Estimación Robusta en Modelos de Indice Simple con Errores Asimétricos

Ana M. Bianco

Instituto de Cálculo, FCEyN, UBA y CONICET

En colaboración con
C. Agostinelli y G. Boente

Single Index Model

Observe independent r. v. $(y, \mathbf{x}^{\mathsf{t}})$

$y \in \mathbb{R}$: response

$\mathbf{x} \in \mathbb{R}^q$: single index variables

$$y = \eta_0(\boldsymbol{\beta}_0^{\mathsf{t}}\mathbf{x}) + \epsilon$$

$\swarrow$ $\searrow$

Unknown $\eta_0 : \mathbb{R} \to \mathbb{R}$ **and** $\boldsymbol{\beta}_0 \in \mathbb{R}^q$

Single Index Model

Observe independent r. v. $(y, \mathbf{x}^{\mathsf{t}})$

$y \in \mathbb{R}$: response

$\mathbf{x} \in \mathbb{R}^q$: single index variables

$$y = \eta_0(\boldsymbol{\beta}_0^{\mathsf{t}}\mathbf{x}) + \epsilon$$

$$\swarrow \qquad\qquad \searrow$$

Unknown $\eta_0 : \mathbb{R} \to \mathbb{R}$ **and** $\boldsymbol{\beta}_0 \in \mathbb{R}^q$

- Dimension reduction technique: reduces the dimensionality by means of the single index $\boldsymbol{\beta}_0^{\mathsf{t}}\mathbf{x}$
- If $\boldsymbol{\beta}_0$ is estimated efficiently $\Rightarrow \boldsymbol{\beta}_0^{\mathsf{t}}\mathbf{x} \in \mathbb{R}$ carrier to estimate η_0 nonparametrically

Single Index Model: Identifiability

For the sake of identifiability

$$y = \eta_0(\boldsymbol{\beta}_0^{\mathsf{t}}\mathbf{x}) + \epsilon$$

we assume

$$\|\boldsymbol{\beta}_0\| = 1 \text{ and w.l.g. } \beta_{0q} > 0$$

Errors Distribution

Classical setting

Usually assumed: $\mathbb{E}(\epsilon) = 0$ and $\mathbb{E}(\epsilon^2) < \infty$

Our setting

ϵ has density

$$g(\epsilon, \alpha) = Q(\alpha) \exp^{\alpha t(\epsilon)} \tag{1}$$

$\alpha > 0$: nuisance parameter

$t(\epsilon)$: continuous function with unique mode at ϵ_0.

Errors Distribution

- Symmetric Errors

$\epsilon \sim F(\cdot/\sigma) \quad F$: symmetric

σ: scale parameter $(\alpha = \sigma)$

Errors Distribution

- Symmetric Errors

$\epsilon \sim F(\cdot/\sigma)$ $\quad F$: symmetric

σ: scale parameter ($\alpha = \sigma$)

- Asymmetric Errors

log–Gamma Regression Model

$$g(\epsilon, \alpha) = \frac{\alpha^{\alpha}}{\Gamma(\alpha)} \exp^{\alpha(\epsilon - \exp(\epsilon))}$$

Asymmetric and unimodal at $\epsilon_0 = 0$

α: shape parameter

Estimation of σ and α is needed to calibrate robust estimators

Most Popular: log–Gamma Regression Model

Parametric Case

$$\left.\begin{array}{r} z \in \mathbb{R}: \text{ response} \\ \\ \mathbf{x} \in \mathbb{R}^q \text{covariates} \end{array}\right\} \qquad \begin{array}{l} z|_{\mathbf{x}} \sim \Gamma(\alpha, \mu(\mathbf{x})) \\ \\ \log \mu(\mathbf{x}) = \mathbf{x}^{\mathsf{t}} \boldsymbol{\beta}_0 \end{array}$$

Hence

$$u = z/\mu(\mathbf{x}) \sim \Gamma(\alpha, 1) \Rightarrow \begin{cases} y = \log(z) \\ \\ \epsilon = \log(u) \end{cases}$$

$\Rightarrow$

$$y = \mathbf{x}^{\mathsf{t}} \boldsymbol{\beta}_0 + \epsilon$$

where $\epsilon \sim \log(\Gamma(\alpha, 1))$, i.e. $g(\epsilon, \alpha) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \exp^{\alpha(\epsilon - \exp(\epsilon))}$

Most Popular: log–Gamma Regression Model

Parametric Case

- Classical Estimators:

Deviance Components: $d(y, a) = \exp(y - a) - (y - a) - 1$

Most Popular: log–Gamma Regression Model

Parametric Case

- Classical Estimators:

Deviance Components: $d(y,a) = \exp(y-a) - (y-a) - 1$

- $M-$Estimators:

$$\phi(y,u,c) = \rho_{\mathrm{T}}\left(\frac{\sqrt{d(y,u)}}{c}\right)$$

where c tuning constant (depending on α)

We borrow some of these ideas to compute a *robust profile method* that involves smoothing techniques.

Given $\boldsymbol{\beta}$, define kernel weights $W_{\boldsymbol{\beta},\,i}(u)$

$$W_{\boldsymbol{\beta},\,i}(u,h) = K\left(\frac{\boldsymbol{\beta}^{\mathsf{t}}\mathbf{x}_i - u}{h}\right)\left\{\sum_{j=1}^{n} K\left(\frac{\boldsymbol{\beta}^{\mathsf{t}}\mathbf{x}_j - u}{h}\right)\right\}^{-1} .$$

K: kernel

h: bandwidth

$W_{\boldsymbol{\beta},\,i}(u,h)$ measures the *closeness* between u and the projection of $\mathbf{x}_i$ in the direction of $\boldsymbol{\beta}$

Given an initial value $\widehat{c}$

- **Step 1**: For each fixed $\boldsymbol{\beta}$, with $\|\boldsymbol{\beta}\| = 1$, let

$$\widehat{\eta}_{\boldsymbol{\beta}}(u) \quad = \quad \underset{a\in\mathbb{R}}{\text{argmin}}\, R_n(a) = \underset{a\in\mathbb{R}}{\text{argmin}} \sum_{i=1}^{n} \rho_{\mathrm{T}}\left(\frac{\sqrt{d(y_i, a)}}{\widehat{c}}\right) W_{\boldsymbol{\beta},\, i}(u, h)$$

Given an initial value $\widehat{c}$

- **Step 1**: For each fixed $\boldsymbol{\beta}$, with $\|\boldsymbol{\beta}\| = 1$, let

$$\widehat{\eta}_{\boldsymbol{\beta}}(u) = \underset{a\in\mathbb{R}}{\operatorname{argmin}}\, R_n(a) = \underset{a\in\mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^{n} \rho_{\mathrm{T}}\left(\frac{\sqrt{d(y_i, a)}}{\widehat{c}}\right) W_{\boldsymbol{\beta},\, i}(u, h)$$

- **Step 2**: Define the estimators $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}_0$ as a minimum of $G_n(\boldsymbol{\beta})$ where

$$G_n(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n} \rho_{\mathrm{T}}\left(\frac{\sqrt{d\left(y_i, \widehat{\eta}_{\boldsymbol{\beta}}\left(\boldsymbol{\beta}^{\mathsf{t}}\mathbf{x}_i\right)\right)}}{\widehat{c}}\right) \tau_{n,\boldsymbol{\beta}}(\mathbf{x}_i)$$

τ: trimming function

Given an initial value $\widehat{c}$

- **Step 1**: For each fixed $\boldsymbol{\beta}$, with $\|\boldsymbol{\beta}\| = 1$, let

$$\widehat{\eta}_{\boldsymbol{\beta}}(u) \quad = \quad \underset{a \in \mathbb{R}}{\operatorname{argmin}}\, R_n(a) = \underset{a \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^{n} \rho_{\mathrm{T}}\left(\frac{\sqrt{d(y_i, a)}}{\widehat{c}}\right) W_{\boldsymbol{\beta},\, i}(u, h)$$

- **Step 2**: Define the estimators $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}_0$ as a minimum of $G_n(\boldsymbol{\beta})$ where

$$G_n(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^{n} \rho_{\mathrm{T}}\left(\frac{\sqrt{d\left(y_i, \widehat{\eta}_{\boldsymbol{\beta}}\left(\boldsymbol{\beta}^{\mathsf{t}}\mathbf{x}_i\right)\right)}}{\widehat{c}}\right) \tau_{n,\boldsymbol{\beta}}(\mathbf{x}_i)$$

  τ: trimming function

- **Step 3**: Define the final estimator $\widehat{\eta}$ of η_0 as $\widehat{\eta}(u) = \widehat{a}(u)$ with

$$(\widehat{a}(u), \widehat{b}(u)) \quad = \quad \underset{(a,b) \in \mathbb{R}^2}{\operatorname{argmin}} \sum_{i=1}^{n} W_{\widehat{\boldsymbol{\beta}},\, i}(u, h) \rho_{\mathrm{T}}\left(\frac{\sqrt{d(y_i, a + b(\widehat{\boldsymbol{\beta}}^{\mathsf{t}}\mathbf{x}_j - u))}}{\widehat{c}}\right)$$

Generalizing the proposal

$$(y_i, \mathbf{x}_i), i = 1, \ldots, n, \text{independent r. v.} : y_i = \eta_0(\boldsymbol{\beta}_0' \mathbf{x}_i) + \epsilon_i$$

$$R_n(\boldsymbol{\beta}, a, u, c) = \sum_{i=1}^{n} W_{\boldsymbol{\beta}, i}(u, h) \underbrace{\rho_{\mathrm{T}} \left(\frac{\sqrt{d(y_i, a)}}{c} \right)}_{\phi(y_i, a, c)}$$

$$G_n(\boldsymbol{\beta}, v, c) = \frac{1}{n} \sum_{i=1}^{n} \underbrace{\rho_{\mathrm{T}} \left(\frac{\sqrt{d(y_i, v(\boldsymbol{\beta}^{\mathsf{t}} \mathbf{x}_i))}}{c} \right)}_{\phi(y_i, v, c)} \tau_{n, \boldsymbol{\beta}}(\mathbf{x}_i)$$

General Case

Given an initial value $\widehat{\alpha}$

- **Step 1**: For each fixed $\boldsymbol{\beta}$, with $\|\boldsymbol{\beta}\| = 1$, let

$$\widehat{\eta}_{\boldsymbol{\beta}}(u) \quad = \quad \operatorname*{argmin}_{a \in \mathbb{R}} R_n(\boldsymbol{\beta}, a, u, \widehat{\alpha})$$

- **Step 2**: Define the estimators $\widehat{\boldsymbol{\beta}}$:

$$\widehat{\boldsymbol{\beta}} = \operatorname*{argmin}_{\|\boldsymbol{\beta}\|=1} G_n(\boldsymbol{\beta}, \widehat{\eta}_{\boldsymbol{\beta}}, \widehat{\alpha})$$

- **Step 3**: Define the final estimator $\widehat{\eta}$ of η_0 as $\widehat{\eta}(u) = \widehat{a}(u)$ with

$$(\widehat{a}(u), \widehat{b}(u)) \quad = \quad \operatorname*{argmin}_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^{n} W_{\widehat{\boldsymbol{\beta}},\, i}(u,h) \phi \left(y_i, a + b(\widehat{\boldsymbol{\beta}}^{\mathsf{t}} \mathbf{x}_j - u), \widehat{\alpha} \right)$$

Initial Estimators

- Suitable initial estimators should be developed in each case.

- In the case of the log–Gamma distribution, we propose a profile procedure that combines smoothing techniques with the $MM-$estimators of B.,García Ben & Yohai (2005).

Asymptotic Behavior

Under regularity conditions (Rodriguez, 2007)

- $\widehat{\eta}_\beta$ and $\widehat{\boldsymbol{\beta}}$ are consistent.
- Denote $\boldsymbol{\beta}^{(q-1)} = (\beta_1, \ldots, \beta_{q-1})^{\mathsf{t}}$. Then:
  - $\sqrt{n}(\widehat{\boldsymbol{\beta}}^{(q-1)} - \boldsymbol{\beta}_0^{(q-1)}) \xrightarrow{D} N(0, \boldsymbol{\Sigma})$
  - $\sqrt{n}(\widehat{\beta}_q - \beta_{0q}) \xrightarrow{p} 0$

Asymptotic Behavior: log–Gamma model

Using the independence between ϵ_i and $\mathbf{x}_i$ and after some algebra, we have

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}^{(q-1)} - \boldsymbol{\beta}_0^{(q-1)}) \xrightarrow{D} N\left(0, \frac{\mathbb{E}_0 \psi^{*\,2}(u_1, c)}{\left(\mathbb{E}\chi^*(u_1, c)\right)^2} \, \widetilde{\mathbf{B}}_1^{-1} \widetilde{\boldsymbol{\Sigma}}_1 (\widetilde{\mathbf{B}}_1^{-1})^{\mathsf{t}}\right)$$

Asymptotic Efficiency

$$e = \frac{\mathbb{E}_0 \psi^{*\,2}(u_1, c)}{\left(\mathbb{E}\chi^*(u_1, c)\right)^2}$$

corresponds to $MM-$estimator B., García Ben & Yohai (2005).

Monte Carlo Experiment

Model: Scheme S_0

$$y_i = \log(z_i) \qquad\qquad 1 \leq i \leq 100$$
$$z_i | x_i \sim \Gamma(3, \mu(x_i))$$

$$\mathbb{E}\,(z_i | x_i) = \mu(x_i)$$
$$\log(\mu(x_i)) = \eta_0(\boldsymbol{\beta}_0^{\mathsf{t}} \mathbf{x}_i)$$

- $\boldsymbol{\beta}_0 = (1,1)/\sqrt{2}$ $\qquad\qquad$ $\mathbf{x}_i \sim \mathcal{U}((0,1) \times (0,1))$
- 1) $\eta_0(t) = \sin(2\pi t)$
  2) $\eta_0(t) = 8(t - \frac{1}{\sqrt{2}})^2$
- $NR = 1000$

Monte Carlo Experiment

We compute the classical and the proposed estimators: *cl* and *r*

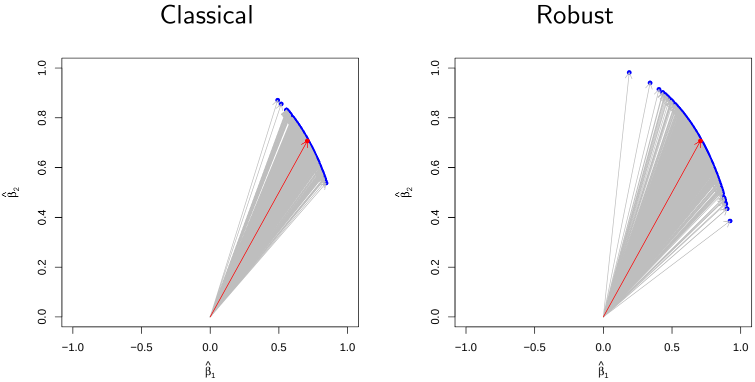
- K: Epanechnikov kernel
- Bandwidths:
  - h_1= 0.15 initial
  - h_{LIN}=0.25 local linear

  We also compute a local constant estimator in Step 3 with
  - h_2=0.20 local constant
- ρ: Tukey's biweight score function

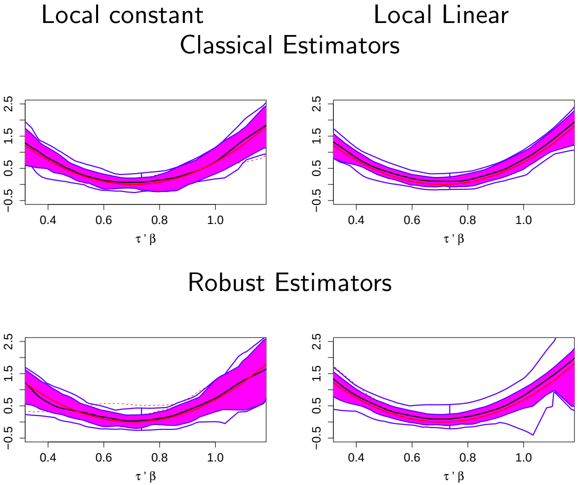
$MSE(\widehat{\boldsymbol{\beta}})$

$\eta(t)$		$\widehat{\boldsymbol{\beta}}_{\mathrm{CL}}$	$\widehat{\boldsymbol{\beta}}_{\mathrm{R}}$
$\sin(2\pi t)$	S_0	**0.003**	0.006
$8(t-\frac{1}{\sqrt{2}})^2$	S_0	**0.006**	0.013

Estimators of $\boldsymbol{\beta}_0$ when $\eta_0(t) = 8(t - \frac{1}{\sqrt{2}})^2$

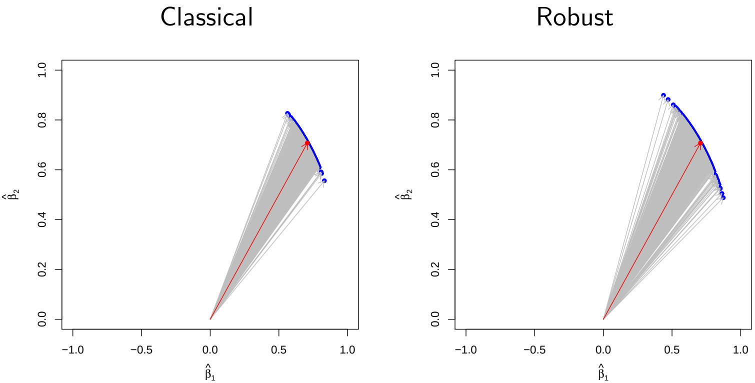


S_0: Estimators of η_0 on a fixed grid $\tau_i^{\mathrm{t}}\boldsymbol{\beta}_0, 1 \leq i \leq 100$



Functionals boxplots when $\eta_0(t) = 8(t - \frac{1}{\sqrt{2}})^2$.

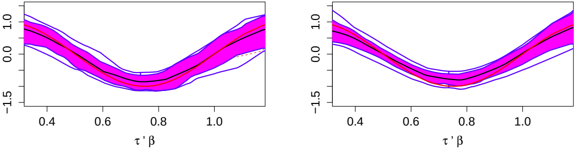
Estimators of $\boldsymbol{\beta}_0$ when $\eta_0(t) = \sin(2\pi t)$



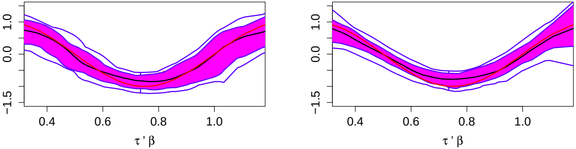
S_0: Estimators of η_0 on a fixed grid $\tau_i^{\mathrm{t}}\boldsymbol{\beta}_0, 1 \leq i \leq 100$

Local constant | Local Linear

Classical Estimators



Robust Estimators



Functionals boxplots when $\eta_0(t) = \sin(2\pi u)$.

Contaminated samples: 10% contamination

Contaminated samples:

we generate $u_i \sim \mathcal{U}(0,1)$ for $1 \leq i \leq 100$ and we take

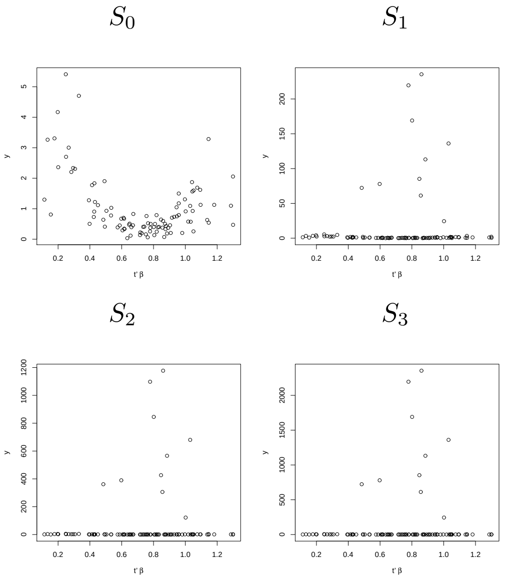
$$z_{i,c} = \begin{cases} z_i & \text{if } u_i \leq 0.90 \\ z_i^\star & \text{if } u_i > 0.90. \end{cases}$$

$$S_1 : z_i^\star \sim \Gamma(3, 100)$$

$$S_2 : z_i^\star \sim \Gamma(3, 500)$$

$$S_3 : z_i^\star \sim \Gamma(3, 1000)$$

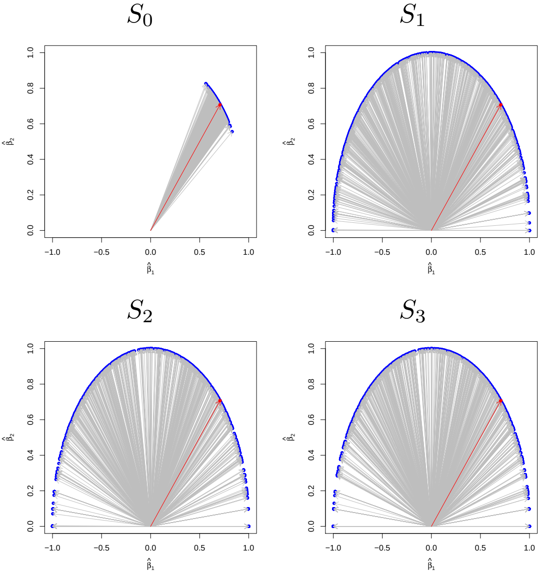
Generated sample when $\eta_0(u) = \sin(2\pi u)$



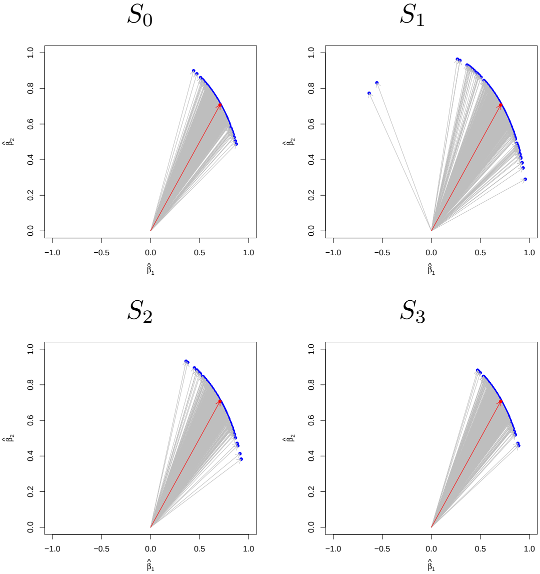
$MSE(\widehat{\boldsymbol{\beta}})$

$\eta(t)$		$\widehat{\boldsymbol{\beta}}_{\mathrm{CL}}$	$\widehat{\boldsymbol{\beta}}_{\mathrm{R}}$
$\sin(2\pi t)$	S_0	**0.003**	0.006
	S_1	1.237	**0.016**
	S_2	1.246	**0.008**
	S_3	1.250	**0.006**
$8(t - \frac{1}{\sqrt{2}})^2$	S_0	**0.006**	0.013
	S_1	1.033	**0.073**
	S_2	1.257	**0.048**
	S_3	1.266	**0.016**

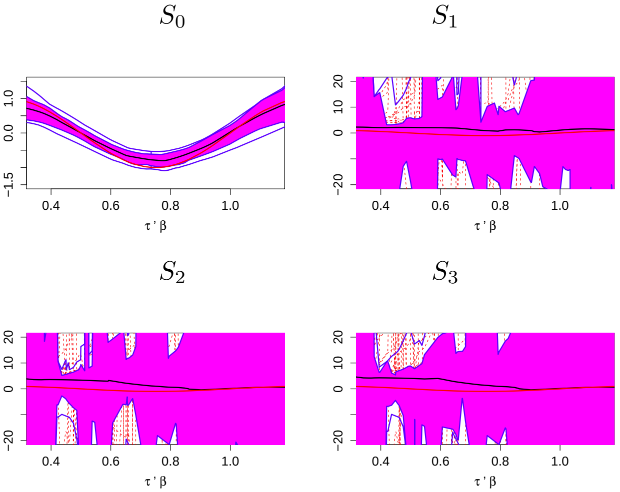
Classical estimators of $\boldsymbol{\beta}_0$ when $\eta_0(t) = \sin(2\pi t)$



Robust estimators of $\boldsymbol{\beta}_0$ when $\eta_0(t) = \sin(2\pi t)$

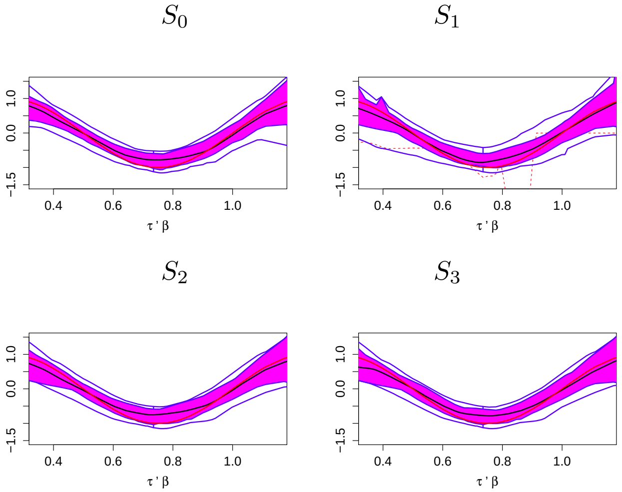


Classical estimators of η_0 when $\eta_0(t) = \sin(2\pi t)$



Functionals boxplots of the Local Linear Estimators of η_0

Robust estimators of η_0 when $\eta_0(t) = \sin(2\pi t)$



Functionals boxplots of the Local Linear Estimators of η_0

Conclusions

- We propose a stepwise robust procedure for the estimation of the link function η_0 and the single index parameter $\boldsymbol{\beta}_0$.
- The resulting estimators are consistent and $\widehat{\boldsymbol{\beta}}^{(q-1)}$ is asymptotically normal.
- The results of our preliminary simulation show the stability of the proposal in the selected contaminated scenarios.

Muchas gracias por su atención!

S_0

$\eta_0(t)$	$\widehat{\boldsymbol{\beta}}_{\mathrm{CL,INI}}$	$\widehat{\boldsymbol{\beta}}_{\mathrm{CL}}$	$\widehat{\boldsymbol{\beta}}_{\mathrm{R,INI}}$	$\widehat{\boldsymbol{\beta}}_{\mathrm{R}}$
$\sin(2\pi t)$	0.004	0.003	0.006	0.006
$8(t-\frac{1}{\sqrt{2}})^2$	0.007	0.006	0.013	0.013

Table 1: MSE of $\widehat{\boldsymbol{\beta}}$.

	$\widehat{\eta}_{\mathrm{CL}}$	$\widehat{\eta}_{\mathrm{CL,LIN}}$	$\widehat{\eta}_{\mathrm{R}}$	$\widehat{\eta}_{\mathrm{R,LIN}}$
$\sin(2\pi t)$	0.031	0.036	0.037	0.037
$8(t-\frac{1}{\sqrt{2}})^2$	0.034	0.024	0.052	0.064

Table 2: Mean over replications of $MSE(\widehat{\eta_0})$.

log–Gamma case: Initial Estimators

Step I.1 For each a, u and $\boldsymbol{\beta}$, compute $s_{n,\boldsymbol{\beta},u}(a)$ solution of

$$\sum_{i=1}^{n} \rho_0 \left(\frac{\sqrt{d(y_i, a)}}{s_{n,\boldsymbol{\beta},u}(a)} \right) W_{\boldsymbol{\beta},\, i}(u, h) = b$$

and $\widetilde{\eta}_{\boldsymbol{\beta}}(u)$ as the value $\widetilde{\eta}_{\boldsymbol{\beta}}(u) = \operatorname{argmin}_a s_{n,\boldsymbol{\beta},u}(a)$

Step I.2 For each $\boldsymbol{\beta}$ let $\widetilde{\sigma}(\boldsymbol{\beta})$ solve

$$\frac{1}{\sum_{i=1}^{n} \tau_{n,\boldsymbol{\beta}}(\mathbf{x}_i)} \sum_{i=1}^{n} \rho_0 \left(\frac{\sqrt{d\left(y_i, \widetilde{\eta}_{\boldsymbol{\beta}}\left(\boldsymbol{\beta}^{\mathsf{t}} \mathbf{x}_i\right)\right)}}{\widetilde{\sigma}(\boldsymbol{\beta})} \right) \tau_{n,\boldsymbol{\beta}}(\mathbf{x}_i) = b$$

$$\widetilde{\boldsymbol{\beta}} = \operatorname*{argmin}_{\|\boldsymbol{\beta}\|=1} \widetilde{\sigma}(\boldsymbol{\beta}) \text{ and } \widehat{s}_n = \widetilde{\sigma}(\widetilde{\boldsymbol{\beta}})$$

Step I.3 Consider $\widehat{\alpha} = S^{\star\, -1}(\widehat{s}_n)$, where $S^{\star}(\alpha)$ is the solution of

$$\mathbb{E}_{\alpha} \rho_0 \left(\frac{\sqrt{-1 - u + \exp(u)}}{S^{\star}(\alpha)} \right) = b$$

General Case

$\phi : \mathbb{R}^3 \to \mathbb{R}$: loss function.

For each $\boldsymbol{\beta}$ and any continuous function $v : \mathbb{R} \to \mathbb{R}$ define

$$\begin{aligned} R(\boldsymbol{\beta}, a, u, \alpha) &= \mathbb{E}_0 \left[\phi\left(y, a, \alpha\right) | \boldsymbol{\beta}^{\mathsf{t}} \mathbf{x} = u \right] \\ G(\boldsymbol{\beta}, v, \alpha) &= \mathbb{E}_0 \left[\phi\left(y, v(\boldsymbol{\beta}^{\mathsf{t}} \mathbf{t}), \alpha\right) \tau_{\boldsymbol{\beta}}(\mathbf{x}) \right] \end{aligned}$$

Denote by $\eta_{\boldsymbol{\beta}}(u) = \text{argmin}_{a \in \mathbb{R}} R(\boldsymbol{\beta}, a, u, \alpha_0)$.

Assume that $\boldsymbol{\beta}_0 = \text{argmin}_{\boldsymbol{\beta} \in \mathbb{R}^q} G(\boldsymbol{\beta}, \eta_0, \alpha_0)$ is the unique minimum.

Mean over replications of $MSE(\widehat{\eta_0})$

$\eta(t)$		$\widehat{\eta}_{\mathrm{CL}}$	$\widehat{\eta}_{\mathrm{CL,LIN}}$	$\widehat{\eta}_{\mathrm{R}}$	$\widehat{\eta}_{\mathrm{R,LIN}}$
$\sin(2\pi t)$	S_0	0.031	0.036	0.037	0.037
	S_1	29.773	630.148	0.460	0.047
	S_2	35.906	227.174	0.356	0.042
	S_3	39.473	350.844	0.072	0.041
$8(t-\frac{1}{\sqrt{2}})^2$	S_0	0.034	0.024	0.052	0.064
	S_1	19.344	2693.669	1.258	0.960
	S_2	28.435	249.353	0.756	0.116
	S_3	31.411	444.717	0.081	0.057