Plug-in empirical likelihood

Leonora Pahirko, Jānis Valeinis

Faculty of Physics and Mathematics, University of Latvia

May 28, 2011

Empirical likelihood function

Nonparametric likelihood

$X_1,\ldots,X_n \sim F_0$ iid, the nonparametric likelihood of the CDF F is

$$L(F) = \prod_{i=1}^{n} p_i = \prod_{i=1}^{n} P(x = X_i).$$

EL ratio function

$$R(F) = \frac{L(F)}{L(\widehat{F}_n)} = \prod_{i=1}^{n} np_i, \quad \text{where } \widehat{F}_n \text{ is ECDF}$$

Profile EL ratio function

$$EL_n(X_i, \theta) = \max\left\{\prod_{i=1}^{n} np_i \mid p_i > 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i m(X_i, \theta) = 0\right\}$$

Plug-in empirical likelihood

(A0) $P(EL_n(\theta_0, \hat{h}) = 0) \to 0$

(A1) $\sum_{i=1}^{n} m(X_i, \theta_0, \hat{h})/\sqrt{n} \to_d U$

(A2) $\sum_{i=1}^{n} m^2(X_i, \theta_0, \hat{h})/n \to_{pr} \sigma^2$

(A3) $\max_{1 \le i \le n} |m(X_i, \theta_0, \hat{h})\sqrt{n}| \to_{pr} 0$

Theorem by Hjort, McKeague, Van Keilegom

If (A0) - (A3) hold, then $-2 \ln EL_n(\theta_0, \widehat{h}) \to_d U^2/\sigma^2$.

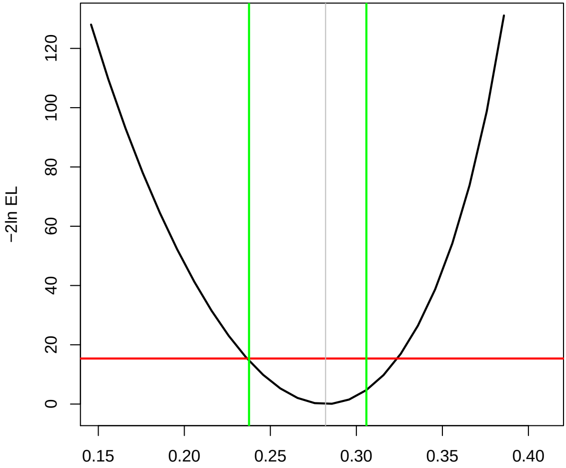
Applications

Example 1

- Asymptotic variance of Hodges-Lehmann estimator of location
- $X_1, \ldots, X_n$ iid from an unknown density f_0
- Parameter of interest $\theta_0 = \int f_0^2 dx$
- Estimating equation $m(X, \theta, f) = f(X) - \theta$, where

$$\widehat{f}(x) = n^{-1} \sum_{i=1}^{n} k\left(\frac{X_i - x}{b}\right) \frac{1}{b}$$

Confidence interval $\theta_0 = \int f_0^2 dx$



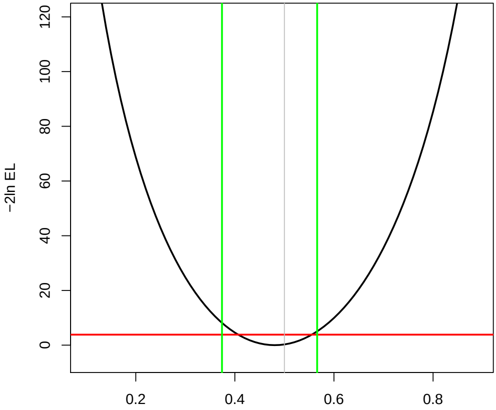
Example 2

- Nonparametric regression model $Y = \mu(X) + \varepsilon$
- F_ε and $\mu(\cdot)$ unknown
- Parameter of interest $\theta_0 = F_\varepsilon(z)$
- Estimating equation $m(X, Y, \theta, \mu) = I_{\{Y-\mu(X)\leq z\}} - \theta$
- Nuisance parameter

$$\widehat{\mu}(x) = \sum_{i=1}^{n} W_{n,i}(x; b_n)Y_i,$$

where $W_{n,i}(x; b_n) = k_{b,x}(X_i)/\sum_{j=1}^{n} k_{b,x}(X_j)$

Confidence interval $\theta_0 = F_\varepsilon(0)$



Two sample plug-in empirical likelihood

Difference between integrals of squared densities

- Parameter of interest $\Delta = \int f_2^2(y)dy - \int f_1^2(x)dx$
- $\theta_0 = \int f_1^2(x)dx$

Estimating equations

$$m_1(X, \theta_0, \Delta) = f_1(X) - \theta_0$$

$$m_2(Y, \theta, \Delta) = f_2(Y) - \theta_0 - \Delta$$

Two sample plug-in empirical likelihood

Difference of smoothed Huber estimators

- Two independent samples X and Y with distribution functions F_1 and F_2
- Parameter of interest $\Delta = \theta_1 - \theta_0$

Estimating equations

$$m_1(X, \theta_0, \Delta) = \tilde{\psi}\left(\frac{X - \theta_0}{\hat{\sigma}_1}\right)$$

$$m_2(Y, \theta_0, \Delta) = \tilde{\psi}\left(\frac{Y - \Delta + \theta_0}{\hat{\sigma}_2}\right)$$

$\hat{\sigma}_1$ and $\hat{\sigma}_2$ are scale estimators, $\tilde{\psi}$ smoothed Huber estimator

See

Two sample plug-in empirical likelihood

Structural relationship models

- Location-scale and Lehmann's alternative models
- Two independent samples X and Y with distribution functions F_1 and F_2
- Parameter of interest $\Delta = F_1(\phi_1(F_2^{-1}(\phi_2(t,h)),h))$
- $\theta_0 = \phi_1(F_2^{-1}(\phi_2(t,\hat{h})),\hat{h})$

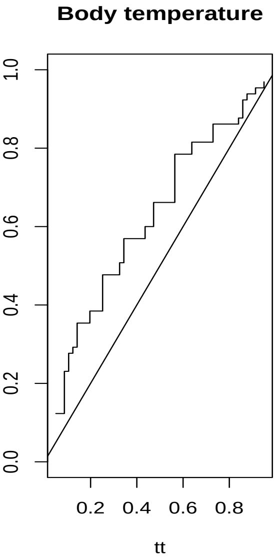
Estimating equations

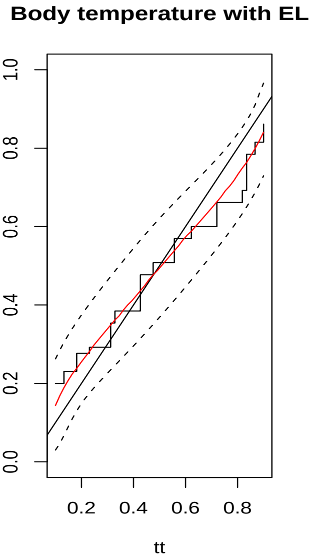
$$m_1(X,\theta_0,\Delta,t,\hat{h}) = I_{\{X\leq\theta_0\}} - \Delta$$

$$m_2(Y,\theta_0,\Delta,t,\hat{h}) = I_{\{Y\leq\phi_1^{-}(\theta_0,\hat{h})\}} - \phi_2(t,\hat{h})$$

See Valeinis, 2007.

Example of location-scale model





Thank you!