

Zadaci za vježbu: pouzdani intervali

Vježbe 8

Napomena. Oznaka q_p označava kvantil s vjerojatnošću p *lijevo* od njega; to vrijedi i za navedene t - i χ^2 -kvantile. Intervale za proporciju računajte normalnom (Waldovom) aproksimacijom. Pri simulacijama postavite `set.seed(108)`.

1 Zadaci za rješavanje olovkom i papirom

Zadatak 1. Poznata preciznost uređaja.

Očitavanja koncentracije imaju normalnu razdiobu s nepoznatim očekivanjem μ i poznatom standardnom devijacijom $\sigma = 0.4$ mg/L. Za 16 očitavanja dobiveno je $\bar{x} = 5.2$ mg/L.

- (a) Izračunajte 95%-tni pouzdani interval za μ , uz $z_{0.975} = 1.96$.
- (b) Izračunajte 90%-tni interval, uz $z_{0.95} = 1.645$, i usporedite širine.
- (c) Objasnite zašto to nije interval koji mora sadržavati 95% pojedinačnih očitavanja.

Zadatak 2. Mali uzorak, nepoznata varijanca.

Vrijeme obrade pet uzoraka (min) je 8, 9, 10, 11, 12. Pretpostavite normalnu populaciju s nepoznatim očekivanjem i varijancom.

- (a) Izračunajte $\bar{x}$, S^2 i standardnu pogrešku sredine.
- (b) Odredite 95%-tni pouzdani interval za očekivano vrijeme; koristite $t_{0.975}(4) = 2.776$.
- (c) Usporedite njegovu širinu s intervalom dobivenim pogrešnom zamjenom tog kvantila s 1.96. Navedite pretpostavku koja opravdava točan t -interval.

Zadatak 3. Udio obojenih sjemenki.

U slučajnom uzorku od 200 sjemenki njih 80 je obojeno. Pretpostavite Bernoullijev model s nepoznatim parametrom p .

- (a) Procijenite p i njegovu standardnu pogrešku.
- (b) Izračunajte približni 95%-tni pouzdani interval te navedite granice kao proporcije i kao postotke. Koristite $z_{0.975} = 1.96$.
- (c) Provjerite da je u uzorku dovoljno opažanja obiju kategorija za normalnu aproksimaciju. Kako bi se širina promijenila za $n = 800$ i isti opaženi udio?

Zadatak 4. Interval za varijancu i standardnu devijaciju.

Za devet nezavisnih normalnih mjerenja mase dobiveno je $S^2 = 4 \text{ mg}^2$. Za normalni uzorak vrijedi $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$.

- (a) Iz nejednakosti $2.180 \leq 8S^2/\sigma^2 \leq 17.535$ izvedite 95%-tni interval za σ^2 . Brojevi su kvantili $\chi_{0.025}^2(8)$ i $\chi_{0.975}^2(8)$.
- (b) Odredite interval za σ i navedite jedinice obaju intervala.
- (c) Jesu li intervali simetrični oko S^2 , odnosno S ? Objasnite zašto pri preuređivanju nejednakosti veći kvantil daje donju granicu.

Zadatak 5. Planiranje preciznosti.

Želi se 95%-tni interval za očekivanje normalnih mjerenja širine najviše 0.8 mg/L. Standardna devijacija poznata je i iznosi 1.2 mg/L.

- (a) Odredite najmanji cijeli broj opažanja koji zadovoljava uvjet, uz $z_{0.975} = 1.96$. Razlikujte širinu od poluširine intervala.
- (b) Za drugi pokus planira se procjena nepoznate proporcije s poluširinom najviše 0.05. Koristeći $p(1-p) \leq 1/4$ odredite konzervativnu veličinu uzorka.
- (c) Koliko se približno mora povećati uzorak da se poluširina prepolovi? Objasnite pomoću faktora $1/\sqrt{n}$.

2 Zadaci za rad u R-u

Zadatak 1. Razina pouzdanosti mijenja širinu.

Koncentracije (mg/L) su:

```
x <- c(4.8, 5.2, 5.1, 4.9, 5.4, 5.0, 5.3, 4.7, 5.1, 5.2)
```

Pretpostavite normalnu populaciju nepoznate varijance.

- (a) Izračunajte 90%, 95% i 99% pouzdane intervale za očekivanje izravno pomoću `mean()`, `sd()` i `qt()`.
- (b) Provjerite intervale s `t.test(x, conf.level = ...)$conf.int`; u ovom zadatku koristite samo interval iz izlaza.
- (c) Nacrtajte tri intervala na jednom grafu i protumačite odnos između razine pouzdanosti i preciznosti.

Zadatak 2. Interval za intenzitet prebrojavanja.

Brojevi čestica u 40 uzoraka jednakog volumena zadani su:

```
x <- c(2, 4, 3, 5, 1, 4, 2, 3, 6, 2, 3, 4, 1, 5, 3, 2, 4, 3, 5, 2,
      4, 3, 2, 6, 1, 3, 4, 2, 5, 3, 2, 4, 3, 1, 5, 4, 2, 3, 4, 3)
```

Pretpostavite nezavisni Poissonov model s istim λ po uzorku.

- (a) Procijenite λ i odredite približni 95%-tni interval $\bar{x} \pm z_{0.975} \sqrt{\bar{x}/n}$.
- (b) Nacrtajte histogram i usporedite sredinu s uzoračkom varijancom; objasnite odakle dolazi standardna pogreška u navedenom intervalu.
- (c) Ako svaki uzorak ima volumen 2 mL, pretvorite procjenu i granice u broj čestica po mL. Razlikujte jedinicu parametra od broja opažanja n .

Zadatak 3. Normalna aproksimacija za proporciju.

U trima nezavisnim istraživanjima opaženo je 40 od 100, 160 od 400 i 2 od 20 jedinki s određenim obilježjem.

- (a) Izračunajte 95%-tne Waldove intervale prema formuli

$$\hat{p} \pm 1.96 \sqrt{\hat{p}(1 - \hat{p})/n}.$$

- (b) Prikažite intervale na grafu s označenim granicama 0 i 1. Usporedite širine prvih dvaju intervala.
- (c) Provjerite broj opažanja obiju kategorija. Objasnite problem trećeg intervala i zašto samo odsijecanje negativne granice ne dokazuje točnu razinu pouzdanosti.

Zadatak 4. Nesigurnost procjene varijabilnosti.

Normalna mjerenja imaju nepoznatu varijancu. Za dva nezavisna uzorka dobiveno je $(n, S^2) = (8, 2.25)$ i $(40, 2.25)$, s varijancom u mm^2 .

- (a) Pomoću `qchisq()` izračunajte 95%-tne intervale:

$$\left[\frac{(n-1)S^2}{\chi_{0.975}^2}, \frac{(n-1)S^2}{\chi_{0.025}^2} \right].$$

- (b) Korjenovanjem granica odredite intervale za standardnu devijaciju i prikažite ih na istom grafu.
- (c) Objasnite zašto jednaka točkovna procjena ne znači jednaku preciznost i zašto je pretpostavka normalnosti ovdje bitna.

Zadatak 5. Što znači 95% pouzdanosti?

Simulirajte 5000 nezavisnih uzoraka veličine 20 iz $N(10, 4)$.

- (a) Za svaki uzorak izračunajte 95%-tni t -interval, tretirajući varijancu kao nepoznatu. Odredite udio intervala koji sadrže pravo očekivanje 10.
- (b) Prvih 50 intervala nacrtajte kao vodoravne dužine, a one koji ne sadrže 10 označite drugom bojom.
- (c) Ponovite za uzorke veličine 80 te usporedite pokrivenost i prosječnu širinu. Objasnite što se u ponavljanjima mijenja: parametar ili interval.