# Estimating a frontier function using a high-order moments method

Gilles STUPFLER (University of Nottingham)
Joint work with Stéphane GIRARD (INRIA Rhône-Alpes) and Armelle GUILLOU (Université de Strasbourg)

EMS 2017 Conference
Helsinki, Finland, 24th July 2017

# Outline

- Context
- Constructing the estimator
- Asymptotic results
- Finite-sample results
- Discussion

# Context

We assume that $Y$ is a univariate random variable recorded along with a finite-dimensional covariate $X$. Suppose that $Y$ given $X = x$ has a finite right endpoint $g(x)$:

$$g(x) := \sup\{y \in \mathbb{R} \,|\, \mathbb{P}(Y \leq y|X = x) < 1\} < \infty.$$

We address the problem of estimating the frontier function $x \mapsto g(x)$.

Practical relevance:

- Temperature/wind speed as a function of 2D/3D coordinates.
- Performance in athletics as a function of age.
- Life span as a function of socioeconomic status.
- Production level as a function of input.

Specifically, assume that the distribution of $(X, Y)$ has support

$$S = \{(x, y) \in \Omega \times \mathbb{R} \,|\, 0 \leq y \leq g(x)\}$$

where

- $X$ has a pdf $f$ on the compact subset $\Omega$ of $\mathbb{R}^d$ having nonempty interior $\text{int}(\Omega)$;
- $g$ is a positive Borel measurable function on $\Omega$.

We consider pointwise estimation of the function $g$ on $\text{int}(\Omega)$, given an $n-$sample of i.i.d. replications of $(X, Y)$.

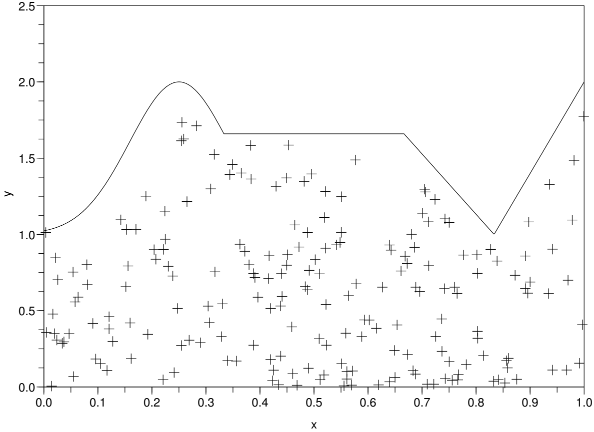


Figure 1: Frontier $g$ (solid line), data points (+). The sample size is $n = 200$.

# Some existing methods

- Extreme-value based estimators: Geffroy (1964), Gardes (2002), Girard and Jacob (2003a, 2003b, 2004), Girard and Menneteau (2005), Menneteau (2008).
- Optimization methods and linear programming: Bouchard *et al.* (2004, 2005), Girard *et al.* (2005), Nazin and Girard (2014).
- Piecewise polynomial estimators: Korostelev and Tsybakov (1993), Korostelev *et al.* (1995), Härdle *et al.* (1995).
- Projection estimators: Jacob and Suquet (1995).

If $g$ is nondecreasing and concave:

- DEA/FDH estimators and improvements: Deprins *et al.* (1984), Farrell (1957), Gijbels *et al.* (1999).
- Robust estimators: Aragon *et al.* (2005), Cazals *et al.* (2002), Daouia and Simar (2005), Daouia *et al.* (2012).
- Local MLE (with random noise): Aigner *et al.* (1976), Fan *et al.* (1996), Kumbhakar *et al.* (2007), Simar and Zelenyuk (2011).

# Constructing the estimator

Assume first that $Y$ is a positive random variable with finite right endpoint $\theta$. Denote by $\mu_p := \mathbb{E}(Y^p)$.

## Proposition 1

*It holds that* $\mu_p/\mu_{p+1} \to 1/\theta$ *as* $p \to \infty$.

Given data points $Y_1, \ldots, Y_n$, this opens a number of ways to estimate $\theta$ from the class of empirical high-order moments

$$\widehat{\mu}_p = \frac{1}{n} \sum_{k=1}^{n} Y_k^p \quad \text{with} \quad p = p_n \to \infty.$$

However, the most direct of such estimators, namely $\widetilde{\theta}_n = \widehat{\mu}_{p_n+1}/\widehat{\mu}_{p_n}$, is in practice too biased to be used.

A better alternative is given by, for some tuning constant $a > 0$,

$$\frac{1}{\widehat{\theta}_n} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}}{\widehat{\mu}_{(a+1)p_n+1}} - (p_n+1)\frac{\widehat{\mu}_{p_n}}{\widehat{\mu}_{p_n+1}}\right].$$

- This estimator is motivated by the elimination of the bias term when the survival function of $Y$ is

$$\forall y \in [0,\theta],\ \overline{F}(y) := \mathbb{P}(Y > y) = \left(1 - \frac{y}{\theta}\right)^{\alpha}.$$

- High-order moments allow to control the bias brought by general survival functions with polynomial decay near the endpoint.

# High-order moments frontier estimator

Our previous construction suggests the following estimator of $g(x)$:

$$\frac{1}{\widehat{g}_n(x)} = \frac{1}{ap_n}\left[((a+1)p_n+1)\frac{\widehat{\mu}_{(a+1)p_n}(x)}{\widehat{\mu}_{(a+1)p_n+1}(x)} - (p_n+1)\frac{\widehat{\mu}_{p_n}(x)}{\widehat{\mu}_{p_n+1}(x)}\right]$$

where $\widehat{\mu}_p(x)$ is a well-behaved estimator of the conditional $p$th order moment $\mu_p(x) := \mathbb{E}(Y^p|X=x)$.

We choose this estimator to be the smoothed estimator

$$\widehat{\mu}_{p,h_n}(x) := \frac{1}{nh_n^d}\sum_{k=1}^{n} Y_k^p K\left(\frac{x-X_k}{h_n}\right).$$

Here, $K$ is a kernel function, *i.e.* a bounded pdf on $\mathbb{R}^d$ with support included in the unit Euclidean ball $B \subset \mathbb{R}^d$, and $h_n > 0$ is a bandwidth sequence that converges to 0.

Our (kernel) estimator $\widehat{g}_n(x)$ of $g(x)$ is then defined by

$$\frac{a p_n}{\widehat{g}_n(x)} = ((a+1)p_n + 1)\frac{\widehat{\mu}_{(a+1)p_n,h_n}(x)}{\widehat{\mu}_{(a+1)p_n+1,h_n}(x)} - (p_n+1)\frac{\widehat{\mu}_{p_n,h_n}(x)}{\widehat{\mu}_{p_n+1,h_n}(x)}.$$

For ease of exposition, assume that we work in the parametric setting

**(P)** $\forall y \in [0, g(x)]$, $\overline{F}(y|x) = (1 - y/g(x))^{-1/\gamma(x)}$, with $\gamma(x) < 0$.

**(A)** $f$, $g$ and $\gamma$ are positive and Hölder continuous on $\Omega$ with respective exponents $\eta_f$, $\eta_g$ and $\eta_\gamma$.

A departure from **(P)** is actually allowed (if we stay within the Hall class).

# Asymptotic results

**Theorem 1 (Pointwise consistency, frontier estimator)**

*If* $np_n^{1/\gamma(x)} h_n^d \to \infty$ *and* $p_n\, h_n^{\eta_g} \to 0$, *then* $\widehat{g}_n(x) \xrightarrow{\mathbb{P}} g(x)$.

**Theorem 2 (Asymptotic normality, frontier estimator)**

*If* $np_n^{1/\gamma(x)} h_n^d \to \infty$, $np_n^{2+1/\gamma(x)} h_n^{d+2\eta_g} \to 0$ *and* $np_n^{1/\gamma(x)} h_n^{d+2\eta_\alpha} \to 0$, *then*

$$\sqrt{np_n^{2+1/\gamma(x)} h_n^d}\left(\frac{\widehat{g}_n(x)}{g(x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\int_B K^2}{f(x)} V(\gamma(x), a)\right)$$

*where* $V(\gamma, a)$ *is explicitly known.*

Uniform consistency results on compacta $E \subset \text{int}(\Omega)$ are also available.

# Finite-sample results

We examined the finite-sample performance of the estimator depending on the value of:

- the frontier function $g$ (smooth or not),
- the extreme-value index function $\gamma$ (constant or not),
- the dimension $d = 1$ or $2$.

We also checked for robustness against a violation of model **(P)**, and we compared the estimator to:

- the block maxima estimator of Geffroy (1964),
- a primitive version of the high order moments method, constructed for a conditional uniform model by Girard and Jacob (2008).

In general, the estimator $\widehat{g}_n$ significantly outperformed these competitors w.r.t. the $L^1$ metric, even though the choices of $a$, $p_n$ and $h_n$ were crude.

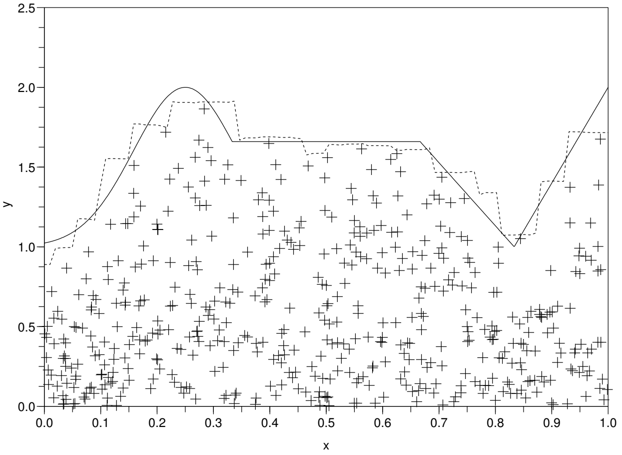


Figure 2: Case $\gamma(x) = -2[2.5 + |\cos(2\pi x)|]^{-1}$: frontier function $g$ (solid line), high-order moments estimate $\widehat{g}_n$ (dotted line) corresponding to the best result among 500 replications of a sample of size 500.

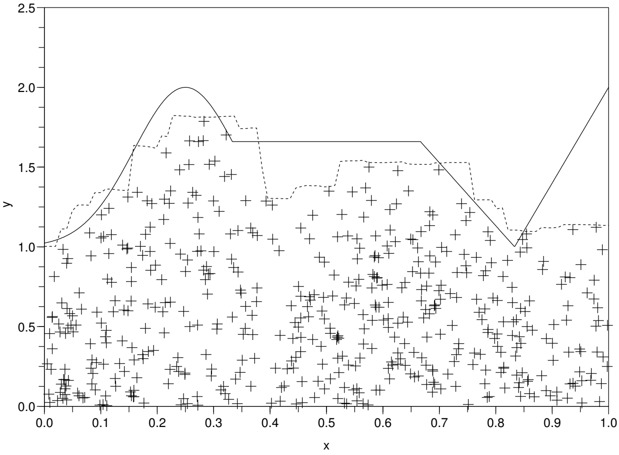


Figure 3: Case $\gamma(x) = -2[2.5 + |\cos(2\pi x)|]^{-1}$: frontier function $g$ (solid line), high-order moments estimate $\widehat{g}_n$ (dotted line) corresponding to the worst result among 500 replications of a sample of size 500.

# Discussion

Conclusions:

- High order moments provide a class of interesting devices when it comes to estimating an endpoint/frontier function.
- The order $p_n$ is a substitute for the effective sample size $k_n$ of extreme-value methods.
- The presented estimator has satisfactory finite-sample performance even with fairly simple choices of tuning parameters.

Some ideas for further studies:

- Development of data-driven choice procedures of tuning parameters;
- Construction of outlier-resistant high order moments procedures;
- Building estimators adapted to high-dimensional data sets.

# References

High-order moments estimator for a frontier function:

Girard, S., Guillou, A., Stupfler, G. (2013). Frontier estimation with kernel regression on high order moments, *Journal of Multivariate Analysis* **116**: 172–189.

Uniform versions of the asymptotic results:

Girard, S., Guillou, A., Stupfler, G. (2014). Uniform strong consistency of a frontier estimator using kernel regression on high order moments, *ESAIM: Probability and Statistics* **18**: 642–666.

**Thanks for listening!**