Matematinė statistika (užduotys naudojant R) 2020/ 2021 m.m.

2 užduotis Parametrų ir jų funkcijų pasikliovimo intervalai

1. Duomenys iš pirmos užduoties 2) punkto, t.y. stebėtas a.d., turintis Veibulo skirstinį su parametrais η ir ν, t. y. a. d. T pasiskirstymo funkcija yra

$$F(t;\eta,\nu) = 1 - e^{-\left(\frac{t}{\eta}\right)^{\nu}},$$

gauta didumo $n = 100$ paprastoji atsitiktinė imtis.

Pastaba. Taškiniai įverčiai buvo rasti pirmoje užduotyje.

Atlikite užduotis:

a) Raskite stebėtą Fišerio informacinę matricą.

b) Raskite parametrų pasikliovimo intervalus. Rezultatus (tikros parametrų reikšmės, taškiniai įverčiai ir pasikliovimo intervalai) pateikite duomenų lentelėje (dataframe).

c) Raskite tikimybės, kad stebėtas a.d. įgis reikšmę didesnę už t, pasikliovimo intervalą. Rezultatus (t, tikra tikimybės reikšmė, tikimybės įvertis, pasikliovimo lygmuo, pasikliovimo intervalas) pateikite duomenų lentelėje (dataframe).

d) Raskite intervalinį medianos įvertį. Rezultatus (tikra medianos reikšmė, taškinis įvertis, pasikliovimo lygmuo ir pasikliovimo intervalas) pateikite duomenų lentelėje (dataframe).

e) Raskite p-ojo kvantilio pasikliovimo intervalą. Rezultatus (kvantilio lygmuo p, tikra kvantilio reikšmė, taškinis įvertis, pasikliovimo lygmuo ir pasikliovimo intervalas) pateikite duomenų lentelėje (dataframe).

2. Tarkime, kad laikas iki įvykio T turi ekstremalių reikšmių (minimalių) skirstinį, t.y. pasiskirstymo funkcija

$$F(x;\mu,\sigma) = 1 - \exp\left(-e^{\frac{x-\mu}{\sigma}}\right)$$

Naudodami atvirkštinės transformacijos metodą, sumodeliuokite paprastąją atsitiktinę imtį (Pastaba. Kai iš konteksto aišku, dažnai žodis realizacija praleidžiamas).

Raskite taškinius ir intervalinius parametrų įverčius. Pasirinktame taške t įvertinkite pasiskirstymo funkciją ir jos pasikliovimo intervalą.

3. Duomenys iš pirmos užduoties 3) punkto, t.y. a.d. X, turinčio binominį skirstinį su parametrais k ir p, t. y.

$$\mathbf{P}(X = m \,|\, k, p) = C_k^m p^m (1-p)^{k-m},$$

dydžio $n = 50$ paprastoji atsitiktinė imtis. Raskite tikslų parametro p pasikliovimo intervalą naudodami beta skirstinio kvantilius.

Papildoma medžiaga ir pavyzdžiai

1. Delta metodo taikymas

Stebėta Fišerio informacinė matrica (observed Fisher information matrix; observed Fisher information)

$$\boldsymbol{J}(\boldsymbol{\theta}) = \left(-\frac{\partial^2 \ln L}{\partial\theta_j \partial\theta_{j'}}\right)_{m\times m}; \tag{1}$$

$$\frac{\boldsymbol{J}(\boldsymbol{\theta})}{n} \approx \boldsymbol{i}(\boldsymbol{\theta}) \tag{2}$$

dideliems n (naudojantis Didžiųjų skaičių dėsniu, vienmačio parametro atveju žr. 33 skaidrę Matemat_statistika_02_Taskiniai_ivertiniai) , t.y. $\boldsymbol{J}(\boldsymbol{\theta}) \approx n\boldsymbol{i}(\boldsymbol{\theta}) = \boldsymbol{I}(\boldsymbol{\theta})$, ieškant asimptotinių pasikliovimo intervalų vietoje $\boldsymbol{I}(\boldsymbol{\theta})$, galime naudoti $\boldsymbol{J}(\boldsymbol{\theta})$.

Pažymėkime $\hat{\boldsymbol{\theta}}$ parametro $\boldsymbol{\theta}$ didžiausiojo tikėtinumo įvertinį.

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} Y \sim N\left(\mathbf{0}, \boldsymbol{i}^{-1}(\boldsymbol{\theta})\right). \tag{3}$$

Naudojant delta metodą:

$$\sqrt{n}(\gamma(\hat{\boldsymbol{\theta}}) - \gamma(\boldsymbol{\theta})) \xrightarrow{d} W \sim N\left(0, \sigma_\gamma^2(\boldsymbol{\theta})\right), \tag{4}$$

čia

$$\boldsymbol{G}_\gamma(\boldsymbol{\theta}) = \left(\frac{\partial\gamma(\boldsymbol{\theta})}{\partial\theta_1}, \cdots, \frac{\partial\gamma(\boldsymbol{\theta})}{\partial\theta_s}\right)^T; \quad \sigma_\gamma^2(\boldsymbol{\theta}) = \boldsymbol{G}_\gamma(\boldsymbol{\theta})^T \boldsymbol{i}^{-1}(\boldsymbol{\theta}) \boldsymbol{G}_\gamma(\boldsymbol{\theta})$$

Tada

$$\frac{\sqrt{n}\left(\gamma(\hat{\boldsymbol{\theta}}) - \gamma(\boldsymbol{\theta})\right)}{\sigma_\gamma(\boldsymbol{\theta})} \xrightarrow{d} Z \sim N(0,\, 1). \tag{5}$$

Pavyzdžiui, $\gamma(\boldsymbol{\theta}) = S(x, \boldsymbol{\theta})$; $\gamma(\boldsymbol{\theta}) = x_p(\boldsymbol{\theta})$.

Pastaba. Išgyvenimo funkcija įgyja reikšmes iš intervalo $(0,\, 1]$, aproksimaciją normaliuoju skirstiniu galima pataisyti naudojant transformaciją

$$Q(x, \boldsymbol{\theta}) = \ln\frac{S(x, \boldsymbol{\theta})}{1 - S(x, \boldsymbol{\theta})}.$$

Atsižvelgiant į tai, kad

$$\left(\ln\frac{u}{1-u}\right)' = \frac{1}{u(1-u)}$$

ir naudojant (5) gauname

$$\frac{\sqrt{n}\left(Q(x, \hat{\boldsymbol{\theta}}) - Q(x, \boldsymbol{\theta})\right)}{\sigma_Q(\theta)} \xrightarrow{d} Z \sim N(0,\, 1). \tag{6}$$

$$\sigma_Q(\boldsymbol{\theta}) = \frac{1}{S(x, \boldsymbol{\theta})(1 - S(x, \boldsymbol{\theta}))}\sigma_\gamma(\boldsymbol{\theta}) = \frac{1}{S(x, \boldsymbol{\theta})(1 - S(x, \boldsymbol{\theta}))}\sigma_S(\boldsymbol{\theta}). \tag{7}$$

Asimptotinis išgyvenimo funkcijos pasikliovimo intervalas, esant pasikliovimo lygmeniui $(1-\alpha)$

$$\left(\left(1 + \frac{1 - S(x, \hat{\boldsymbol{\theta}})}{S(x, \hat{\boldsymbol{\theta}})} e^{\sigma_Q(\hat{\boldsymbol{\theta}})\frac{z_{\alpha/2}}{\sqrt{n}}}\right)^{-1},\ \left(1 + \frac{1 - S(x, \hat{\boldsymbol{\theta}})}{S(x, \hat{\boldsymbol{\theta}})} e^{-\sigma_Q(\hat{\boldsymbol{\theta}})\frac{z_{\alpha/2}}{\sqrt{n}}}\right)^{-1}\right)$$

čia z_α yra standartinio normaliojo skirstinio α lygmens kritinė reikšmė.

Pastaba. Jeigu $\gamma(\boldsymbol{\theta}) = x_p(\boldsymbol{\theta})$, tai funkcija γ įgyja teigiamas reikšmes ir aproksimacija yra pataisomas naudojant transformaciją $K(\boldsymbol{\theta}) = \ln \gamma(\boldsymbol{\theta})$.

2. R pavyzdžiai

R funkcijos, skirtos rasti ekstremumo taškus, pateikia Hessian matricą, t.y. antrųjų tikslo funkcijos išvestinių matricą:

```
rez <- nlm(ln_l, theta <- c(1), hessian=TRUE)
rez$estimate
rez$hessian

mle <- maxLik(logLik = ln_l, start = c(lambda = 1))
coef(mle)
hessian(mle)

rez_optim <- optim(init, fn=ln_l2, method="L-BFGS-B",
      lower = c(1.0e-13, 1.0e-13), upper=c(Inf, Inf),
      control=list(fnscale=-1), hessian=TRUE)
rez_optim$par
rez_optim$hessian
```

3. R. Rezultatų apjungimas į lentelę

```
data.frame(v1, v2, v3)
```