
Bootstrap pro distribuční funkce

Barbora Šáchová, Alena Ulmanová

9. března 2026

- $X_1, \dots, X_n \sim F$
- $\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n I[X_i \geq x]$
- $X_1^*, \dots, X_n^* \sim \hat{F}$
- $R_n = R_n(X_1, \dots, X_n), R_n^* = R_n(X_1^*, \dots, X_n^*)$
- $H_n(x) = H_{F,n}(x) = P(R_n \leq x)$
- $H_n^*(x) = H_{\hat{F},n}(x) = P^*(R_n^* \leq x)$

Bootstrap

Pro $b = 1, \dots, B$:

- generujeme X_1^*, \dots, X_n^* nezávisle s rozdělením \hat{F}
- počítáme $R_{n,b}^* = R_n(X_1^*, \dots, X_n^*)$

Odhadneme $H_n^*(x)$ pomocí

$$H_{n,B}^*(x) = \frac{1}{B} \sum_{b=1}^B I[R_{n,b}^* \leq x]$$

- $H_{n,B}^*(x) \xrightarrow{s.j.} H_n^*(x)$ pro $B \rightarrow \infty$, podmíněně na původním výběru

- Nechť $R \sim H$
- Konvergence v distribuci $R_n \xrightarrow{D} R$ znamená, že $\lim_{n \rightarrow \infty} H_n(x) = H(x)$ ve všech bodech spojitosti H .
- Budeme značit $R_n^* \xrightarrow{D^*} R$ skoro jistě (v pravděpodobnosti), pokud $H_n^*(x) \xrightarrow{s.j./P} H(x)$ ve všech bodech spojitosti H .

- $X_1, \dots, X_n \sim F, \mu = EX_1, \sigma^2 = \text{Var}X_1 < \infty$

- CLV : $R_n = \sqrt{n}(\bar{X}_n - \mu) \xrightarrow{D} R \sim N(0, \sigma^2)$

-

$$\varphi_{R_n}(z) = \left(1 - \frac{\sigma^2 z^2}{2n} + \lambda_n(z) \right)^n$$

- $\lambda_n(z) = o\left(\frac{1}{n}\right)$

- Tedy $\varphi_{R_n}(z) \rightarrow \varphi_R(z) = e^{-\frac{\sigma^2 z^2}{2}}$

- Lévy-Cramér : $H_n(x) \rightarrow H(x) = \Phi\left(\frac{x}{\sigma}\right)$ pro všechna $x \in \mathbb{R}$.

Věta o konvergenci distribučních funkcí

Nechť jsou X_1, X_2, \dots nezávislé, stejně rozdělené náhodné veličiny s rozptylem $\sigma^2 > 0$. Pro náhodný výběr X_1, \dots, X_n mějme neparametrické bootstrap pozorování $X_1^*, \dots, X_n^* \sim \hat{F}$.
Nechť

$$R_n^* = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n X_i^* - \bar{X}_n \right).$$

Definujeme bootstrap charakteristickou funkci

$$\varphi_{R_n^*}(z) = E^* \left[\exp \left(iz \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j^* - \bar{X}_n) \right) \right],$$

poté pro každé $z \in \mathbb{R}$, $\varphi_{R_n^*}(z)$ konverguje s.j. k charakteristické funkci $N(0, \sigma^2)$ -distribuce

$$\varphi_R(z) = e^{-\frac{\sigma^2 z^2}{2}}.$$

Věta o podmíněné slabé konvergenci

$R_n^* \xrightarrow{D^*} R$ skoro jistě (nebo v pravděpodobnosti) právě tehdy, když pro všechny (stejně) spojitě, omezené funkce $h: \mathbb{R} \rightarrow \mathbb{R}$ platí

$$E^* [h(R_n^*)] \longrightarrow E [h(R)], \quad n \rightarrow \infty$$

skoro jistě (nebo v pravděpodobnosti).

Věta o podmíněné spojitě transformaci

Nechť $R_n^* \xrightarrow{D^*} R$ skoro jistě (nebo v pravděpodobnosti) a $g : \mathbb{R} \rightarrow \mathbb{R}$ je spojitá. Poté

$$g(R_n^*) \xrightarrow{D^*} g(R)$$

skoro jistě (nebo v pravděpodobnosti).

Podmíněná verze Sluckého věty

Předpokládejme, že rozdělení náhodných veličin Y_n^* a Z_n^* závisí na X_1, \dots, X_n . Nechť $Z_n^* \xrightarrow{D^*} Z$ skoro jistě (nebo v pravděpodobnosti). Nechť pro nějakou konstantu c a pro každé $\varepsilon > 0$ platí

$$P^*(|Y_n^* - c| > \varepsilon) \xrightarrow{s.j./p} 0.$$

Pak

$$Y_n^* + Z_n^* \xrightarrow{D^*} c + Z, \quad Y_n^* Z_n^* \xrightarrow{D^*} cZ$$

skoro jistě (nebo v pravděpodobnosti).

- **cíl:** $H_n^*(x) \xrightarrow{P} H(x)$ v bodech spojitosti
- použijeme analogický postup jako při důkazu klasické delta věty
- mějme $R_n = \sqrt{n}(g(\bar{X}_n) - g(\mu)) = \nabla g(\mu)\sqrt{n}(\bar{X}_n - \mu) + o_P(1)$
- předpokládáme, že $\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \xrightarrow{D^*} N(0, \Sigma)$ skoro jistě
- $R_n^* = \sqrt{n}(g(\bar{X}_n^*) - g(\bar{X}_n))$ je bootstrap verze R_n
- ukážeme, že $R_n^* = \widetilde{R}_n^* + r_n^*$, kde $P^*(\widetilde{R}_n^* \leq x) \xrightarrow{P} H(x)$ a $r_n^* \xrightarrow{P} 0$.

Delta věta pro bootstrap

Nechť $(\bar{X}_n - \mu)/c_n$ konverguje v distribuci k R , kde $c_n \rightarrow 0$ je deterministická posloupnost. Nechť je dále

$$R_n^* = \frac{g(\bar{X}_n^*) - g(\bar{X}_n)}{c_n}$$

bootstrap verze $R_n = \sqrt{n}(g(\bar{X}_n) - g(\mu))$ a $g: \mathbb{R}^k \rightarrow \mathbb{R}^\ell$ měřitelné zobrazení, spojitě diferencovatelné na okolí μ .

Nechť dále $(\bar{X}_n^* - \bar{X}_n)/c_n \xrightarrow{D^*} R$ v pravděpodobnosti (nebo skoro jistě). Poté

$$\frac{g(\bar{X}_n^*) - g(\bar{X}_n)}{c_n} \xrightarrow{D^*} g'(\mu)R \quad \text{v pasti.}$$

Pokud navíc \bar{X}_n konverguje skoro jistě k μ , pak tato konvergence platí skoro jistě.

- X_1, \dots, X_n iid,
- Necht existují čtvrté momenty.
- Odhadneme $\sigma^2 > 0$ výběrovým rozptylem

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = g(\bar{X}_n, \bar{X}_n^2),$$

kde $g(x, y) = y - x^2$.

- Necht $\mu = (EX_1, EX_1^2)$.
- Pak z Delta věty : $\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2) \xrightarrow{D} R \sim N(0, g'(\mu)\Sigma g'(\mu)^T)$
- Bootstrap verze : $\sqrt{n}(\hat{\sigma}_n^{*2} - \hat{\sigma}_n^2) \xrightarrow{D^*} R \sim N(0, g'(\mu)\Sigma g'(\mu)^T)$