Adaptive estimation of survival function in the convolution model on $\mathbb{R}^+$

Gwennaëlle MABON

CREST - ENSAE & Université Paris Descartes

April, 20th 2016

Motivations: one-sided error in convolution models (a.k.a. additive measurement errors).

→ Application to back calculation problems in AIDS research
Groeneboom and Wellner (1992), van Es et al. (1998), Jongbloed (1998), Groeneboom and Jongbloed (2003).

→ Application in finance (nonparametric regression)
Jirak, Meister and Reiß (2014), Reiß & Selk (2015) .

Motivations: one-sided error in convolution models (a.k.a. additive measurement errors).

$\longrightarrow$ Application to back calculation problems in AIDS research
Groeneboom and Wellner (1992), van Es et al. (1998), Jongbloed (1998), Groeneboom and Jongbloed (2003).

$\longrightarrow$ Application in finance (nonparametric regression)
Jirak, Meister and Reiß (2014), Reiß & Selk (2015) .

Motivations: one-sided error in convolution models (a.k.a. additive measurement errors).

⟶ Application to back calculation problems in AIDS research
Groeneboom and Wellner (1992), van Es et al. (1998), Jongbloed (1998), Groeneboom and Jongbloed (2003).

⟶ Application in finance (nonparametric regression)
Jirak, Meister and Reiß (2014), Reiß & Selk (2015) .

- We study the following model:

$$Z_i = X_i + Y_i, \quad i = 1, \dots, n, \tag{1}$$

  - X_i's i.i.d. nonnegative variables with **unknown** density f, survival function S_X.
  - Y_i's i.i.d. nonnegative variables with **known** density g, survival function S_Y.
  - $(X_i)_i \perp\!\!\!\perp (Y_i)_i$, $Z_i \sim h$, survival function S_Z.

Target: estimation of S_X when the Z_i's are observed and g is known.

- **Assumptions**: S_X, S_Y and g belong to $\mathbb{L}^2(\mathbb{R}^+)$.
- Find an appropriate orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$,$(\varphi_k)_{k\geq 0}$,

$$S_X(x) = \sum_{k\geq 0} a_k(S_X)\varphi_k(x).$$

$a_k(S_X)$: k-th component of S_X in the orthonormal basis.

- Study the MISE of the estimator in this basis.

$$\mathbb{E}\|S_X - \hat{S}_{X,m}\|^2 \leq ?$$

- Build a model selection procedure *à la* Birgé and Massart.

$$\widehat{m} = \arg\min_{m\in\mathcal{M}} \gamma_n(\hat{S}_{X,m}) + \mathrm{pen}(m).$$

- **Assumptions**: S_X, S_Y and g belong to $\mathbb{L}^2(\mathbb{R}^+)$.
- Find an appropriate orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$,$(\varphi_k)_{k\geq 0}$,

$$S_X(x) = \sum_{k\geq 0} a_k(S_X)\varphi_k(x).$$

$a_k(S_X)$: k-th component of S_X in the orthonormal basis.

- Study the MISE of the estimator in this basis.

$$\mathbb{E}\|S_X - \hat{S}_{X,m}\|^2 \leq ?$$

- Build a model selection procedure *à la* Birgé and Massart.

$$\widehat{m} = \arg\min_{m\in\mathcal{M}} \gamma_n(\hat{S}_{X,m}) + \mathrm{pen}(m).$$

- **Assumptions**: S_X, S_Y and g belong to $\mathbb{L}^2(\mathbb{R}^+)$.
- Find an appropriate orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$,$(\varphi_k)_{k\geq 0}$,

$$S_X(x) = \sum_{k\geq 0} a_k(S_X)\varphi_k(x).$$

$a_k(S_X)$: k-th component of S_X in the orthonormal basis.

- Study the MISE of the estimator in this basis.

$$\mathbb{E}\|S_X - \hat{S}_{X,m}\|^2 \leq ?$$

- Build a model selection procedure *à la* Birgé and Massart.

$$\widehat{m} = \arg\min_{m\in\mathcal{M}} \gamma_n(\hat{S}_{X,m}) + \mathrm{pen}(m).$$

- **Assumptions**: S_X, S_Y and g belong to $\mathbb{L}^2(\mathbb{R}^+)$.
- Find an appropriate orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$,$(\varphi_k)_{k\geq 0}$,

$$S_X(x) = \sum_{k\geq 0} a_k(S_X)\varphi_k(x).$$

$a_k(S_X)$: k-th component of S_X in the orthonormal basis.

- Study the MISE of the estimator in this basis.

$$\mathbb{E}\|S_X - \hat{S}_{X,m}\|^2 \leq ?$$

- Build a model selection procedure *à la* Birgé and Massart.

$$\widehat{m} = \arg\min_{m\in\mathcal{M}} \gamma_n(\hat{S}_{X,m}) + \mathrm{pen}(m).$$

Let $z \geq 0$, by definition $S_Z(z) = \mathbb{P}(Z > z)$, we get

$$
\begin{aligned}
S_Z(z) = \mathbb{P}(X+Y>z) &= \iint \mathbb{1}_{x+y>z}\, f(x)\mathbb{1}_{x\geq 0}\, g(y)\mathbb{1}_{y\geq 0}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int \left(\int_{z-y}^{+\infty} f(x)\,\mathrm{d}x\right) g(y)\mathbb{1}_{y\geq 0}\mathbb{1}_{z-y\geq 0}\,\mathrm{d}y \\
&\quad + \int \left(\int_{0}^{+\infty} f(x)\,\mathrm{d}x\right) g(y)\mathbb{1}_{y\geq 0}\mathbb{1}_{z-y\leq 0}\,\mathrm{d}y \\
&= \int_0^z S_X(z-y)g(y)\,\mathrm{d}y + S_Y(z).
\end{aligned}
$$

$$
\boxed{S_Z(z) = S_X \star g(z) + S_Y(z)} \tag{2}
$$

Let $z \geq 0$, by definition $S_Z(z) = \mathbb{P}(Z > z)$, we get

$$\begin{aligned} S_Z(z) = \mathbb{P}(X + Y > z) &= \iint \mathbb{1}_{x+y>z}\, f(x) \mathbb{1}_{x\geq 0}\, g(y) \mathbb{1}_{y \geq 0}\, \mathrm{d}x\, \mathrm{d}y \\ &= \int \left(\int_{z-y}^{+\infty} f(x)\, \mathrm{d}x \right) g(y) \mathbb{1}_{y\geq 0} \mathbb{1}_{z-y\geq 0}\, \mathrm{d}y \\ &\quad + \int \left(\int_{0}^{+\infty} f(x)\, \mathrm{d}x \right) g(y) \mathbb{1}_{y\geq 0} \mathbb{1}_{z-y\leq 0}\, \mathrm{d}y \\ &= \int_0^z S_X(z-y) g(y)\, \mathrm{d}y + S_Y(z). \end{aligned}$$

$$\boxed{S_Z(z) = S_X \star g(z) + S_Y(z)} \tag{2}$$

- For $\mathbb{R}^+$-supported functions, the convolution product writes

$$S_X \star g(z) = \int_0^z S_X(u)g(z-u)\,\mathrm{d}u$$
$$= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} a_k(S_X)a_j(g)\int_0^z \varphi_k(u)\varphi_j(z-u)\,\mathrm{d}u.$$

- We introduce the Laguerre basis defined for $k \in \mathbb{N}, x \geq 0$, by

$$\varphi_k(x) = \sqrt{2}L_k(2x)e^{-x} \quad \text{with} \quad L_k(x) = \sum_{j=0}^{k}\binom{k}{j}\frac{(-x)^j}{j!}.$$

The $(\varphi_k)_k$'s form an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$.

- What makes the Laguerre basis relevant is the relation

$$\int_0^x \varphi_k(u)\varphi_j(x-u)\,\mathrm{d}u = 2^{-1/2}\left(\varphi_{k+j}(x) - \varphi_{k+j+1}(x)\right).$$

- For $\mathbb{R}^+$-supported functions, the convolution product writes

$$S_X \star g(z) = \int_0^z S_X(u) g(z-u)\,\mathrm{d}u$$
$$= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_k(S_X) a_j(g) \int_0^z \varphi_k(u)\varphi_j(z-u)\,\mathrm{d}u.$$

- We introduce the Laguerre basis defined for $k \in \mathbb{N}, x \geq 0$, by

$$\varphi_k(x) = \sqrt{2} L_k(2x) e^{-x} \quad \text{with} \quad L_k(x) = \sum_{j=0}^{k} \binom{k}{j} \frac{(-x)^j}{j!}.$$

The $(\varphi_k)_k$'s form an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$.

- What makes the Laguerre basis relevant is the relation

$$\int_0^x \varphi_k(u)\varphi_j(x-u)\,\mathrm{d}u = 2^{-1/2}\left(\varphi_{k+j}(x) - \varphi_{k+j+1}(x)\right).$$

- For $\mathbb{R}^+$-supported functions, the convolution product writes

$$S_X \star g(z) = \int_0^z S_X(u) g(z-u)\,\mathrm{d}u$$
$$= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_k(S_X) a_j(g) \int_0^z \varphi_k(u)\varphi_j(z-u)\,\mathrm{d}u.$$

- We introduce the Laguerre basis defined for $k \in \mathbb{N}, x \geq 0$, by

$$\varphi_k(x) = \sqrt{2} L_k(2x) e^{-x} \quad \text{with} \quad L_k(x) = \sum_{j=0}^{k} \binom{k}{j} \frac{(-x)^j}{j!}.$$

The $(\varphi_k)_k$'s form an orthonormal basis of $\mathbb{L}^2(\mathbb{R}^+)$.

- What makes the Laguerre basis relevant is the relation

$$\int_0^x \varphi_k(u)\varphi_j(x-u)\,\mathrm{d}u = 2^{-1/2}\left(\varphi_{k+j}(x) - \varphi_{k+j+1}(x)\right).$$

- It yields

$$S_X \star g(z) = \frac{1}{\sqrt{2}} \sum_{k=0}^{\infty} \varphi_k(z) \Big(a_k(S_X) a_0(g) + \sum_{l=0}^{k} \Big(a_{k-l}(g) - a_{k-l-1}(g) \Big) a_l(S_X) \Big) .$$

- Equation implies (2)

$$S_X \star g(z) = S_Z(z) - S_Y(z) = \sum_{k \geq 0} (a_k(S_Z) - a_k(S_Y)) \varphi_k(z)$$

- We obtain for any m that

$$\mathbf{G}_m \vec{S}_{X,m} = \vec{S}_{Z,m} - \vec{S}_{Y,m}$$

$\vec{S}_{\bullet,m} = {}^t(a_0(S_\bullet), \ldots, a_{m-1}(S_\bullet))$.

- $\mathbf{G}_m$ is the lower triangular Toeplitz matrix with elements

$$\mathbf{G}_m = \frac{1}{\sqrt{2}} \begin{cases} a_0(g) & \text{if } i = j, \\ a_{i-j}(g) - a_{i-j-1}(g) & \text{if } j < i, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

- It yields

$$S_X \star g(z)$$
$$= \frac{1}{\sqrt{2}} \sum_{k=0}^{\infty} \varphi_k(z) \Big(a_k(S_X) a_0(g) + \sum_{l=0}^{k} \Big(a_{k-l}(g) - a_{k-l-1}(g) \Big) a_l(S_X) \Big) .$$

- Equation implies (2)

$$S_X \star g(z) = S_Z(z) - S_Y(z) = \sum_{k \geq 0} (a_k(S_Z) - a_k(S_Y)) \varphi_k(z)$$

- We obtain for any m that

$$\mathbf{G}_m \vec{S}_{X,m} = \vec{S}_{Z,m} - \vec{S}_{Y,m}$$

$\vec{S}_{\bullet,m} = {}^t(a_0(S_\bullet), \ldots, a_{m-1}(S_\bullet))$.

- $\mathbf{G}_m$ is the lower triangular Toeplitz matrix with elements

$$\mathbf{G}_m = \frac{1}{\sqrt{2}} \begin{cases} a_0(g) & \text{if } i = j, \\ a_{i-j}(g) - a_{i-j-1}(g) & \text{if } j < i, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

- It yields

$$S_X \star g(z)$$

$$= \frac{1}{\sqrt{2}} \sum_{k=0}^{\infty} \varphi_k(z) \Big(a_k(S_X) a_0(g) + \sum_{l=0}^{k} \Big(a_{k-l}(g) - a_{k-l-1}(g) \Big) a_l(S_X) \Big) .$$

- Equation implies (2)

$$S_X \star g(z) = S_Z(z) - S_Y(z) = \sum_{k \geq 0} (a_k(S_Z) - a_k(S_Y)) \varphi_k(z)$$

- We obtain for any m that

$$\mathbf{G}_m \vec{S}_{X,m} = \vec{S}_{Z,m} - \vec{S}_{Y,m}$$

$$\vec{S}_{\bullet,m} = {}^t(a_0(S_\bullet), \ldots, a_{m-1}(S_\bullet)).$$

- $\mathbf{G}_m$ is the lower triangular Toeplitz matrix with elements

$$\mathbf{G}_m = \frac{1}{\sqrt{2}} \begin{cases} a_0(g) & \text{if } i = j, \\ a_{i-j}(g) - a_{i-j-1}(g) & \text{if } j < i, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

- It yields

$$S_X \star g(z)$$

$$= \frac{1}{\sqrt{2}} \sum_{k=0}^{\infty} \varphi_k(z) \Big(a_k(S_X) a_0(g) + \sum_{l=0}^{k} \Big(a_{k-l}(g) - a_{k-l-1}(g) \Big) a_l(S_X) \Big) .$$

- Equation implies (2)

$$S_X \star g(z) = S_Z(z) - S_Y(z) = \sum_{k \geq 0} (a_k(S_Z) - a_k(S_Y)) \varphi_k(z)$$

- We obtain for any m that

$$\mathbf{G}_m \vec{S}_{X,m} = \vec{S}_{Z,m} - \vec{S}_{Y,m}$$

$\vec{S}_{\bullet,m} = {}^t(a_0(S_\bullet), \dots, a_{m-1}(S_\bullet))$.

- $\mathbf{G}_m$ is the lower triangular Toeplitz matrix with elements

$$\mathbf{G}_m = \frac{1}{\sqrt{2}} \begin{cases} a_0(g) & \text{if } i = j, \\ a_{i-j}(g) - a_{i-j-1}(g) & \text{if } j < i, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

- $\mathbf{G}_m$ is a lower triangular matrix and is invertible *iff* the coefficients of the diagonal are different from 0.

$$a_0(g) = \sqrt{2}\int_{\mathbb{R}^+} g(u)e^{-u}\,\mathrm{d}u = \sqrt{2}\mathbb{E}[e^{-Y}] > 0.$$

- It yields

$$\boxed{\vec{S}_{X,m} = \mathbf{G}_m^{-1}\left(\vec{S}_{Z,m} - \vec{S}_{Y,m}\right)}$$

- Remark:

$$\begin{aligned} a_k(S_Z) &= \int_{\mathbb{R}^+} S_Z(u)\varphi_k(u)\,\mathrm{d}u = \int_{\mathbb{R}^+}\varphi_k(u)\left(\int_u^{+\infty} h(v)\,\mathrm{d}v\right)\mathrm{d}u \\ &= \int_{\mathbb{R}^+}\left(\int_0^v \varphi_k(u)\,\mathrm{d}u\right)h(v)\,\mathrm{d}v = \mathbb{E}\left[\Phi_k(Z_1)\right] \end{aligned}$$

with Φ_k a primitive of φ_k defined as $\Phi_k(x) = \int_0^x \varphi_k(u)\,\mathrm{d}u$.

- $\mathbf{G}_m$ is a lower triangular matrix and is invertible *iff* the coefficients of the diagonal are different from 0.

$$a_0(g) = \sqrt{2}\int_{\mathbb{R}^+} g(u)e^{-u}\,\mathrm{d}u = \sqrt{2}\mathbb{E}[e^{-Y}] > 0.$$

- It yields

$$\boxed{\vec{S}_{X,m} = \mathbf{G}_m^{-1}\left(\vec{S}_{Z,m} - \vec{S}_{Y,m}\right)}$$

- Remark:

$$a_k(S_Z) = \int_{\mathbb{R}^+} S_Z(u)\varphi_k(u)\,\mathrm{d}u = \int_{\mathbb{R}^+}\varphi_k(u)\left(\int_u^{+\infty} h(v)\,\mathrm{d}v\right)\mathrm{d}u$$
$$= \int_{\mathbb{R}^+}\left(\int_0^v \varphi_k(u)\,\mathrm{d}u\right)h(v)\,\mathrm{d}v = \mathbb{E}\left[\Phi_k(Z_1)\right]$$

with Φ_k a primitive of φ_k defined as $\Phi_k(x) = \int_0^x \varphi_k(u)\,\mathrm{d}u$.

- $\mathbf{G}_m$ is a lower triangular matrix and is invertible *iff* the coefficients of the diagonal are different from 0.

$$a_0(g) = \sqrt{2}\int_{\mathbb{R}^+} g(u)e^{-u}\,\mathrm{d}u = \sqrt{2}\mathbb{E}[e^{-Y}] > 0.$$

- It yields

$$\boxed{\vec{S}_{X,m} = \mathbf{G}_m^{-1}\left(\vec{S}_{Z,m} - \vec{S}_{Y,m}\right)}$$

- Remark:

$$\begin{aligned} a_k(S_Z) &= \int_{\mathbb{R}^+} S_Z(u)\varphi_k(u)\,\mathrm{d}u = \int_{\mathbb{R}^+} \varphi_k(u)\left(\int_u^{+\infty} h(v)\,\mathrm{d}v\right)\,\mathrm{d}u \\ &= \int_{\mathbb{R}^+}\left(\int_0^v \varphi_k(u)\,\mathrm{d}u\right) h(v)\,\mathrm{d}v = \mathbb{E}\left[\Phi_k(Z_1)\right] \end{aligned}$$

with Φ_k a primitive of φ_k defined as $\Phi_k(x) = \int_0^x \varphi_k(u)\,\mathrm{d}u$.

- Let $\mathcal{S}_m = \overline{\text{span}}\{\varphi_k\}_{k\in\{0,\dots,m-1\}}$ and consider $S_{X,m}$ the projection of S_X on $\mathcal{S}_m$

$$S_{X,m}(x) = \sum_{k=0}^{m-1} a_k(S_X)\varphi_k(x). \tag{4}$$

Definition (Projection estimator)

$$\widehat{S}_{X,m}(x) = \sum_{k=0}^{m-1} \hat{a}_k\varphi_k(x) \tag{5}$$

$$^t(\hat{a}_0, \dots, \hat{a}_{m-1}) = \widehat{\vec{S}}_{X,m} \quad \text{and} \quad \widehat{\vec{S}}_{Z,m} = {}^t(\hat{a}_0(Z), \dots, \hat{a}_{m-1}(Z))$$

$$\text{with} \quad \widehat{\vec{S}}_{X,m} = \mathbf{G}_m^{-1}\left(\widehat{\vec{S}}_{Z,m} - \vec{S}_{Y,m}\right) \quad \text{and} \quad \hat{a}_k(Z) = \frac{1}{n}\sum_{i=1}^{n}\Phi_k(Z_i),$$

Proposition (M.(2014))

If S_X and $g \in \mathbb{L}^2(\mathbb{R}^+)$ and $\mathbb{E}[Z_1] < \infty$, for $\mathbf{G}_m$ defined by (3) and $\widehat{S}_{X,m}$ defined by (5), the following result holds

$$\mathbb{E}\|S_X - \widehat{S}_{X,m}\|^2 \leq \|S_X - S_{X,m}\|^2 + \frac{\mathbb{E}[Z_1]}{n}\varrho^2(\mathbf{G}_m^{-1}). \quad (6)$$

$\varrho^2(\mathbf{A})$ is the largest eigenvalue of a matrix ${}^t\mathbf{A}\mathbf{A}$ in absolute value.

Consequence

- *m plays the same role as a bandwith parameter.*
  - *m too small $\Rightarrow$ dominant bias.*
  - *m too big $\Rightarrow$ dominant variance.*
- *Choose m to have a trade-off between the bias and the variance.*

Proposition (M.(2014))

If S_X and $g \in \mathbb{L}^2(\mathbb{R}^+)$ and $\mathbb{E}[Z_1] < \infty$, for $\mathbf{G}_m$ defined by (3) and $\widehat{S}_{X,m}$ defined by (5), the following result holds

$$\mathbb{E}\|S_X - \widehat{S}_{X,m}\|^2 \leq \|S_X - S_{X,m}\|^2 + \frac{\mathbb{E}[Z_1]}{n}\varrho^2(\mathbf{G}_m^{-1}). \tag{6}$$

$\varrho^2(\mathbf{A})$ is the largest eigenvalue of a matrix ${}^t\mathbf{A}\mathbf{A}$ in absolute value.

Consequence

- *m plays the same role as a bandwith parameter.*

  m too small $\Rightarrow$ dominant bias.

  m too big $\Rightarrow$ dominant variance.

- *Choose m to have a trade-off between the bias and the variance.*

Goal: define an empirical version of the upper bound on the MISE

$$\|S_X - S_{X,m}\|^2 + \frac{\mathbb{E}[Z_1]}{n}\varrho^2(\mathbf{G}_m^{-1})$$

→ Approximation of the bias term by

$$-\|\widehat{S}_{X,m}\|^2$$

→ Approximation of the variance term by

$$\mathrm{pen}(m) = \frac{\kappa\mathbb{E}[Z_1]}{n}\varrho^2\left(\mathbf{G}_m^{-1}\right)\log n$$

Goal: define an empirical version of the upper bound on the MISE

$$\|S_X - S_{X,m}\|^2 + \frac{\mathbb{E}[Z_1]}{n}\varrho^2(\mathbf{G}_m^{-1})$$

$\longrightarrow$ Approximation of the bias term by

$$-\|\widehat{S}_{X,m}\|^2$$

$\longrightarrow$ Approximation of the variance term by

$$\mathrm{pen}(m) = \frac{\kappa\mathbb{E}[Z_1]}{n}\varrho^2\left(\mathbf{G}_m^{-1}\right)\log n$$

Goal: define an empirical version of the upper bound on the MISE

$$\|S_X - S_{X,m}\|^2 + \frac{\mathbb{E}[Z_1]}{n}\varrho^2(\mathbf{G}_m^{-1})$$

$\longrightarrow$ Approximation of the bias term by

$$-\|\widehat{S}_{X,m}\|^2$$

$\longrightarrow$ Approximation of the variance term by

$$\mathrm{pen}(m) = \frac{\kappa\mathbb{E}[Z_1]}{n}\varrho^2\left(\mathbf{G}_m^{-1}\right)\log n$$

(B1) $\mathcal{M}_n = \left\{1 \leq m \leq n, \varrho^2\left(\mathbf{G}_m^{-1}\right) \log n \leq Cn\right\}$, where $C > 0$.

(B2) $0 < \mathbb{E}[Z_1^3] < \infty$.

Theorem (M.(2014))

If S_X and $g \in \mathbb{L}^2(\mathbb{R}^+)$, let us suppose that **(B1)-(B2)** *are true. Let $\widehat{S}_{X,\widehat{m}}$ be defined by (5) and*

$$\widehat{m} = \operatorname*{argmin}_{m \in \mathcal{M}_n} \left\{-\|\widehat{S}_{X,m}\|^2 + \mathrm{pen}(m)\right\},$$

with $\mathrm{pen}(m) = \dfrac{\kappa \mathbb{E}[Z_1]}{n} \varrho^2\left(\mathbf{G}_m^{-1}\right) \log n$, *then there exists a positive numerical constant* $\kappa \geq \kappa_0$ *such that*

$$\mathbb{E}\|S_X - \widehat{S}_{X,\widehat{m}}\|^2 \leq 4 \inf_{m \in \mathcal{M}_n} \left\{\|S_X - S_{X,m}\|^2 + \mathrm{pen}(m)\right\} + \frac{C}{n},$$

where C is a constant depending on $\mathbb{E}[Z_1^3]$.

(B1) $\mathcal{M}_n = \left\{1 \leq m \leq n, \varrho^2\left(\mathbf{G}_m^{-1}\right)\log n \leq Cn\right\}$, where $C > 0$.

(B2) $0 < \mathbb{E}[Z_1^3] < \infty$.

Theorem (M.(2014))

If S_X and $g \in \mathbb{L}^2(\mathbb{R}^+)$, let us suppose that **(B1)-(B2)** *are true. Let $\widehat{S}_{X,\widehat{m}}$ be defined by (5) and*

$$\widehat{m} = \underset{m\in\mathcal{M}_n}{\text{argmin}} \left\{-\|\widehat{S}_{X,m}\|^2 + \mathrm{pen}(m)\right\},$$

with $\mathrm{pen}(m) = \dfrac{\kappa\mathbb{E}[Z_1]}{n}\varrho^2\left(\mathbf{G}_m^{-1}\right)\log n$, *then there exists a positive numerical constant* $\kappa \geq \kappa_0$ *such that*

$$\mathbb{E}\|S_X - \widehat{S}_{X,\widehat{m}}\|^2 \leq 4 \inf_{m\in\mathcal{M}_n}\left\{\|S_X - S_{X,m}\|^2 + \mathrm{pen}(m)\right\} + \frac{C}{n},$$

where C is a constant depending on $\mathbb{E}[Z_1^3]$.

Corollary (M.(2014))

If S_X and $g \in \mathbb{L}^2(\mathbb{R}^+)$, let us suppose that **(B1)-(B2)** *are true. Let $\widehat{S}_{X,\widetilde{m}}$ be defined by (5) and*

$$\widetilde{m} = \underset{m \in \mathcal{M}_n}{\operatorname{argmin}} \left\{ -\|\widehat{S}_{X,m}\|^2 + \widehat{\mathrm{pen}}(m) \right\}$$

$$\widehat{\mathrm{pen}}(m) = \frac{2\kappa \bar{Z}_n}{n} \varrho^2 \left(\mathbf{G}_m^{-1} \right) \log n \quad \text{where} \quad \bar{Z}_n = \frac{1}{n} \sum_{i=1}^{n} Z_i,$$

then there exists a positive numerical constant $\kappa \geq \kappa_0$ such that

$$\mathbb{E}\|S_X - \widehat{S}_{X,\widetilde{m}}\|^2 \leq 4 \inf_{m \in \mathcal{M}_n} \left\{ \|S_X - S_{X,m}\|^2 + \mathrm{pen}(m) \right\} + \frac{C}{n}$$

where C is a constant depending on $\mathbb{E}[Z_1]$, $\mathbb{E}[Z_1^2]$, $\mathbb{E}[Z_1^3]$.

- Estimation of the survival function in a global setting on $\mathbb{R}^+$.
- Related works:
  - Estimation of the density $\rightarrow$ Mabon (2014).
  - Estimation of linear functionals of the density (c.d.f, pointwise estimation of the density, Laplace transform) $\rightarrow$ Mabon (2015).
- Perspectives:
  - Estimation when g is unknown $\rightarrow$ work in progress.
  - Goodness-of-fit test.

Thank you for your attention.

- Estimation of the survival function in a global setting on $\mathbb{R}^+$.
- Related works:
  - Estimation of the density $\rightarrow$ Mabon (2014).
  - Estimation of linear functionals of the density (c.d.f, pointwise estimation of the density, Laplace transform) $\rightarrow$ Mabon (2015).
- Perspectives:
  - Estimation when g is unknown $\rightarrow$ work in progress.
  - Goodness-of-fit test.

Thank you for your attention.

- van Es, B. (2011). Combining kernel estimators in the uniform deconvolution problem. *Statistica Neerlandica*, 65(3):275–296.
- Groeneboom, P. and Jongbloed, G. (2003). Density estimation in the uniform deconvolution model. *Statistica Neerlandica*, 57(1):136–157.
- Jirak, M, Meister, A. and Reiß, M. (2014). Adaptive function estimation in nonparametric regression with one-sided errors. *The Annals of Statistics*, 42(5):1970–2002.
- Jongbloed, G. (1998). Exponential deconvolution: two asymptotically equivalent estimators. *Statistica Neerlandica*, 52(1):6–17.
- Mabon, G. (2014). Adaptive deconvolution on the nonnegative real line. *preprint MAP5 2014-33*. In revision
- Mabon, G. (2015). Adaptive deconvolution of linear functionals on the nonnegative real line. *preprint MAP5 2015-24*. In revision
- Reiß & Selk (2015). Efficient estimation of functionals in nonparametric boundary models. *To appear in Bernoulli*.