Empīriskā ticamības funkcijas metode vienas un divu izlašu gadījumā

Sandra Vucāne

Latvijas Universitāte

2010.gada 29.septembris

EL metode vienai izlasei

- $X_1, \ldots, X_n$ i.i.d. gadījuma lielumi ar nezināmu sadalījuma funkciju F un parametru θ;

- F un θ: $g(X, \theta)$, $E\{g(X, \theta)\} = 0$
  1. vidējai vērtībai: $g(X_i, \theta) = X_i - \mu$;
  2. kvantilēm $\theta_q = F^{-1}(q)$:
     - $g(X_i, \theta_q) = I_{\{X \leq \theta_q\}} - q$,
     - $\hat{F}_{h,p}(\theta_q) = \sum_{i=1}^{n} p_i G_h(\theta_q - X_i)$.
       $g(X_i, \theta_q) = G_h(\theta_q - X_i) - q$, kur $G_h(x) = G(x/h)$ un h ir joslas platums.

- Funkciju $L(F) = \prod_{i=1}^{n} p_i$ maksimizē pie ierobežojumiem:

$$p_i \geq 0, \ \sum_i p_i = 1, \ \sum_i p_i g(X_i, \theta) = 0.$$

EL metode vienai izlasei

- Empīriskā ticamības funkcija parametram θ

$$L(\theta) = \prod_{i=1}^{n} p_i = \prod_{i=1}^{n} n^{-1}\{1 + \lambda(\theta) g(X_i, \theta)\}^{-1}$$

- Empīriskās ticamības attiecība parametram θ

$$R(\theta) = \frac{L(\theta)}{L(\hat{\theta})} = \prod_{i=1}^{n}\{1 + \lambda(\theta) g(X_i, \theta)\}^{-1},$$

Logaritmiskā empīriskās ticamības attiecība

$$W(\theta_0) = -2\log R(\theta_0) \to_d \chi^2_1,$$

kad $n \to \infty$ un $H_0 : \theta = \theta_0$ ir spēkā.

EL metode divām izlasēm

- $X_1, \ldots, X_{n_1}$ ir i.i.d. $\sim F_1$ un $Y_1, \ldots, Y_{n_2}$ ir i.i.d. $\sim F_2$;
- θ_0 un Δ_0 $\left(\Delta_0 = \theta_1 - \theta_0\right)$:

  $w_1(X_i, \theta_0, \Delta_0, t)$ un $w_2(Y_j, \theta_0, \Delta_0, t)$, kurām ir spēkā

  $\mathbb{E}_{F_1} w_1(X, \theta_0, \Delta_0, t) = 0$, $\mathbb{E}_{F_2} w_2(Y, \theta_0, \Delta_0, t) = 0$.
- Piemēri:

  1. <u>Vidējo vērtību starpība</u>:

     $\theta_0 = \int x dF_1(x)$ un $\Delta_0 = \int y dF_2(y) - \int x dF_1(x)$.
     $w_1 = X - \theta_0$, $w_2 = Y - \theta_0 - \Delta_0$.

  2. <u>Sadalījuma funkciju starpība</u>:

     $\theta_0 = F_1(t)$ un $\Delta_0 = F_2(t) - F_1(t)$.
     $w_1 = I_{\{X \leq t\}} - \theta_0$, $w_2 = I_{\{Y \leq t\}} - \theta_0 - \Delta_0$.

EL metode divām izlasēm

- Empīriskās ticamības attiecība

$$R(\Delta, \theta) = \sup_{\theta, p, q} \prod_{i=1}^{n_1} (n_1 p_i) \prod_{j=1}^{n_2} (n_2 q_j), \text{ kur}$$

$$p_i = \{n_1(1 + \lambda_1(\theta) w_1(X_i, \theta, \Delta, t))\}^{-1}, \quad i = 1, \ldots, n_1.$$
$$q_j = \{n_2(1 + \lambda_2(\theta) w_2(Y_j, \theta, \Delta, t))\}^{-1}, \quad j = 1, \ldots, n_2,$$

- $\lambda_1(\theta)$ un $\lambda_2(\theta)$:

$$\sum_{i=1}^{n_1} \frac{w_1(X_i, \theta, \Delta, t)}{1 + \lambda_1(\theta) w_1(X_i, \theta, \Delta, t)} = 0$$
$$\sum_{j=1}^{n_2} \frac{w_2(Y_j, \theta, \Delta, t)}{1 + \lambda_2(\theta) w_2(Y_j, \theta, \Delta, t)} = 0.$$

EL metode divām izlasēm

- Empīriskā logaritmiskā ticamības attiecība

$$W(\Delta,\theta) = -2\log R(\Delta,\theta) = 2\sum_{i=1}^{n_1}\log(1+\lambda_1(\theta)w_1(X_i,\theta,\Delta,t)) + 2\sum_{j=1}^{n_2}\log(1+\lambda_2(\theta)w_2(Y_j,\theta,\Delta,t)).$$

- Parametra θ novērtējumu $\hat{\theta} = \hat{\theta}(\Delta)$, kas maksimizē $R(\Delta,\theta)$ fiksētam parametram Δ, var iegūt no vienādojuma $\partial W(\Delta,\theta)/\partial\theta$

Izpildoties zināmiem nosacījumiem

$$-2\log R(\Delta_0,\hat{\theta}) \to_d \chi_1^2,$$

3. Kvantiļu funkciju starpība:
   $\theta_0 = F_1^{-1}(t)$ un $\Delta_0 = F_2^{-1}(t) - F_1^{-1}(t)$.
   $w_1 = I_{\{X \le \theta_0\}} - t$, $w_2 = I_{\{Y \le \theta_0 + \Delta_0\}} - t$.

4. P-P grafiks:
   $\theta_0 = F_2^{-1}(t)$ un $\Delta_0 = F_1(F_2^{-1}(t))$, kas ir funkciju F_1 un F_2 P-P grafiks.
   $w_1 = I_{\{X \le \theta_0\}} - \Delta_0$, $w_2 = I_{\{Y \le \theta_0\}} - t$.

5. ROC līkne:
   $\theta_0 = F_2^{-1}(1-t)$ un $\Delta_0 = 1 - F_1(F_2^{-1}(1-t))$.
   $w_1 = I_{\{X \le \theta_0\}} + \Delta_0 - 1$, $w_2 = I_{\{Y \le \theta_0\}} + t - 1$.

6. Q-Q grafiks:
   $\theta_0 = F_2(t)$ un $\Delta_0 = F_1^{-1}(F_2(t))$.
   $w_1 = I_{\{X \le \Delta_0\}} - \theta_0$, $w_2 = I_{\{Y \le t\}} - \theta_0$.

7. Strukturālo attiecību modeļi:
   $F_1(t) = \phi_2^-(F_2(\phi_1^-(t,h)),h), \quad t \in \mathbb{R}, \quad h \in \mathcal{H} \subseteq \mathbb{R}^l$

   - $\Delta := \Delta(t) = F_1(\phi_1(F_2^{-1}(\phi_2(t,h)),h))$ vispārina P-P grafiku.
   - $w_1 = I_{\{X \le \theta_0\}} - \Delta_0$, $w_2 = I_{\{Y \le \phi_1^-(\theta_0,h_0)\}} - \phi_2(t,h_0)$, kur $\theta_0 = \phi_1(F_2^{-1}(\phi_2(t,h_0)),h_0)$.

Gludinātā EL 2 izlašu gadījumā

- $\hat{F}_{b_1,p}(x) = \sum_{i=1}^{n_1} p_i H_{b_1}(x - X_i)$ un $\hat{F}_{b_2,q}(y) = \sum_{j=1}^{n_2} q_j H_{b_2}(y - Y_j)$, kur $H_{b_j}(t) = H_j(t/b_j)$, $b_1 = b_1(n_1)$ un $b_2 = b_2(n_2)$,
- $R^{(sm)}(\Delta, \theta) = \sup_{p,q} \prod_{i=1}^{n_1}(n_1 p_i) \prod_{j=1}^{n_2}(n_2 q_j)$

Piemērs	$w_1(X_i, \theta, \Delta, t)$	$w_2(Y_i, \theta, \Delta, t)$
2	$H_{b_1}(t - X_i) - \theta_0$	$H_{b_2}(t - Y_j) - \theta_0 - \Delta_0$
3	$H_{b_1}(\theta_0 - X_i) - t$	$H_{b_2}(\theta_0 + \Delta_0 - Y_j) - t$
4	$H_{b_1}(\theta_0 - X_i) - \Delta_0$	$H_{b_2}(\theta_0 - Y_j) - t$
5	$H_{b_1}(\theta_0 - X_i) - 1 + \Delta_0$	$H_{b_2}(\theta_0 - Y_j) - 1 + t$
6	$H_{b_1}(\Delta_0 - X_i) - \theta_0$	$H_{b_2}(t - Y_j) - \theta_0$
7	$H_{b_1}(\theta_0 - X_i) - \Delta_0$	$H_{b_2}(\phi_1^-(\theta_0, h_0) - Y_j) - \phi_2(t, h_0)$

$$w_1(X_i, \theta, \Delta, t) = H_{b_1}(\xi_1(\theta) - X_i) - \xi_2(\theta),$$
$$w_2(Y_j, \theta, \Delta, t) = H_{b_2}(\psi_1(\theta) - X_i) - \psi_2(\theta).$$

- Δ: $w_1(X_i, \theta_0, \Delta, t, h)$ un $w_2(Y_j, \theta_0, \Delta, t, h)$, kur h ir traucējošais parametrs.

- Mallows attālums: $\hat{h} = \arg\min_{h \in \mathcal{H}_0} \{\frac{1}{b-a} \int_a^b (F_{1n_1}^{-1}(u) - \phi_1(F_{2n_2}^{-1}(\phi_2(u, h)), h))^2 du\}$, kur $0 \leq a < b \leq 1$.

- P-P grafiks
  Lokācijas modeļa gadījumā: $F_1(t)$, $F_2(t - \hat{h})$.

Paldies par uzmanību!