Sekvenční intervaly spolehlivosti dané délky

Terézia Hatalová Tomáš Kremla

25. března 2025

Předpoklady

Nechť $\{X_i\}_{i=1}^{\infty}$ je posloupnost iid n. veličin z rozdělení s hustotou $f(x;\theta)$, $\theta \in (\underline{\theta}, \overline{\theta})$. Ať L_n, U_n jsou statistiky s následujícími vlastnostmi:

(i) $L_n = L_n(X_1, \ldots, X_n) < U_n = U_n(X_1, \ldots, X_n), \quad \text{s.j.}, \ \forall n \in \mathbb{N}$

(ii) $\exists A > 0 : \ \sqrt{n}(U_n - L_n) \longrightarrow 2\, u_{1-\alpha/2}\, A, \quad \text{s.j. pro } n \to \infty,$

(iii) $\sqrt{n}(U_n - \theta) - Z_n A - u_{1-\alpha/2}\, A \longrightarrow 0, \quad \text{s.j. pro } n \to \infty,$

kde Z_n je standardizovaný součet iid náhodných veličin s konečným druhým momentem.

Předpoklady

Ať $d > 0$, definujme nyní pravidlo pro ukončení výběru následovně:

$$N(d) = \min\{n;\ U_n - L_n \leq 2d\}.$$

Přidejme předpoklad:

(iv) Náhodné veličiny $\{N(d)d^2;\ d > 0\}$ jsou stejnoměrně integrovatelné.

Za hledaný sekvenční interval spolehlivosti vezmeme:

$$(L_{N(d)}, U_{N(d)}).$$

Věta

Věta

Nechť platí předchozí předpoklady, potom

(a) pro všechna $d > 0$ platí $N(d) < +\infty$ s.j. a $\mathbb{E}\, N(d) < +\infty$. Dále $N(d)$ je nerostoucí funkce d s.j., $\lim_{d\to 0} N(d) = +\infty$ s.j. a

$$\lim_{d\to 0} \mathbb{E}\, N(d) = +\infty$$

(b) $\lim_{d\to 0} N(d)d^2 = u^2_{1-\alpha/2}\, A^2$ s.j.,

$$\lim_{d\to 0} \mathbb{E}\, N(d)d^2 = u^2_{1-\alpha/2}\, A^2$$

(c) $\lim_{d\to 0} P(L_{N(d)} \leq \theta \leq U_{N(d)}) = 1 - \alpha$.

Príklad - Normálne rozdelenie

Nech $\{X_i\}_{i=1}^{\infty}$ je postoupnost iid n. veličín z $N(\mu, \sigma^2)$ pre μ, σ^2 neznáme. Chceme interval spoľahlivosti pre μ dĺžky $2d$, $d > 0$ s pravdepodobnosťou pokrytia $1 - \alpha$. Uvažujme:

$$L_n = \overline{X}_n - \frac{S_n u_{1-\alpha/2}}{\sqrt{n}}, \quad U_n = \overline{X}_n + \frac{S_n u_{1-\alpha/2}}{\sqrt{n}},$$

kde

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \quad S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 + \frac{1}{n}.$$

Príklad - Normálne rozdelenie

Definujeme rozhodovacie pravidlo pre ukončenie výberu ako

$$N(d) = \min\{n;\ \frac{S_n u_{1-\alpha/2}}{\sqrt{n}} \leq d\}.$$

Dostávame požadovaný sekvenčný interval spoľahlivosti tvaru

$$\left(\overline{X}_{N(d)} - \frac{S_{N(d)} u_{1-\alpha/2}}{\sqrt{N(d)}}, \overline{X}_{N(d)} + \frac{S_{N(d)} u_{1-\alpha/2}}{\sqrt{N(d)}} \right).$$

(Ne)optimalita CI při nesekvenčním přístupu

Nechť $\mathbb{X} = (X_i)_{i=1}^n$ jsou náhodné veličiny se sdruženou hustotou

$$\sigma^{-n} f\left(\frac{x_1-\mu}{\sigma}, \dots, \frac{x_n-\mu}{\sigma}\right),$$

kde $\mu \in \mathbb{R}$ a $\sigma > 0$ jsou neznámé parametry. Chceme ukázat, že neexistuje interval spolehlivosti $S(\mathbb{x})$ pro μ splňující:

(a) $\mathbb{P}_{\mu,\sigma}(S(\mathbb{X}) \ni \mu) \geq 1 - \alpha$ pro $\forall \mu \in \mathbb{R}$ a $\sigma > 0$,

(b) $\operatorname{diam}(S(\mathbb{x})) < 2d$ pro $\forall \mathbb{x} \in \mathbb{R}^n$,

pro fixní $d > 0$ a $\alpha \in (0, 1)$.

Steinova procedúra

Nech $\{X_i\}_{i=1}^{\infty}$ je postoupnost iid n. veličín z $N(\mu, \sigma^2)$ pre μ, σ^2 neznáme. Chceme interval spoľahlivosti pre μ dĺžky $2d$, $d > 0$ s pravdepodobnosťou pokrytia $1 - \alpha$.

Prvý krok

(1) vezmime počiatočný výber s rozsahom n_0: $X_1, \dots, X_{n_0}$, $\overline{X}_0 = \frac{1}{n_0} \sum_{i=1}^{n_0} X_i$ a $S_0^2 = \frac{1}{n_0 - 1} \sum_{i=1}^{n_0} (X_i - \overline{X}_0)^2$

(2) $N = \max\{n_0 + 1, \lfloor \frac{S_0^2}{z} \rfloor + 1\}$

(3) vezmime $a_0, a_{n_0+1}, \dots, a_N$ tak, že platí

$$a_0 + \sum_{k=n_0+1}^{N} a_k = 1, \tag{1}$$

$$\frac{a_0^2}{n_0} + \sum_{k=n_0+1}^{N} a_k^2 = \frac{z}{S_0^2}. \tag{2}$$

Steinova procedúra

Druhý krok

(1) Pre jednoznačnosť predpokladajme napríklad
$a_{n_0+1} = \cdots = a_N$,

(2) vezmime náhodný výber s rozsahom $N - n_0$: $X_{n_0+1}, \ldots, X_N$,

(3) uvažujme štatistiku

$$I_N = a_0 \overline{X}_0 + \sum_{k=n_0+1}^{N} a_k X_k. \tag{3}$$

Steinova procedúra

Veta

Nech I_N definovaná ako (3), potom $z^{-1/2}(I_N - \mu)$ má t_{n_0-1} rozdelenie. Ďalej platí

$$\mathbb{P}\left[-z^{1/2}t_{n_0-1}\left(1-\frac{\alpha}{2}\right) < I_N - \mu < z^{1/2}t_{n_0-1}\left(1-\frac{\alpha}{2}\right)\right] = 1-\alpha.$$

Ak z zvolíme tak, aby $-z^{1/2}t_{n_0-1}(1-\frac{\alpha}{2}) = d$, dostávame interval spoľahlivosti pre μ s pravdepodobnosťou pokrytia $1-\alpha$ $(I_N - d, I_N + d)$ požadovanej dĺžky $2d$.