Nonparametric estimation under shape constraints, part 2

Piet Groeneboom, Delft University

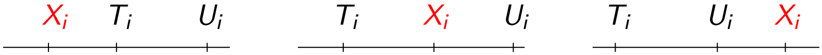
August 7, 2013

What to expect?

- Theory and open problems for interval censoring, case 2.
- Same for the bivariate current status model.
- Convex regression.
- The convex envelope of one-sided Brownian motion without drift.
- Does bootstrapping from the Grenander estimate work for global statistics?

Interval censoring, case 2

$X_1, X_2, \ldots, X_n \sim F_0$.
Instead of observing the X_i's, one only observes $X_i \leq T_i$ or $X_i \in (T_i, U_i]$ or $X_i > U_i$, for some random pair (T_i, U_i), where $T_i < U_i$, (T_i, U_i) independent of X_i.



$\delta_{i1} = 1$ $\qquad$ $\delta_{i2} = 1$ $\qquad$ $\delta_{i1} = \delta_{i2} = 0$

So, instead of observing X_i's, one observes

$$(T_i, U_i, \delta_{i1}, \delta_{i2}) = (T_i, U_i, 1_{\{X_i \leq T_i\}}, 1_{\{X_i \in (T_i, U_i]\}}).$$

where (T_i, U_i) is independent of X_i.

Interval censoring model

We want to estimate the unknown distribution function F_0 of X_i, based on the data $(T_i, U_i, \delta_{i1}, \delta_{i2})$.

The log likelihood function in F (conditional on the (T_i, U_i)'s) is, taking $\delta_{i3} = 1 - \delta_{i1} - \delta_{i2}$:

$$\sum_{i=1}^{n} \left\{\delta_{i1} \log F(T_i) + \delta_{i2} \log(F(U_i) - F(T_i)) + \delta_{i3} \log(1 - F(U_i))\right\}.$$

The (nonparametric) maximum likelihood estimator (MLE) $\hat{F}_n$ maximizes the log likelihood over the class of *all* distribution functions F.

Interval censoring, case 2, algorithms

Algorithms for computing the MLE:

1. EM algorithm
   Start for example with the discrete uniform distribution on a subset of the observation points and iterate:

   $$F^{(m+1)}(t) = n^{-1} \sum_{i=1}^{n} P^{(m)} \{X \le t | T_j, U_j, \delta_{j1}, \delta_{j2}, j = 1, \ldots, n\}$$

   Very slow!
2. Iterative convex minorant algorithm (Groeneboom (1991), using the modification in Jongbloed (1998)).
3. Support reduction algorithm (Groeneboom, Jongbloed, and Wellner (2008)).
   Iterative cone projection, starting with a minimal "feasible" (finite likelihood) solution. Available in R (MLEcens).

Interval censoring, case 2, non-separated case

Asymptotic local distribution?

Conjecture (Groeneboom (1991))

Let G be the distribution function of (T_i, U_i) and let F_0 and G be continuously differentiable at t_0 and (t_0, t_0), respectively, with strictly positive derivatives $f_0(t_0)$ and $g(t_0, t_0)$, Let $\hat{F}_n$ be the MLE of F_0. Then

$$(n \log n)^{1/3} \left\{ \hat{F}_n(t_0) - F_0(t_0) \right\} \big/ \left\{ 6 f_0(t_0)^2 / g(t_0, t_0) \right\}^{1/3} \xrightarrow{\mathcal{D}} Z,$$

where $Z = \text{argmax}_t \{W(t) - t^2\}$.

Still not proved!

Local rate

1. Shown in Groeneboom (1991): the conjecture is true for a "toy" estimator, obtained by doing one step of the iterative convex minorant algorithm, starting the iterations at the underlying distribution function F_0.
   Birgé (1999) has constructed a histogram-type estimator, achieving the local rate $(n \log n)^{1/3}$ in this model. (Minimax) rate is faster than the rate in the current status model!
2. If the times T_i and U_i are separated, that is:
   $$\mathbb{P}\{U_i - T_i < \epsilon\} = 0,$$
   for some $\epsilon > 0$, the rate drops to $n^{1/3}$.
3. The asymptotic distribution of the MLE can in this case be proved to be the same as the distribution of the toy estimator (Groeneboom (1996)). Limit distribution is again $Z = \text{argmax}\{W(t) - t^2\}$. Variance: Wellner (1995).

Lucien Birgé and Jon Wellner

Smooth functionals for interval censoring

- The nonlinear aspect of the functional is negligible.

$$\sqrt{n}\left\{K(\hat{F}_n) - K(F_0)\right\} = \sqrt{n}\int \kappa_{F_0} d\big(\hat{F}_n - F_0\big) + o_p(1).$$

- Transformation to the observation space measure.

$$\int \kappa_{F_0}\, d\big(\hat{F}_n - F_0\big) = -\int \theta_{\hat{F}_n}(t,\delta)\, dQ_0(t,\delta),$$

where $\theta_{F_0}(t,\delta)$ and $\theta_{\hat{F}_n}(t,\delta)$ are defined via the solutions of integral equations. No explicit solutions for $\theta_{F_0}(t,\delta)$ and $\theta_{\hat{F}_n}(t,\delta)$!

Smooth functionals for interval censoring

- Use that $\hat F_n$ is the MLE.
  Replace $\theta_{\hat F_n}$ by $\bar\theta_{\hat F_n}(t,\delta)$, where $\bar\theta_{\hat F_n}(t,\delta)$ satisfies:

$$\int \bar\theta_{\hat F_n}(t,\delta)\,dQ_n = 0, \tag{1}$$

  and write:

$$\int \kappa_{F_0}\,d\big(\hat F_n - F_0\big) \overset{\text{step 2}}{=} -\int \theta_{\hat F_n}\,dQ_0$$

$$\overset{(1)}{=} \int \bar\theta_{\hat F_n}\,d\big(Q_n - Q_0\big) - \int \Big\{\theta_{\hat F_n} - \bar\theta_{\hat F_n}\Big\}\,dQ_0.$$

- Asymptotic variance equals information lower bound.

$$\int \bar\theta_{\hat F_n}\,d\big(Q_n - Q_0\big) = \int \theta_{F_0}\,d\big(Q_n - Q_0\big) + o_p\left(n^{-1/2}\right),$$

In: Geskus and Groeneboom (1996, 1997 and 1999).

SMLE and MSLE?

Once the MLE $\hat{F}_n$ is computed, we can easily compute the SMLE by

$$\int \mathbb{K}_h(t-x)\,d\hat{F}_n(x), \qquad \mathbb{K}_h(u) = \int_{-\infty}^{u/h} K(x)\,dx.$$

Theory has to use local smooth functional theory again, see Groeneboom and Ketelaars (2011).

Computing the MSLE is harder. Local limit for separated case is determined in Groeneboom (2012). Proof is based on the solution of a non-linear integral equation.

Deconvolution

$$Z_i = X_i + Y_i \sim h_0, \quad h_0(z) = \int g(z-x)\,dF_0(x),\ z \geq 0,$$

g is a known decreasing continuous density on $[0,\infty)$.
F_0 has support, contained in $[0,\infty)$.
MLE maximizes over F:

$$\sum_{i=1}^{n} \log \int g(Z_i - x)\,dF(x).$$

Conjecture (Groeneboom (1991))

At an interior point t of the support of F_0:

$$n^{1/3}\left(\frac{g(0)^2}{4f_0(t)h_0(t)}\right)^{1/3}\left\{\hat{F}_n(t) - F_0(t)\right\} \longrightarrow Z,$$

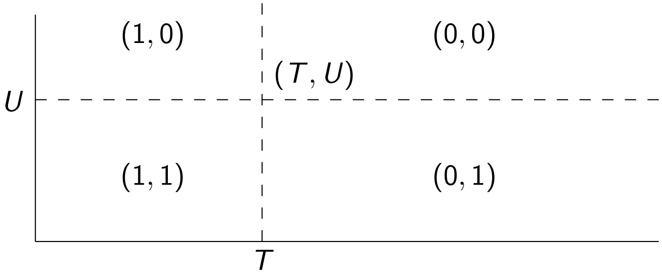
where $Z = \text{argmax}\{W(t) - t^2\}$.

Bivariate current status I

In the bivariate current status model the observations consist of a quadruple $(T, U, \delta_1, \delta_2)$, where

$$\delta_1 = 1_{\{X \le T\}},\, \delta_2 = 1_{\{Y \le U\}}, \tag{2}$$

and (X, Y) is independent of the observation (T, U).



Bivariate current status II

A maximum likelihood estimator $\hat{F}_n$ of F_0, the distribution function of (X, Y), maximizes

$$\int \delta_1\delta_2 \log F(u,v)\, d\,\mathbb{P}_n + \int \delta_1(1-\delta_2)\log\{F_1(u) - F(u,v)\}\; d\,\mathbb{P}_n$$
$$+ \int (1-\delta_1)\delta_2 \log\{F_2(v) - F(u,v)\}\; d\,\mathbb{P}_n$$
$$+ \int (1-\delta_1)(1-\delta_2)\log\{1 - F_1(u) - F_2(v) + F(u,v)\}\; d\mathbb{P}_n$$

over F, where F_1 and F_2 are the first and second marginal dfs of F, respectively, and $\mathbb{P}_n$ is the empirical measure of the observations.
Difficulty: we cannot assume that the mass is located in the observation points.
Preliminary reduction algorithm to find the points of possible mass.

Theorem (Groeneboom (2013))

Consider an interior point (t, u), and define the square A_n, with midpoint (t, u), by:

$$A_n = [t - n^{-1/6}, t + n^{-1/6}] \times [u - n^{-1/6}, u + n^{-1/6}].$$

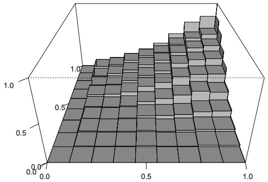
Then, under some regularity conditions, the plug-in estimator

$$\tilde{F}_n(t, u) \stackrel{def}{=} \frac{\int_{A_n} \delta_1 \delta_2 \, d\mathbb{P}_n(v, w, \delta_1, \delta_2)}{\int_{A_n} d\mathbb{G}_n(v, w)}, \qquad (3)$$

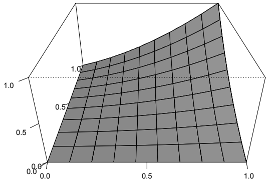
where $\mathbb{G}_n$ is the empirical distribution function of the observations (T_i, U_i), satisfies:

$$n^{1/3} \left\{ \tilde{F}_n(t, u) - F_0(t, u) \right\} \xrightarrow{\mathcal{D}} N\left(\beta, \sigma^2\right),$$

where $N(\beta, \sigma^2)$ is a normal distribution with (specified) parameters β and σ^2.



(a) Plug-in estimate

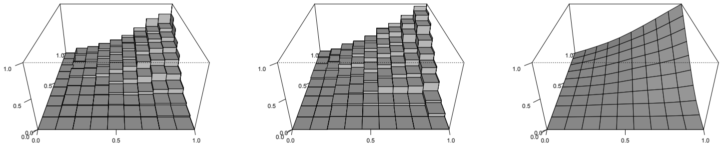


(b) $F_0(x,y) = \frac{1}{2}xy(x+y)$

Figure: Plug-in estimate for a sample of size $n = 1000$ from $F_0(x,y) = \frac{1}{2}xy(x+y)$ on $[0,1]^2$.

The grid has a width of order $n^{-1/3}$, but the binwidth of the estimator is of order $n^{-1/6}$!
The plug-in estimate is not a distribution function (for has -some- negative masses)!

The plug-in estimate is compared with the MLE on a grid and smoothed maximum likelihood estimator (SMLE) (taking bandwidths $n^{-1/6}$ in both directions) in a simulation study in Groeneboom (2013).



(a) Plugin (b) MLE (c) SMLE

Figure: MLE, SMLE and plugin estimate for a sample of size $n = 1000$ from $F_0(x, y) = \frac{1}{2}xy(x + y)$ on $[0, 1]^2$.

The SMLE is (modulo boundary correction) defined by

$$\hat{F}_{nh}^{(SML)}(t,u) = \int \mathbb{K}_h(t-v)\mathbb{K}_h(u-w)\,d\hat{F}_n(v,w),$$

where, for a symmetric kernel K,

$$\mathbb{K}_h(x) = \int_{-\infty}^{x/h} K(y)\,dy,$$

and $\hat{F}_n$ is the maximum likelihood estimator on a grid.

1. SMLE will have rate $n^{-1/3}$, if bandwidth $\asymp n^{-1/6}$.
2. We can probably achieve higher rates for the SMLE, but then we have to use higher order kernels.
3. Rate of the MLE is unknown.
4. Certain minimax calculations suggest that the rate for the MLE will contain logarithmic factors, causing a rate slower than $n^{-1/3}$.

Convex regression

The relationship between age and log(income) for Canadian income data can be expected to be concave.

We can estimate this relationship only using the concavity restriction.

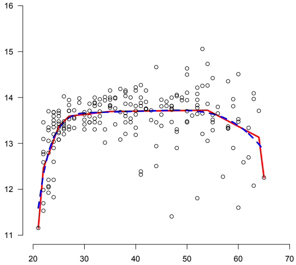


Figure: Concave cubic spline estimate with 5 knots at equal quantile distances (Meyer (2008), blue, dashed) and nonparametric isotonic estimate (red)

- For the usual cubic spline estimation one would have to specify the location of the knots in advance. For example, the estimate in Meyer (2008) uses equal quantile distances.
- The isotonic least squares estimate chooses the locations of the knots automatically. It minimizes the criterion

$$\sum_{i=1}^{n} \{Y_i - f(t_i)\}^2$$

just under the restriction that f is convex or concave.

The nonparametric convex LS estimate

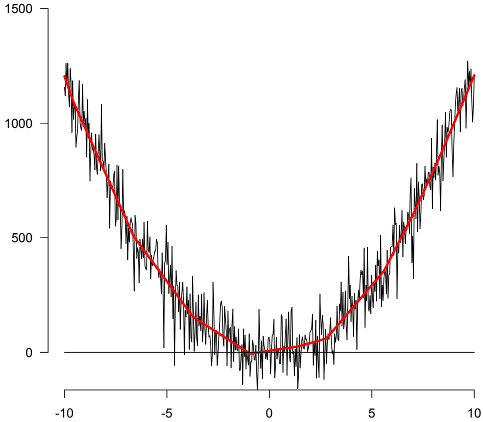


Figure: $12x^2+$ normals

Local limit distribution of the convex regression estimate

$n^{2/5}\{\hat{f}_n(t) - f_0(t)\}$ converges in distribution to the value at zero of the limit, as $c \to \infty$, of the minimizer f_c of the quadratic form

$$\frac{1}{2}\int_{-c}^{c} f(x)^2\,dx - \int_{-c}^{c} f(x)\,d\left(W(x) + 4x^3\right), \qquad f(\pm c) = 12c^2,$$

where W is standard two-sided Brownian motion, originating from zero.

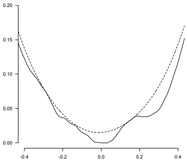


Figure: The functions Y (solid) and H (dashed) for standard two-sided Brownian motion on $[-0.4, 0.4]$.

The limit is the second derivative of the unique cubic spline H lying above and touching Y = integrated Brownian motion $+t^4$, the invelope: Groeneboom, Jongbloed, and Wellner (2001b)

Open problem: are the points of touch isolated?

The nonparametric convex LS estimate

The nonparametric convex LS estimate will probably by asymptotically locally minimax in the same way as the Grenander estimator, but a proof of this fact is still lacking!

Figure: Eric Cator

The concave majorant of one-sided Brownian motion

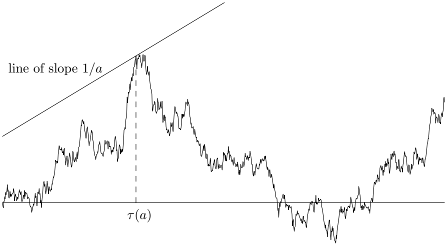


Figure: $\tau(a) = \text{argmax}_x \{W(x) - \frac{x}{a} : x \geq 0\}$

We have again the switch relation:

$$S_t \geq 1/a \iff t \leq \tau(a),$$

where S_t is the slope of the concave majorant at time t.

The argmax process $a \mapsto \tau(a)$ for Brownian motion

Theorem (Groeneboom (1983))

1. *The argmax process $a \mapsto \tau(a)$ is a time inhomogeneous process with independent increments, and, for $u > 0$,*

$$\frac{\mathbb{P}\left\{\tau(a+h) - \tau(a) \in du \,\middle|\, \tau(a) = t\right\}}{h} \sim \frac{e^{-u/(2a^2)}}{a^2\sqrt{2\pi u}}\, du,\ h \downarrow 0.$$

2. *Let $N(a,b)$ be the number of jumps of τ in $[a,b]$. Then*

$$N(a,b) \stackrel{\mathcal{D}}{=} \mathrm{Poisson}\left(\log(b/a)\right).$$

3. *As a consequence of 2:*

$$\left\{N(a,b) - \log(b/a)\right\} / \sqrt{\log(b/a)} \xrightarrow{\mathcal{D}} N(0,1),\ b/a \to \infty.$$

Corollary (Groeneboom (1983))

1. *Brownian motion on $[0, \infty)$ can be decomposed into the argmax process τ and Brownian excursions.*
2. *If S_n is the slope of the concave majorant of the uniform empirical process $U_n = \sqrt{n}\{\mathbb{F}_n - F\}$ on $[0, 1]$, then*

$$\left\{ \int_0^1 S_n(t)^2 \, dt - \log n \right\} / \sqrt{3 \log n} \xrightarrow{\mathcal{D}} N(0, 1).$$

Part 2 uses Doob's transformation (to go from Brownian motion to Brownian bridge) and Hungarian embedding.

Similar methods yield for the number of jumps N_n of the concave majorant of the uniform empirical process $U_n = \sqrt{n}\{\mathbb{F}_n - F\}$:

Theorem (Sparre Andersen (1954))

$$\{N_n - \log n\} / \sqrt{\log n} \xrightarrow{\mathcal{D}} N(0, 1).$$

Ronald Pyke

Theorem (Groeneboom and Pyke (1983))

If S_n is the slope of the concave majorant of the uniform empirical process $U_n = \sqrt{n}\{\mathbb{F}_n - F\}$ on $[0,1]$, then

$$\left\{\int_0^1 S_n(t)^2\,dt - \log n\right\} / \sqrt{3\log n} \xrightarrow{\mathcal{D}} N(0,1).$$

Order the induced spacings between locations of vertices of the least concave majorant:

$$D_{n0}; D_{n1,1}, \ldots, D_{n1,J_{n1}}; \ldots; D_{ni,1}, \ldots, D_{ni,J_{ni}}; \ldots$$

where there are J_{ni} i-step spacings. Then:

$$\int_0^1 S_n(t)^2\,dt = \sum_{i=1}^{n} \left\{ \sum_{j=1}^{J_{ni}} \frac{i^2}{nD_{n,ij}} - 1 \right\}.$$

Replace this by:

$$\sum_{i=1}^{n} \left\{ \sum_{j=1}^{N_i} \frac{i^2}{S_{n,ij}} - 1 \right\}.$$

where (independently) $N_i \sim \text{Poisson}(1/i)$, $S_{n,ij} \sim \Gamma(1, i)$ and condition on

$$\sum_{i=1}^{n} iN_i = n, \qquad \sum_{i=1}^{n} \sum_{j=1}^{N_i} S_{n,ij} \sim n.$$

Further work on this topic

Pitman (1983): interpretation in terms of Bessel processes and path decomposition (David Williams).

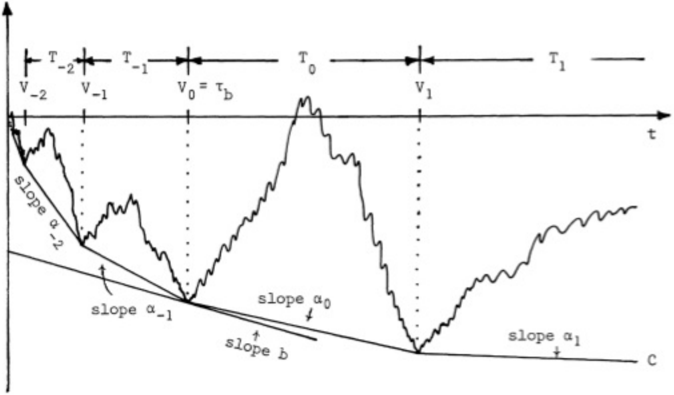


Figure 1. Convex minorant of B is C. The vertex set consists of the points ..., V_{-2}, V_{-1}, V_0, V_1, V_2,...

Theorem (Pitman (1983))

Fix $b \in (-\infty, 0)$ and let $T_0 = \tau_b = argmin\{W(x) - bx\}$. Then:

(i) *The next slope α_0 of the convex minorant is uniformly distributed on $(b, 0)$, and conditionally on $\alpha_0, \dots, \alpha_n$, the next slope α_{n+1} is uniform on $(\alpha_n, 0)$.*

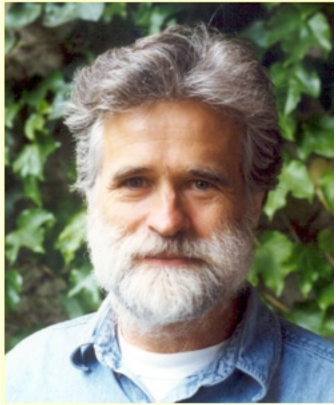
(ii) *The preceding slope α_{-1} has density $|b|x^{-2}$ on the interval $(-\infty, b)$, and, conditional on $\alpha_{-n}, \dots, \alpha_{-1}$, α_{-n-1} has density $|\alpha_{-n}|x^{-2}$ on $(-\infty, \alpha_{-n})$.*

(iii) *The sequences $\{\alpha_i,\ i < 0\}$ and $\{\alpha_i,\ i \geq 0\}$ are independent.*

(iv) *Conditional on all the slopes α_i, the lengths of the segments T_i are independent, and T_i has a gamma$(\frac{1}{2}, \frac{1}{2}\alpha_i^2)$ distribution:*

$$\mathbb{P}\left\{ T_i \in dt \;\middle|\; \alpha_i = a,\ \alpha_j, j \neq i \right| = \frac{|a|}{\sqrt{2\pi t}}\, e^{-\frac{1}{2}a^2 t} dt.$$

Jim Pitman



Çinlar (1992): connection with queueing systems ("Sunset over Brownistan"),

Balabdaoui and Pitman (2011): maximal difference between Brownian bridge and its concave majorant,

Pitman and Ross (2012): greatest convex minorant of Brownian motion, meander, and bridge,

Pitman and Uribe Bravo (2012): the convex minorant of a Lévy process.

The process $a \mapsto V(a) = \text{argmax}\{W(t) - (t-a)^2\}$

Figure 4.1 in Groeneboom (1989) (PTRF):

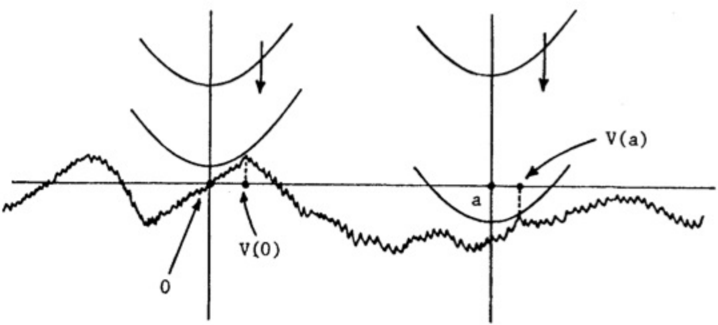


Fig. 4.1

(Stationary) point process of locations of maxima:

$$V(a) = \mathrm{argmax}_x \{W(x) - (x-a)^2 \text{ is maximal}\}.$$

Theorem (Groeneboom (1985), Groeneboom, Hooghiemstra, and Lopuhaä (1999))

Let f be a twice differentiable decreasing density on [0,1]. Then (under some additional conditions) we have, with $\mu = E|V(0)| \int_0^1 |4f'(t)f(t)|^{1/3}\,dt$,

$$n^{1/6}\left\{n^{1/3}\int_0^1 |\hat{f}_n(t) - f(t)|\,dt - \mu\right\} \xrightarrow{\mathcal{D}} N(0,\sigma^2),$$

where $\sigma^2 = 8\int_0^\infty \mathrm{covar}(|V(0)|, |V(c)-c|)\,dc$.

Durot-Lopuhaä k-sample tests

Setting: we want to test the hypothesis

$$H_0 : f_1 = f_2 = \cdots = f_J \quad \text{against} \quad H_1 : f_i \neq f_j, \text{ for some } i \neq j,$$

where $f_j : [0, B] \to \mathbb{R}$ is decreasing. The f_j can be densities, regression functions, etc. Consider the test statistic:

$$T_N = \sum_{j=1}^{J} \int_0^B \left| \hat{f}_{j,n_j}(t) - \hat{f}_N(t) \right| dt, \qquad N = \sum_{j=1}^{J} n_j,$$

where $\hat{f}_{j,n_j}$ is the isotonic estimate in the jth sample of size n_j (for example the Grenander estimate), and $\hat{f}_N$ is the isotonic estimate in the combined samples.
We want to use the bootstrap to find a critical value c such that (for example):

$$\mathbb{P}\{T_N \geq c\} = 0.05.$$

Rik Lopuhaä and Cécile Durot

Durot and Lopuhaä: under a number of regularity conditions there are constants μ and σ such that

$$N^{1/6}\left\{N^{1/3}T_N-\mu\right\}/\sigma \xrightarrow{\mathcal{D}} N(0,1), \qquad (4)$$

under H_0.

μ depends on f and f', if f is the common density or regression function. So we cannot use (4) to find the critical value without first estimating f'.

Consider: $f_\lambda(x)=\lambda e^{-\lambda x}/(1-e^{-3\lambda})$ on $[0,3]$.

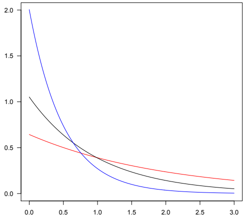


Figure: Truncated exponential densities with λ equal to 0.5 (red), 1

Bootstrapping the critical value

Will bootstrapping from the Grenander estimator work to find the critical value? We try it out!

General procedure: We generate 10,000 sets of 3 samples of size 100, where the first two samples are from a truncated standard exponential density ($\lambda = 1$) on $[0, 3]$ and the third sample from a truncated exponential density with varying parameter λ.

Bootstrap procedure: For each of the original samples we generate 10,000 bootstrap samples from the Grenander estimate $\hat{f}_N$ for the combined original samples. Next we count how many times the test statistics in the original sample exceed the 95% percentile of the test statistics in the 10,000 bootstrap samples. This gives an estimate of the power of the tests.

Verification procedure

1. We first generate 10,000 samples of size 300 from the mixture density:

$$g_\lambda(x) = \frac{2}{3}\,\frac{e^{-x}}{1-e^{-3}} + \frac{1}{3}\,\frac{e^{-\lambda x}}{1-e^{-3\lambda}}, \qquad x \in [0,3],$$

   and compute the test statistics for each sample. Then we determine the 95% percentiles for the values so obtained for the two statistics. This gives the critical values for step 2:

2. We generate 10,000 sets of 3 samples of size 100, where the first two are generated from a standard truncated exponential and the third from an truncated exponential density with parameter λ and count how many times the test statistics exceed the critical values obtained in the first step. This gives estimates of the power functions.

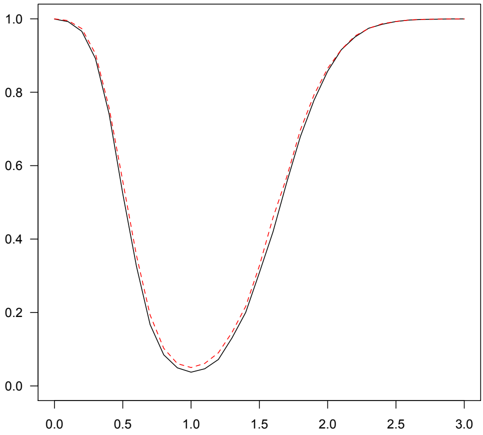


Figure: Estimate of the power of the Durot-Lopuhaä test on the interval $[0, 3]$. Solid: bootstrap estimate from the Grenander estimate; red and dashed: direct estimate of the real power.

Concluding remarks

- Local limit distribution of MLE for interval censoring, case 2, non-separated case, is still only conjectured.
- Local limit of MLE for deconvolution is still only conjectured.
- Local limit of MLE for bivariate current status is still unknown. There are estimators, achieving the $n^{1/3}$-rate.
- Limit of LS estimator for convex regression has been determined, but the structure of the limiting process has not been analytically determined.
- Convex envelope of Brownian motion without drift is very different from convex envelope of Brownian motion with parabolic drift.
- Possibly the bootstrapping of global statistics can be done from the Grenander estimator, as suggested by simulations.

Bibliography I

Fadoua Balabdaoui and Jim Pitman. The distribution of the maximal difference between a Brownian bridge and its concave majorant. *Bernoulli*, 17(1):466–483, 2011.

L. Birgé. Interval censoring: a nonasymptotic point of view. *Math. Methods Statist.*, 8(3):285–298, 1999.

Erhan Çinlar. Sunset over Brownistan. *Stochastic Process. Appl.*, 40(1):45–53, 1992.

P. Groeneboom. Estimating a monotone density. In *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983)*, Wadsworth Statist./Probab. Ser., pages 539–555, Belmont, CA, 1985. Wadsworth.

P. Groeneboom. Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Related Fields*, 81(1):79–109, 1989.

Bibliography II

P. Groeneboom. Nonparametric maximum likelihood estimators for interval censoring and deconvolution. Technical Report 378, Department of Statistics, Stanford University, 1991. URL http://statistics.stanford.edu/~ckirby/reports/1990_1999/reports1991.html.

P. Groeneboom. Lectures on inverse problems. In *Lectures on probability theory and statistics (Saint-Flour, 1994)*, volume 1648 of *Lecture Notes in Math.*, pages 67–164. Springer, Berlin, 1996.

P. Groeneboom. Maximum smoothed likelihood estimators for the interval censoring model. Submitted, 2012.

P. Groeneboom. The bivariate current status model. *Electron. J. Stat.*, 7:1797–1845, 2013.

P. Groeneboom, G. Hooghiemstra, and H.P. Lopuhaä. Asymptotic normality of the L_1 error of the Grenander estimator. *Ann. Statist.*, 27(4):1316–1347, 1999.

Bibliography III

P. Groeneboom, G. Jongbloed, and J.A. Wellner. A canonical process for estimation of convex functions: the "invelope" of integrated Brownian motion $+t^4$. *Ann. Statist.*, 29(6): 1620–1652, 2001b.

P. Groeneboom and T. Ketelaars. A study of two estimators for the interval censoring problem. Electronic Journal of Statistics, 2011, **5**, 1797-1845., 2011.

Piet Groeneboom. The concave majorant of Brownian motion. *Ann. Probab.*, 11(4):1016–1027, 1983.

Piet Groeneboom, Geurt Jongbloed, and Jon A. Wellner. The support reduction algorithm for computing non-parametric function estimates in mixture models. *Scand. J. Statist.*, 35(3): 385–399, 2008.

Piet Groeneboom and Ronald Pyke. Asymptotic normality of statistics based on the convex minorants of empirical distribution functions. *Ann. Probab.*, 11(2):328–345, 1983.

Bibliography IV

Geurt Jongbloed. The iterative convex minorant algorithm for nonparametric estimation. *J. Comput. Graph. Statist.*, 7(3): 310–321, 1998.

Mary C. Meyer. Inference using shape-restricted regression splines. *Ann. Appl. Stat.*, 2(3):1013–1033, 2008.

J. W. Pitman. Remarks on the convex minorant of Brownian motion. In *Seminar on stochastic processes, 1982 (Evanston, Ill., 1982)*, volume 5 of *Progr. Probab. Statist.*, pages 219–227. Birkhäuser Boston, Boston, MA, 1983.

Jim Pitman and Nathan Ross. The greatest convex minorant of Brownian motion, meander, and bridge. *Probab. Theory Related Fields*, 153(3-4):771–807, 2012.

Jim Pitman and Gerónimo Uribe Bravo. The convex minorant of a Lévy process. *Ann. Probab.*, 40(4):1636–1674, 2012.

Erik Sparre Andersen. On the fluctuations of sums of random variables. II. *Math. Scand.*, 2:195–223, 1954.

Bibliography V

J.A. Wellner. Interval censoring, case 2: alternative hypotheses. In *Analysis of censored data (Pune, 1994/1995)*, volume 27 of *IMS Lecture Notes Monogr. Ser.*, pages 271–291. Inst. Math. Statist., Hayward, CA, 1995.