Dvoustupňové metody

Vojtěch Urx

28. března 2023

Hustota

- f - n-rozměrná spojitá hustota
- Náhodné veličiny $Y_1, \dots Y_n$ mají sdruženou hustotu

$$\sigma^{-n} f\left(\frac{y_1-\mu}{\sigma}, \dots, \frac{y_n-\mu}{\sigma}\right) = \sigma^{-n} f\left(\sigma^{-1}(\mathbf{y}-\mu\mathbf{1_n})\right)$$

pro nějaké $\mu \in \mathbb{R}$ a $0 < \sigma < +\infty$

Limitace pevného rozsahu výběru

Tvrzení 1

Nechť $0 < 1 - \alpha < 1$, $0 < d < +\infty$. Pak neexistuje interval spolehlivosti pro μ takový, že pro každé skutečné μ a σ má spolehlivost alespoň $1 - \alpha$ a maximální průměr menší než $2d$.

Tvrzení 2

Nechť $0 < \alpha, \beta < 1$, $\mu_0, \mu_1 \in \mathbb{R}$. Pak neexistuje netriviální test hypotézy $H_0 : \mu = \mu_0$ proti alternativě $H_1 : \mu = \mu_1$, takový, že pro každé skutečné σ má hladinu α a sílu β.

Vlastnost σ roztoucího do nekonečna

Pro $\mu_1 \neq \mu_2$ platí

$$\sup_{E \subset \mathbb{R}^n} \mid \mathbb{P}_{\mu_1}(\mathbb{Y} \in E) - \mathbb{P}_{\mu_2}(\mathbb{Y} \in E) \mid \underset{\sigma \to \infty}{\longrightarrow} 0,$$

neboť

$$\sup_{E \subset \mathbb{R}^n} \mid \int_E \sigma^{-n} f\left(\sigma^{-1}(\mathbf{y} - \mu_1 \mathbf{1_n})\right) d\mathbf{y} - \int_E \sigma^{-n} f\left(\sigma^{-1}(\mathbf{y} - \mu_2 \mathbf{1_n})\right) d\mathbf{y} \mid =$$

$$= \sup_{E \subset \mathbb{R}^n} \mid \int_E f(\mathbf{x}) - f\left(\mathbf{x} + \sigma^{-1}(\mu_1 - \mu_2)\mathbf{1_n}\right) d\mathbf{x} \mid \underset{\sigma \to \infty}{\overset{Sheffe}{=\!\!\longrightarrow}} 0$$

Dvojstupňový Intervalový odhad

- Nyní nechť $Y_1, Y_2, Y_3, \ldots \overset{iid}{\sim} N(\mu, \sigma^2)$
- Chceme najít interval s předepsanou spolehlivostí $1 - \alpha$ a předepsaným průměrem 2d.
- Začneme s počátečním výběrem rozsahu $n_0 \geq 2$ a s jeho pomocí určíme finální rozsah výběru, který potřebujeme.
- $\nu_0 = n_0 - 1$

Věta o Steinově první proceduře

Věta 3

$$\frac{I_N - \mu}{\sqrt{z}} \sim t(\nu_0)$$

Tedy

$$\mathbb{P}\left(-t(\nu_0, 1 - \frac{\alpha}{2})\sqrt{z} < I_N - \mu < t(\nu_0, 1 - \frac{\alpha}{2})\sqrt{z}\right) = 1 - \alpha$$

Jestli na začátku zvolíme z splňující

$$t(\nu_0, 1 - \frac{\alpha}{2})\sqrt{z} = d,$$

pak platí

$$\mathbb{P}\left(I_N - d < \mu < I_N + d\right) = 1 - \alpha$$

Testování hypotéz

- Nechť $\mu_0 \in \mathbb{R}$.
- $H_0 : \mu = \mu_0, H_1 : \mu > \mu_0$
- Testová statistika je

$$t = \frac{I_N - \mu_0}{\sqrt{z}} \overset{H_0}{\sim} t(\nu_0)$$

- Obecně

$$t = \frac{I_N - \mu}{\sqrt{z}} + \frac{\mu - \mu_0}{\sqrt{z}} \sim t(\nu_0) + \frac{\mu - \mu_0}{\sqrt{z}}$$

Síla testu

- Pro testování H_0 proti H_1 s hladinou α, zamítneme H_0 právě, když $t > t(\nu_0, 1-\alpha)$. Síla testu je

$$\mathbb{P}(t(\nu_0) > t(\nu_0, 1-\alpha) - \frac{\mu - \mu_0}{\sqrt{z}})$$

  což nezávisí na σ^2 a je ryze roztoucí v μ.
- Sílu testu mimo okolí μ_0 můžeme libovolně zdola omezit volbou z
  - Pro předepsané $\delta > 0$, $0 < \beta < 1$ zvolíme z takové, že

$$\mathbb{P}(t(\nu_0) > t(\nu_0, 1-\alpha) - \frac{\delta}{\sqrt{z}}) = \beta.$$

  - Pak náš test má sílu β pro každé $\mu \geq \mu_0 + \delta$.
- Tento test zároveň funguje jako test $H_0' : \mu \leq \mu_0$ proti $H_1 : \mu > \mu_0$.

Chybovost odhadu

- Nechť $W(|a-\mu|)$ je neklesající funkce na $[0,\infty)$ popisující chybu způsobenou odhadnutím μ pomocí a, splňující $W(0)=0, W(\infty)\leq\infty$ a $\mathbb{E}W(c|t(\nu_0)|)$ existuje pro všechna $c>0$.
- Podle věty 3 očekávaná chyba I_N je

$$\mathbb{E}W(|I_N-\mu|)=\mathbb{E}W(\sqrt{z}|t(\nu_0)|)$$

- Pro každé $0<\omega<W(\infty)$ lze najít z takové, že

$$\mathbb{E}W(|I_N-\mu|)=\omega.$$