Statistični praktikum

Intervali zaupanja

Asistent dr. Kristina Veljković

Fakulteta za računalništvo in informatiko

NEZNANI PARAMETRI PORAZDELITVE

- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec.
- Z $(x_1, x_2, \ldots, x_n)$ označimo dobljeni vzorec konkretnih vrednosti.

NEZNANI PARAMETRI PORAZDELITVE

- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec.
- Z $(x_1, x_2, \ldots, x_n)$ označimo dobljeni vzorec konkretnih vrednosti.
- Ocenjevanje neznanih parametrov porazdelitve

Parameter	Cenilka $f(X_1, X_2, \ldots X_n)$	Ocena $f(x_1, x_2, \ldots x_n)$
Pričakovana vrednost μ	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Standardni odklon σ	$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$	$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$
Verjetnost p	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ $(X_i$ – indikatorji dogodka $A)$	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Interval zaupanja za μ (σ je znan)

- Za konstrukcijo intervala zaupanja uporabljamo dejstvo

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1).$$

- Slučajna spremenljivka X je normalno porazdeljena ali imamo dovolj veliki vzorec (za uporabo CLI).

Interval zaupanja za μ (σ je znan)

- Za konstrukcijo intervala zaupanja uporabljamo dejstvo

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1).$$

- Slučajna spremenljivka X je normalno porazdeljena ali imamo dovolj veliki vzorec (za uporabo CLI).
- Z verjetnostjo $1 - \alpha$ se μ nahaja v intervalu

$$[\overline{X} - \varepsilon, \overline{X} + \varepsilon],$$

oziroma odstopa za $\pm\varepsilon$ od cenilke $\hat{\mu} = \overline{X}$.

Interval zaupanja za μ (σ je znan)

- Za konstrukcijo intervala zaupanja uporabljamo dejstvo
$$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1).$$
- Slučajna spremenljivka X je normalno porazdeljena ali imamo dovolj veliki vzorec (za uporabo CLI).
- Z verjetnostjo $1-\alpha$ se μ nahaja v intervalu
$$[\overline{X}-\varepsilon, \overline{X}+\varepsilon],$$
oziroma odstopa za $\pm\varepsilon$ od cenilke $\hat{\mu}=\overline{X}$.
- Dobi se
$$\varepsilon = c\frac{\sigma}{\sqrt{n}},$$
kjer je $c = F^{-1}\left(1-\frac{\alpha}{2}\right)$ kvantil standardne normalne porazdelitve.

Interval zaupanja za μ (σ je znan)

- Interval zaupanja za μ

$$I_\mu = \left[\overline{X} - c\frac{\sigma}{\sqrt{n}},\ \overline{X} + c\frac{\sigma}{\sqrt{n}}\right],$$

  kjer je $c = F^{-1}\left(1 - \frac{\alpha}{2}\right)$.
- $1 - \alpha$ je stopnja zaupanja, α stopnja tveganja.

Interval zaupanja za μ (σ je znan)

- Interval zaupanja za μ

$$I_\mu = \left[\overline{X} - c\frac{\sigma}{\sqrt{n}},\ \overline{X} + c\frac{\sigma}{\sqrt{n}} \right],$$

  kjer je $c = F^{-1}\left(1 - \frac{\alpha}{2}\right)$.
- $1 - \alpha$ je stopnja zaupanja, α stopnja tveganja.
- Širina (dolžina) intervala zaupanja

$$\ell = \overline{X} + c\frac{\sigma}{\sqrt{n}} - \left(\overline{X} - c\frac{\sigma}{\sqrt{n}} \right) = 2c\frac{\sigma}{\sqrt{n}}.$$

- S povečanjem stopnje zaupanja se dobi širši interval zaupanja.

INTERVAL ZAUPANJA ZA μ (σ JE ZNAN)

- Primer: Signal intenzitete μ je poslan z lokacije A. Na lokaciji B se beleži sprejet signal. Zaradi šumenja signal zaznamo z naključno napako. Intenziteta signala na lokaciji B je normalno porazdeljena slučajna spremenljivka s povprečjem μ in standardnim odklonom 3. Da bi zmanjšali napako, isti signal neodvisno beležimo 10-krat. Dobili smo naslednje vrednosti intenzitete signala na lokaciji

$$B : 17, 21, 20, 18, 19, 22, 20, 21, 16, 19$$

Določi 95% interval zaupanja za μ.

INTERVAL ZAUPANJA ZA μ (σ JE ZNAN)

- Primer: Iz prejšnjih izkušenj vemo, da je teža lososov, gojenih v komercialni ribogojnici normalno porazdeljena s povprečjem, ki varira od sezone do sezone in konstantnim standardnim odklonom 140 g. Kako veliki vzorec potrebujemo, če želimo biti 90% prepričani, da je ocena povprečne teže lososov natančna na ± 50 g?

INTERVAL ZAUPANJA ZA μ (σ NI ZNAN)

- Za konstrukcijo intervala zaupanja uporabljamo dejstvo

$$\frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

- Slučajna spremenljivka X je normalno porazdeljena ali imamo dovolj veliki vzorec (za uporabo CLI).

INTERVAL ZAUPANJA ZA μ (σ NI ZNAN)

- Za konstrukcijo intervala zaupanja uporabljamo dejstvo

$$\frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

- Slučajna spremenljivka X je normalno porazdeljena ali imamo dovolj veliki vzorec (za uporabo CLI).
- Interval zaupanja za μ s stopnjo zaupanja $1 - \alpha$ je enak

$$I_\mu = \left[\overline{X} - c\frac{S}{\sqrt{n}},\ \overline{X} + c\frac{S}{\sqrt{n}}\right],$$

kjer je $c = t_{n-1;1-\frac{\alpha}{2}}$ kvantil Studentove porazdelitve z $n-1$ prostostnimi stopnjami.

INTERVAL ZAUPANJA ZA μ (σ NI ZNAN)

- Primer: Ameriška agencija za zaščito okolja je zaskrbljena zaradi količine polikloriranih bifenilov (PCB) v mleku doječih mater. Na vzorcu 30 žensk so dobili povprečje 6 in popravljeni standardni odklon 5 količine PCB-jev, merjeno v delih na milijon (ppm). Določi 99% interval zaupanja za povprečno količino PCB-jev v mleku populacije doječih mater. (Poliklorirani bifenili so strupene kemikalije, ki se kopičijo v maščobnem tkivu in mleku doječih mater.)

Interval zaupanja za delež p

- p je delež populacije z določeno lastnostjo.
- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec, kjer je X_i indikatorska slučajna spremenljivka

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}, i = 1, 2, \ldots, n.$$

INTERVAL ZAUPANJA ZA DELEŽ p

- p je delež populacije z določeno lastnostjo.
- Naj bo $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec, kjer je X_i indikatorska slučajna spremenljivka

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}, i = 1, 2, \ldots, n.$$

- Neznani delež p ocenjujemo z vzorčnim deležem

$$\hat{p} = \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Interval zaupanja za delež p

- Za konstrukcijo intervala zaupanja uporabljamo dejstvo (za dovolj veliko n)

$$\hat{p} \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

.

Interval zaupanja za delež p

- Za konstrukcijo intervala zaupanja uporabljamo dejstvo (za dovolj veliko n)
$$\hat{p} \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$
.
- Interval zaupanja za delež p s stopnjo zaupanja $1-\alpha$ je enak
$$I_p = \left[\hat{p} - c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right],$$
kjer je $c = F^{-1}\left(1 - \frac{\alpha}{2}\right)$ kvantil standardne normalne porazdelitve.

INTERVAL ZAUPANJA ZA DELEŽ p

- *Primer:* Uvoznik vina je dobil priložnost poceni nakupa velike količine vina Chateau Lafite Rothschild letnik 1947. Zaradi starosti vina se je del buteljk vina spremenil v kis (po videzu buteljke se ne more zaključiti, katere so pokvarjene). Uvoznik se je odločil pri nakupu naključno izbrati 200 buteljk. Od teh je bilo 18 pokvarjenih. Izračunaj 90% interval zaupanja za delež pokvarjenih buteljk vina.