje niza – rastućim ili opadajućim redosledom.
2. određivanje mesta medijane (M Med) – određivanje jedinice posmatranja koja se u tako uredjenom nizu nalazi u sredini.
3. očitavanje (za negrupisane podatke $\frac{n+1}{2}$ ili za grupisane podatke $\frac{\Sigma(f+1)}{2}$) ili izračunavanje medijane .
Ako je niz podataka paran, Med se računa kao aritmetička sredina vrednosti obeležja na mestu medijane.
Ako je niz podataka neparan, vrednost medijane se očitava direktno nakon određivanja mesta medijane.

MEDIJANA NE ZAVISI OD VREDNOSTI VEĆ OD BROJA OBELEŽJA POSMATRANJA.

Mod

Tipična srednja vrednost – vrednost obeležja posmatranja kojoj odgovara najveći broj jedinica posmatranja. To je vrednost obeležja posmatranja koja se najčešće javlja , tj. **ima najveću učestalost** (f).

MOD ZAVISI ISKLJUČIVO OD UČESTALOSTI.

III MERE VARIJABILITETA

Varijabilitet je promenljivost obeležja posmatranja od jedinice do jedinice posmatranja. Odstupanje vrednosti od centralne tendencije naziva se disperzija , disocijacija ili rasturanje. Parametri koji pokazuju gustinu grupisanja ili odstupanja pojedinih vrednosti obeležja posmatranja od aritmetičke sredine su :

- A. apsolutne mere varijabiliteta (izražavaju se istim jedinicama kao i jedinice posmatranja):
 - interval varijacije, raspon
 $I = \max - \min$
daje informacije samo o ekstremnim vrednostima.
 - standardna devijacija SD
 $SD = + \sqrt{SD^2}$
pozitivna vrednost korena varijanse
 - varijansa SD^2 – prava mera varijabiliteta predstavlja prosečno standardno odstupanje od aritmetičke sredine
 $SD^2 = \frac{\sum (x - \bar{x})^2}{n}$
- B. relativne mere varijabiliteta
 - koeficijent varijacije

$$CV = \frac{SD}{\bar{x}} \cdot 100$$

Standardna devijacija izražena u % aritmetičke sredine. Predstavlja meru homogenosti. Skup je homogen ako je $CV \leq 30\%$ **VARIJABILITET JE MANJI TAMO GDE JE MANJI CV.**

- standardizovana „z“ vrednost – standardizovano odstupanje od aritmetičke sredine izraženo u delovima standardne devijacije. Daje informaciju o odnosu jedinice posmatranja prema skupu. „z“ vrednosti ima onoliko koliko i jedinica posmatranja.

$$z = \frac{x - \bar{x}}{sd}$$

C. specifične

- standardna greška – najvažnija mera varijabiliteta u okviru analitike.
- kovarijansa SD_{xy} – MERA ZAJEDNIČKOG VARIJABILITETA DVA OBELEŽJA POSMATRANJA