STATISTIKA

STATISTIČKA OCENA NEPOZNATIH PARAMETARA OSNOVNOG SKUPA

POJAM STATISTIČKE OCENE

- Statističko ocenjivanje ima za cilj donošenje zaključaka o vrednostima nepoznatih parametara osnovnog skupa na osnovu uzorka.
- U tu svrhu iz osnovnog skupa izvlači se slučajan uzorak veličine n i na bazi njegovih realizacija $x_1, x_2, \ldots\ldots, x_n$ izračunava se približna vrednost nepoznatog parametra θ.
- Približna vrednost nepoznatog parametra θ određena na bazi uzoračke vrednosti naziva se **vrednost ocene nepoznatog parametra θ** i označava se sa $\hat{\theta}$

POJAM STATISTIČKE OCENE

- Statistička ocena nepoznatog parametra osnovnog skupa može biti **tačkasta i intervalna.**
- Tačkasta ocena parametra θ daje se u vidu samo jedne numeričke vrednosti
- Intervalna ocena se iskazuje u vidu intervala vrednosti lociranog između dveju tačaka na liniji realnih brojeva.
- Da bi ostvarena vrednost jedne statistike uzorka bila najbolja ocena nekog parametra skupa, potrebno je da ta statistika ima sledeća svojstva: **nepristrasnost, konzistentnost, efikasnost i dovoljnost.**

POJAM STATISTIČKE OCENE

- ***Ocena parametra skupa na osnovu statistike uzorka smatra se <u>nepristrasnom</u> ako je njeno matematičko očekivanje jednako parametru skupa, čija se vrednost ocenjuje***, odnosno ako je

$$M(\hat{\Theta}) = \theta$$

- Ocena koja ne ispunjava ovaj zahtev je pristrasna, što znači da sadrži sistematsku grešku.
- Razlika između matematičkog očekivanja ocene i vrednosti ocenjivanog parametra naziva se **pristrasnost ili bias ocene** i označava se sa

$$B(\hat{\Theta}) = M(\hat{\Theta}) - \theta$$

POJAM STATISTIČKE OCENE

- Umesto nepristrasne ocene može se koristiti **asimptotski nepristrasna ocena**. To je ocena čije matematičko očekivanje teži ocenjivanom parametru pri povećanju obima uzorka.
- **Konzistentnom ili saglasnom ocenom smatra se vrednost statistike uzorka koja pri povećanju obima uzorka teži u verovatnoći ocenjivanom parametru**, tj. za koju je:

$$\lim_{n\to\infty} P\left\{\left|\theta - \hat{\Theta}\right| < \varepsilon\right\} = 1$$

POJAM STATISTIČKE OCENE

- ***Efikasna ocena je ona nepristrasna ocena nepoznatog parametra θ koja ima najmanju varijansu, između svih mogućih nepristrasnih ocena traženog parametra.***
- ***Vrednost statistike uzorka je dovoljna ocena parametra skupa ako koristi sve informacije koje uzorak sadrži o tom parametru***.
- Osnovne metode za dobijanje tačkastih ocena su: metod momenata, metod maksimalne verodostojnosti i metod najmanjih kvadrata.

POJAM STATISTIČKE OCENE

- U slučaju malog broja posmatranja tačkasta ocena je slabo pouzdana.
- Zbog toga se preporučuje upotreba intervalnih ocena za koje se može odrediti stepen pouzdanosti.
- Tačkasta ocena je bolja kada je razlika $\left|\theta - \widehat{\Theta}\right|$ mala.
- Pozitivan broj ε koji karakteriše granicu ove razlike naziva se **tačnost ocene**.
- Što je ε manje ocena je tačnija. Tačnost ocene zavisi od obima uzorka.

POJAM STATISTIČKE OCENE

- Verovatnoća sa kojom greška ocene ne prelazi tačnost ocene naziva se **pouzdanost ili verovatnoća poverenja**

$$P\{\hat{\Theta}-\varepsilon<\theta<\hat{\Theta}+\varepsilon\}=1-\alpha.$$

- Verovatnoća suprotnog događaja α naziva se rizik.
- Stepen pouzdanosti određuje istraživač. U statistici se ocene najčešće daju sa pouzdanošću 0,90; 0,95 i 0,99.

POJAM STATISTIČKE OCENE

- Interval

$$I_{1-\alpha} = \left[\widehat{\Theta} - \varepsilon ; \widehat{\Theta} + \varepsilon\right]$$

koji sa verovatnoćom p=1-α pokriva nepoznati parametar, naziva se **interval poverenja ili interval pouzdanosti.**

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI POZNATOJ VARIJANSI

- Interval poverenja ocene aritmetičke sredine osnovnog skupa predstavlja interval koji pokriva srednju vrednost osnovnog skupa sa zadatom verovatnoćom $1-\alpha$.
- Nepoznati parametar μ ocenjuje se na osnovu slučajnog uzorka izvučenog iz osnovnog skupa.
- Na osnovu izvučenog uzorka izračunava se uzoračka sredina kao realizacija tačkaste ocene nepoznatog parametra μ.

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI POZNATOJ VARIJANSI

- Za uzorke veće od 30 jedinica njihove sredine raspoređuju se po normalnoj raspodeli oko sredine osnovnog skupa sa srednjom vrednošću μ i varijansom $\sigma_{\bar{x}}^2$
- Zato, iz tablica normalne raspodele za određeni nivo značajnosti (1-α) mogu se naći kvantili $\pm z_{\alpha/2}$ za koje je

$$P\left\{-z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma_{\bar{x}}} < z_{\alpha/2}\right\} = 1-\alpha$$

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI POZNATOJ VARIJANSI

- Odnosno

$$P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) = 1 - \alpha ,$$

- Odakle rešavanjem dobijamo

$$2F(z_{\alpha/2}) - 1 = 1 - \alpha$$

- odnosno

$$F(z_{\alpha/2}) = 1 - \alpha / 2.$$

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI POZNATOJ VARIJANSI

- Transformacijom prethodne jednakosti dobija se interval pouzdanosti za aritmetičku sredinu osnovnog skupa:

$$\bar{x} - z_{\alpha/2}\sigma_{\bar{x}} < \mu < \bar{x} + z_{\alpha/2}\sigma_{\bar{x}}$$

$\bar{x}$ - aritmetička sredina uzorka

z - faktor pouzdanosti

$\sigma_{\bar{x}}$ - standardna greška sredine

μ - aritmetička sredina osnovnog skupa

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI POZNATOJ VARIJANSI

- Kao što je ranije rečeno, za konačan osnovni skup standardna greška, u slučaju uzorka sa vraćanjem je: $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
- a u slučaju uzorka bez vraćanja: $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$
- Popravni faktor $\sqrt{\frac{N-n}{N-1}}$ se primenjuje ukoliko je stopa izbora veća od 0,05 ($n/N > 0{,}05$).

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI POZNATOJ VARIJANSI

- Za poznati obim osnovnog skupa može se odrediti i interval poverenja za ukupnu vrednost posmatranog obeležja X u analiziranom osnovnom skupu na sledeći način:

$$N(\bar{x} - z_{\alpha/2}\sigma_{\bar{x}}) < N\mu < N(\bar{x} + z_{\alpha/2}\sigma_{\bar{x}})$$

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI NEPOZNATOJ VARIJANSI

- Osnovni skup je normalno raspoređen sa nepoznatom sredinom i nepoznatom varijansom.
- Nepoznati parametri se izračunavaju iz uzorka kao nepristrasne tačkaste ocene.
- Statistika $T = \dfrac{\overline{X} - \mu}{S_{\overline{x}}}$

ima Studentovu t-raspodelu sa ν=n-1 stepeni slobode, pa se u ovom slučaju t-raspodela koristi za određivanje intervala poverenja nepoznatog parametra μ.

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI NEPOZNATOJ VARIJANSI

- Za zadati nivo značajnosti (1-α), iz tablica t-raspodele, moguće je naći kvantile $\pm t_{\alpha/2;\nu}$ za koje je

$$P\left\{-t_{\alpha/2;\nu} < \frac{\overline{X}-\mu}{S_{\bar{x}}} < t_{\alpha/2;\nu}\right\} = 1-\alpha$$

- Transformacijom ove relacije dobija se interval poverenja koji pokriva aritmetičku sredinu osnovnog skupa μ sa zadatom pouzdanošću (1-α).

$$\bar{x} - t_{\alpha/2;n-1} S_{\bar{x}} \le \mu \le \bar{x} + t_{\alpha/2;n-1} S_{\bar{x}}$$

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI NEPOZNATOJ VARIJANSI

- Formule za standardnu grešku sredine $s_{\bar{x}}$:
- za negrupisane podatke je

$$s_{\bar{x}} = \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n(n-1)}} \quad \text{ili} \quad s_{\bar{x}} = \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n(n-1)} \cdot \frac{N-n}{N}}$$

- a za grupisane podatke

$$s_{\bar{x}} = \sqrt{\frac{\sum f_i x_i^2 - n\bar{x}^2}{n(n-1)}} \quad \text{ili} \quad s_{\bar{x}} = \sqrt{\frac{\sum f_i x_i^2 - n\bar{x}^2}{n(n-1)} \cdot \frac{N-n}{N}}$$

INTERVAL POVERENJA ZA SREDNJU VREDNOST OSNOVNOG SKUPA - PRI NEPOZNATOJ VARIJANSI

- ukupna vrednost analiziranog obeležja određuje se tako što se granice intervala množe sa veličinom osnovnog skupa.

$$N\cdot(\bar{x}-t_{\alpha/2,n-1}S_{\bar{x}})\leq N\mu\leq N\cdot(\bar{x}+t_{\alpha/2,n-1}S_{\bar{x}})$$

- Kod velikih uzoraka (n≥30), za određivanje intervala poverenja nepoznate aritmetičke sredine osnovnog skupa pri nepoznatoj varijansi osnovnog skupa, mogu se koristiti kvantili normalne raspodele.