Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 126 (2016) 3854–3864

stochastic processes and their applications

www.elsevier.com/locate/spa

A law of the iterated logarithm for Grenander's estimator

Lutz Dümbgen [a], Jon A. Wellner [b,*], Malcolm Wolff [b]

[a] University of Bern, Institute of Mathematical Statistics and Actuarial Science, Alpeneggstrasse 22, CH-3012 Bern, Switzerland

[b] Department of Statistics, University of Washington, Seattle, WA 98195-4322, United States

Received 12 April 2014; accepted 18 August 2015
Available online 29 April 2016

Abstract

In this note we prove the following law of the iterated logarithm for the Grenander estimator of a monotone decreasing density: If $f(t_0) > 0$, $f'(t_0) < 0$, and f' is continuous in a neighborhood of t_0, then

$$\limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/3} (\widehat{f}_n(t_0) - f(t_0)) = \left|f(t_0) f'(t_0)/2\right|^{1/3} 2M$$

almost surely where

$$M \equiv \sup_{g\in\mathcal{G}} T_g = (3/4)^{1/3} \quad \text{and} \quad T_g \equiv \underset{u}{\operatorname{argmax}}\{g(u) - u^2\};$$

here $\mathcal{G}$ is the two-sided Strassen limit set on $\mathbb{R}$. The proof relies on laws of the iterated logarithm for local empirical processes, Groeneboom's switching relation, and properties of Strassen's limit set analogous to distributional properties of Brownian motion; see Strassen [26].

MSC: 60F15; 60F17; 62E20; 62F12; 62G20

Keywords: Grenander; Monotone density; Law of iterated logarithm; Limit set; Strassen; Switching; Strong invariance theorem; Limsup; Liminf; Local empirical process

* Corresponding author.
E-mail address: jaw@stat.washington.edu (J.A. Wellner).

http://dx.doi.org/10.1016/j.spa.2016.04.012
0304-4149/

1. Introduction: the MLE of a monotone density

Nonparametric estimation of a monotone density was first considered by Grenander [9,10]. Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function F on $[0, \infty)$ having a decreasing density f. Grenander showed that the maximum likelihood estimator $\hat{f}_n$ of f is the (left-) derivative of the least concave majorant of the empirical distribution function $\mathbb{F}_n$ (see Figs. 1 and 2)

$$\widehat{f_n} = \{\text{left derivative of the least concave majorant of } \mathbb{F}_n\}.$$

The asymptotic distribution of $\hat{f}_n(t_0)$ at a fixed point t_0 with $f'(t_0) < 0$ was obtained by Prakasa Rao [22,23], and given a somewhat different proof by Groeneboom [11]; also see [13, sections 3.2 and 3.6]. If $f'(t_0) < 0$ and f' is continuous in a neighborhood of t_0, then

$$n^{1/3}(\widehat{f_n}(t_0) - f(t_0)) \to_d \left|\frac{1}{2} f(t_0) f'(t_0)\right|^{1/3} 2\mathbb{Z}, \tag{1}$$

where

$$\begin{aligned} 2\mathbb{Z} &= \text{slope at 0 of the least concave majorant of } W(t) - t^2 \\ &\stackrel{d}{=} \text{slope at 0 of the greatest convex minorant of } W(t) + t^2 \\ &\stackrel{d}{=} 2 \operatorname*{argmin}_{t\in\mathbb{R}}\{W(t) + t^2\}; \end{aligned} \tag{2}$$

here $\{W(t) : t \in \mathbb{R}\}$ is a two-sided Brownian motion process starting at 0. In fact, the convergence in (1) can be extended to weak convergence of the (local) Grenander process as follows. Let $\{\mathbb{S}_{a,b}(t) : t \in \mathbb{R}\}$ denote the slope process corresponding to the least concave majorant of $X_{a,b}(t) = aW(t) - bt^2$, with $a = \sqrt{f(t_0)}$ and $b = |f'(t_0)|/2$. Then for fixed t_0 with $f'(t_0) < 0$ and f' continuous in a neighborhood of t_0,

$$n^{1/3}(\hat{f}_n(t_0 + n^{-1/3}t) - f(t_0)) \Rightarrow \mathbb{S}_{a,b}(t)$$

in the Skorokhod topology on $D[-K, K]$ for every finite $K > 0$; see e.g. [12,18], and [17]. Groeneboom [12] gives a complete analytic characterization of the limiting distribution $\mathbb{Z}$ and further, the distributional structure of the process $\mathbb{S}$. The distribution of $\mathbb{Z} = \mathbb{S}(0)/2$ has been studied numerically by Groeneboom and Wellner [16] which relies heavily on [11,12]. Balabdaoui and Wellner [4] show that the distribution of $\mathbb{Z}$ is log-concave. Note that there is an "invariance principle" involved here: the centered slope of the least concave majorant of $\mathbb{F}_n$ converges weakly to a constant times the slope of the least concave majorant of $X(t) = W(t) - t^2$. We can regard the slope in this Gaussian limit problem, $2\mathbb{Z}$, as an "estimator" of the slope of the line $2t$ in the Gaussian problem of "estimating" the "canonical" linear function $2t$ in "Gaussian white noise" $dW(t)$ since

$$dX(t) = 2t dt + dW(t).$$

2. A law of the iterated logarithm for the Grenander estimator

Our main goal is to prove the following Law of the Iterated Logarithm (LIL) for the Grenander estimator corresponding to the limiting distribution result in (1).

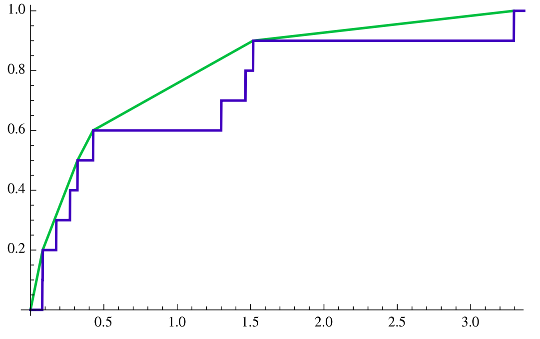

Fig. 1. Empirical distribution and least concave majorant, $n = 10$.

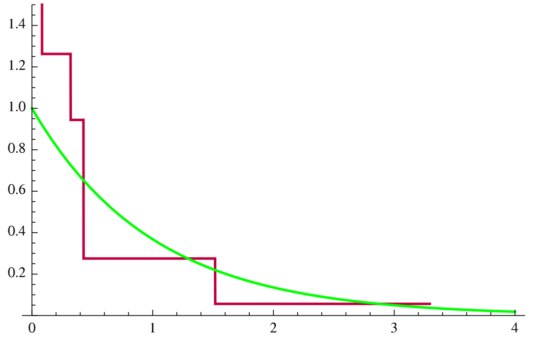

Fig. 2. Grenander estimator and Exp(1) density, $n = 10$.

Theorem 1. *Suppose that* $f(t_0) > 0$, $f_0'(t_0) < 0$ *with* f' *continuous in a neighborhood of* t_0. *Then*

$$\limsup_{n\to\infty} \frac{n^{1/3}(\widehat{f}_n(t_0) - f(t_0))}{(2\log\log n)^{1/3}} = \left|\frac{1}{2} f(t_0) f'(t_0)\right|^{1/3} 2M$$

almost surely where

$$M \equiv \sup_{g\in\mathcal{G}} \operatorname*{argmax}_{t\in\mathbb{R}}\{g(t) - t^2\} = \left(\frac{3}{4}\right)^{1/3};$$

here $\mathcal{G}$ *is the two-sided Strassen limit set on* $\mathbb{R}$ *given by*

$$\mathcal{G} = \left\{ g : \mathbb{R} \to \mathbb{R} \,\middle|\, g(t) = \int_0^t \dot{g}(s)ds,\ t \in \mathbb{R},\ \int_{-\infty}^{\infty} \dot{g}^2(s)ds \le 1 \right\}. \tag{3}$$

Our proof of Theorem 1 will rely on functional laws of the iterated logarithm for the local empirical process established by Mason [20]; see also [5,7,6], and [21]. Along the way we will also prove several lemmas concerning the limit set $\mathcal{G}$.

Proof. We begin the proof of Theorem 1 with a switching argument. Let $b_n \equiv (n^{-1} 2\log\log n)^{1/3}$. Then we want to find a number x_0 such that

$$P(b_n^{-1}(\widehat{f}_n(t_0) - f(t_0)) > x \text{ i.o.}) = \begin{cases} 0, & \text{if } x > x_0, \\ 1, & \text{if } x < x_0. \end{cases} \tag{4}$$

Now we let

$$\widehat{s}_n(a) \equiv \operatorname*{argmax}_s \{\mathbb{F}_n(s) - as\}, \quad a \geq 0, \tag{5}$$

and note that

$$\{\widehat{f}_n(t_0) > a\} = \{\widehat{s}_n(a) > t_0\}, \tag{6}$$

an identity known as the "switching relation" (see e.g. [11], [27, page 296] and [2, Theorem 2.1, page 881]). Thus the event on the left side in (4) can be rewritten via (6) as

$$\left\{\widehat{f}_n(t_0) > f(t_0) + b_n x \text{ i.o.}\right\} = \{\widehat{s}_n(f(t_0) + b_n x) > t_0 \text{ i.o.}\}. \tag{7}$$

But, by letting $s = t_0 + b_n h$ in (5) we see that

$$\widehat{s}_n(f(t_0) + b_n x) - t_0 = b_n \operatorname*{argmax}_h \{\mathbb{F}_n(t_0 + b_n h) - (f(t_0) + b_n x)(t_0 + b_n h)\},$$

and hence the right side of (7) can be rewritten as $\{\widehat{h}_n > 0 \text{ i.o.}\}$ where

$$\begin{aligned} \widehat{h}_n &= \operatorname*{argmax}_h \{\mathbb{F}_n(t_0 + b_n h) - (f(t_0) + b_n x)(t_0 + b_n h)\} \\ &= \operatorname*{argmax}_h \Big\{ b_n^{-2}[\mathbb{F}_n(t_0 + b_n h) - \mathbb{F}_n(t_0) - (F(t_0 + b_n h) - F(t_0))] \\ &\qquad + b_n^{-2}\{F(t_0 + b_n h) - F(t_0) - f(t_0) b_n h\} - xh \Big\}. \end{aligned} \tag{8}$$

The second term on the right side in (8) converges to $f'(t_0)h^2/2$ as $n \to \infty$. To handle the first term we appeal to (a slight extension of) Theorem 2 of [20]; see also [5, Theorem A and Theorem 1.1, pages 1620–1621]: by considering $h \in \mathbb{R}$ and introducing the two-sided version $\mathcal{G}$ of the Strassen limit set given in (3) much as in [28], we see that the sequence of functions

$$\left\{ b_n^{-2}[\mathbb{F}_n(t_0 + b_n h) - \mathbb{F}_n(t_0) - (F(t_0 + b_n h) - F(t_0))] : \ h \in \mathbb{R} \right\}$$

is almost surely relatively compact with limit set

$$\{g(f(t_0)\cdot) : \ g \in \mathcal{G}\}$$

where $\mathcal{G}$ is given by (3).

This is most easily seen as follows: let $\mathbb{G}_n$ be the empirical d.f. of $\xi_1, \ldots, \xi_n$ i.i.d. Uniform(0, 1). As in [5], with $n^{-1}k_n \equiv b_n$ so that $k_n = nb_n = n^{2/3}(2\log\log n)^{1/3} \nearrow \infty$ and $n^{-1}k_n = b_n \searrow 0$, the processes

$$\frac{\xi_n(s)}{\sqrt{2\log\log n}} = \frac{n^{1/2}}{\sqrt{k_n/n}} \frac{\left\{\mathbb{G}_n(F(t_0 + n^{-1}k_n s)) - \mathbb{G}_n(F(t_0)) - (F(t_0 + n^{-1}k_n s) - F(t_0))\right\}}{\sqrt{2\log\log n}}$$

with $s \geq 0$ are almost surely relatively compact with limit set $\mathcal{K}_\infty(c) \equiv \{t \mapsto g(ct) : \ g \in \mathcal{K}_\infty\}$ with $c = f(t_0)$. Here we also note that

$$\frac{n^{1/2}}{\sqrt{k_n/n}\sqrt{2\log\log n}} = \frac{n^{2/3}}{(2\log\log n)^{2/3}} = b_n^{-2}.$$

Thus the processes involved in the argmax in (8) are almost surely relatively compact with limit set

$$\{g(f(t_0)h) + 2^{-1} f'(t_0)h^2 - xh : \ g \in \mathcal{G}\},$$

and by Lemma 1 this set is equal to

$$\left\{ag(h) - bh^2 - xh : \; g \in \mathcal{G}\right\}$$

where $a \equiv \sqrt{f(t_0)}$, and $b = |f'(t_0)|/2$. Thus by Lemma 2, the set of limits for the argmax in (8) equals

$$\left\{(a/b)^{2/3} \operatorname*{argmax}_h \{g(h) - h^2\} - x/(2b) : \; g \in \mathcal{G}\right\}$$

where

$$\left(\frac{a}{b}\right)^{2/3} = \left(\frac{\sqrt{f(t_0)}}{2^{-1}|f'(t_0)|}\right)^{2/3} = \left(\frac{4f(t_0)}{|f'(t_0)|^2}\right)^{1/3}.$$

Hence, with $T_g = \mathrm{argmax}_h\{g(h) - h^2\}$,

$$\begin{aligned}
\left\{\widehat{h}_n > 0 \text{ i.o.}\right\} &\overset{a.s.}{=} \left\{\left(\frac{a}{b}\right)^{2/3} \sup_{g \in \mathcal{G}} T_g > \frac{x}{2b}\right\} \\
&= \left\{2b\left(\frac{a}{b}\right)^{2/3} \sup_{g \in \mathcal{G}} T_g > x\right\} \\
&= \emptyset
\end{aligned}$$

if

$$x > x_0 \equiv 2b\left(\frac{a}{b}\right)^{2/3} \sup_{g \in \mathcal{G}} T_g = \left|\frac{1}{2} f(t_0) f'(t_0)\right|^{1/3} 2 \sup_{g \in \mathcal{G}} T_g.$$

It remains only to show that $\sup_{g \in \mathcal{G}} T_g = (3/4)^{1/3}$. This follows from Lemma 3 in Section 4. □

Lemma 1. *Let $c > 0$ and $d \in \mathbb{R}$. Then*

$$\{t \mapsto g(ct + d) - g(d) : \; g \in \mathcal{G}\} = \sqrt{c}\mathcal{G}.$$

Proof. If $g \in \mathcal{G}$, then

$$\begin{aligned}
g(ct + d) - g(d) &= \int_d^{ct+d} \dot{g}(s)ds = \int_0^{ct} \dot{g}(v + d)dv = \int_0^t \dot{g}(cu + d)c\,du \\
&= \sqrt{c} \int_0^t \sqrt{c}\dot{g}(cu + d)du \\
&= \sqrt{c}\tilde{g}(t)
\end{aligned}$$

where $\tilde{g} \in \mathcal{G}$ since

$$\int_{-\infty}^{\infty} (\sqrt{c}\dot{g}(cu + d))^2 du = \int_{-\infty}^{\infty} \dot{g}^2(w)dw \le 1.$$

This shows that the set of functions $t \mapsto g(ct + d) - g(d)$, $g \in \mathcal{G}$, is contained in $\sqrt{c}\mathcal{G}$. On the other hand, any function $\tilde{g} \in \mathcal{G}$ with derivative $\dot{\tilde{g}}$ may be written as $\tilde{g}(t) = \int \sqrt{c}\dot{g}(cu+d)du$ with $\dot{g}$ given by $\dot{g}(s) \equiv \sqrt{c^{-1}}\dot{\tilde{g}}(c^{-1}s - c^{-1}d)$ and satisfying $\int_{-\infty}^{\infty} \dot{g}(s)^2 ds = \int_{-\infty}^{\infty} \dot{\tilde{g}}(s)^2 ds \le 1$. □

Lemma 2. *Let α, β be positive constants and $\gamma \in \mathbb{R}$. Then*

$$\Big\{\operatorname*{argmax}_h\{\alpha g(h) - \beta h^2 - \gamma h\} : \ g \in \mathcal{G}\Big\}$$
$$= \Big\{(\alpha/\beta)^{2/3} \operatorname*{argmax}_h\{g(h) - h^2\} - \gamma/(2\beta) : \ g \in \mathcal{G}\Big\}. \tag{9}$$

Proof. Note first that

$$\begin{aligned} M_g &\equiv \operatorname*{argmax}_h \Big\{\alpha g(h) - \beta h^2 - \gamma h\Big\} \\ &= \operatorname*{argmax}_h \Big\{\alpha g(h) - \beta (h + \gamma/(2\beta))^2\Big\} \\ &= \operatorname*{argmax}_h \Big\{g(h) - (\beta/\alpha)(h + \gamma/(2\beta))^2\Big\} \\ &= \operatorname*{argmax}_v \Big\{g(v + d) - (\beta/\alpha)v^2\Big\} + d \end{aligned}$$

with $d := -\gamma/(2\beta)$. Moreover, for any $c > 0$ and

$$\tilde{g}(u) \equiv c^{-1/2}\,(g(cu + d) - g(d))$$

we may write

$$\begin{aligned} M_g &= c \operatorname*{argmax}_u \Big\{g(cu + d) - g(d) - (\beta/\alpha)c^2u^2\Big\} + d \\ &= c \operatorname*{argmax}_u \Big\{c^{1/2}\tilde{g}(u) - (\beta/\alpha)c^2u^2\Big\} + d \\ &= c \operatorname*{argmax}_u \Big\{\tilde{g}(u) - (\beta/\alpha)c^{3/2}u^2\Big\} + d. \end{aligned}$$

In case of $c = (\alpha/\beta)^{2/3}$ we obtain

$$M_g = (\alpha/\beta)^{2/3} \operatorname*{argmax}_u \Big\{\tilde{g}(u) - u^2\Big\} - \gamma/(2\beta).$$

Now the claim follows from Lemma 1, because the set $\{\tilde{g} : g \in \mathcal{G}\}$ equals $\mathcal{G}$. □

3. Some comparisons and connections

As noted in the introduction,

$$2\mathbb{Z} \stackrel{d}{=} \text{slope at zero of the least concave majorant of } W(t) - t^2.$$

This suggests that with $T_g = \operatorname{argmax}_t\{g(t) - t^2\}$ we have

$$\Big\{2 \sup T_g : \ g \in \mathcal{G}\Big\} = \sup\{\text{slope at 0 of the least concave majorant of } g(t) - t^2 : \ g \in \mathcal{G}\}.$$

4. Proof for the variational problem

It is natural to conjecture that $\sup_{g\in\mathcal{G}} T_g = (3/4)^{1/3} \approx 0.90856\ldots$. This is motivated by the asymptotic behavior of Chernoff's density; see [12, Corollary 3.4, page 94]: since the density

$$f_{\mathbb{Z}}(z) \sim \frac{1}{2Ai'(a_1)} 4^{4/3} z \exp\left(-\frac{2}{3}z^3 + 3^{1/3}a_1 z\right)$$

as $z \to \infty$, the tail probability $P(\mathbb{Z} > z)$ satisfies

$$P(\mathbb{Z} > z) \sim \frac{1}{2Ai'(a_1)} 4^{4/3} \frac{1}{z} \exp\left(-\frac{2}{3}z^3\right)$$

as $z \to \infty$ where $a_1 \doteq -2.3381$ is the largest zero of the Airy function Ai and $Ai'(a_1) \doteq 0.7022$. Thus from (1) we expect that

$$\limsup_{n\to\infty} \frac{n^{1/3}(\widehat{f}_n(t_0) - f(t_0))}{((3/2)\log\log n)^{1/3}} = \left|\frac{1}{2} f(t_0) f'(t_0)\right|^{1/3} 2,$$

or, equivalently,

$$\begin{aligned}\limsup_{n\to\infty} \frac{n^{1/3}(\widehat{f}_n(t_0) - f(t_0))}{(2\log\log n)^{1/3}} &= \left|\frac{1}{2} f(t_0) f'(t_0)\right|^{1/3} 2 \cdot \frac{1}{2^{1/3}} \cdot \left(\frac{3}{2}\right)^{1/3} \\ &= \left|\frac{1}{2} f(t_0) f'(t_0)\right|^{1/3} 2 \cdot \left(\frac{3}{4}\right)^{1/3}.\end{aligned}$$

On the other hand the proof of Theorem 1 above leads to

$$\limsup_{n\to\infty} \frac{n^{1/3}(\widehat{f}_n(t_0) - f(t_0))}{(2\log\log n)^{1/3}} = \left|\frac{1}{2} f(t_0) f'(t_0)\right|^{1/3} 2 \cdot M \quad \text{a.s.}$$

where

$$M \equiv \sup_{g\in\mathcal{G}} \operatorname{argmax}_{t\in\mathbb{R}}\{g(t) - t^2\} \equiv \sup_{g\in\mathcal{G}} T_g.$$

Thus we conjecture that $M = (3/4)^{1/3}$.

Lemma 3. *Let $t_0 > 0$ be an arbitrary positive number and let $\dot{g} \in L_1([0, t_0])$ be an arbitrary function satisfying*

$$\int_0^{t_0} \dot{g}(s)ds - t_0^2 \ge \int_0^t \dot{g}(s)ds - t^2 \quad \textit{for } 0 \le t \le t_0.$$

Then

$$\int_0^{t_0} \dot{g}(u)^2 du \ge \int_0^{t_0} (2u)^2 du = \frac{4t_0^3}{3}.$$

Proof. Let $\dot{g}_0(u) \equiv 2u$. The claimed inequality is trivial if the integral on the left side is infinite, so we may view $\dot{g}$ and $\dot{g}_0$ as elements of the Hilbert space $L_2([0, t_0])$. Then the assumption on $\dot{g}$ may be rewritten as

$$\langle \dot{g} - \dot{g}_0, 1\rangle \ge \langle \dot{g} - \dot{g}_0, 1_{[0,t]}\rangle \quad \text{for } 0 \le t \le t_0.$$

In other words,

$$\langle \dot{g} - \dot{g}_0, 1_{(t,t_0]}\rangle \geq 0 \quad \text{for } 0 \leq t \leq t_0,$$

and this is equivalent to

$$\langle \dot{g} - \dot{g}_0, f\rangle \geq 0$$

for all functions f in the closed convex cone $\mathbb{K}$ generated by the indicator functions $1_{(t,t_0]}$. This is the set of non-negative and non-decreasing functions on $[0, t_0]$. In particular, $\dot{g}_0 \in \mathbb{K}$, so

$$\langle \dot{g} - \dot{g}_0, \dot{g}_0\rangle \geq 0.$$

Together with the Cauchy–Schwarz inequality we obtain

$$0 \leq \langle \dot{g} - \dot{g}_0, \dot{g}_0\rangle = \langle \dot{g}, \dot{g}_0\rangle - \|\dot{g}_0\|^2 \leq \|\dot{g}\|\,\|\dot{g}_0\| - \|\dot{g}_0\|^2, \tag{10}$$

so $\|\dot{g}\| \geq \|\dot{g}_0\|$. This inequality is strict unless $\dot{g} = \lambda\dot{g}_0$ for some $\lambda \in \mathbb{R}$. In this special case (10) yields $0 \leq (\lambda - 1)\|\dot{g}_0\|^2$, so $\lambda \geq 1$ and $\|\dot{g}\| = \lambda\|\dot{g}_0\|$ with equality if, and only if, $\lambda = 1$ and $\dot{g} = \dot{g}_0$. □

Example 1. If we take $f(x) = e^{-x}1_{[0,\infty)}(x)$ and $t_0 = \log 2$, then

$$\left|\frac{1}{2}f(t_0)f'(t_0)\right|^{1/3} \cdot 2 = (2^{-3})^{1/3} \cdot 2 = 1,$$

so the limit superior is just $\sup_{g\in\mathcal{G}} T_g = (3/4)^{1/3}$.

Example 2. If we take $f(x) = (1 + x)^{-2}1_{[0,\infty)}(x)$, then $-f'(x) = 2(1 + x)^{-3}$ and hence with $t_0 = 1$ we have $f(1) = 1/4 = -f'(1)$. Then

$$\left|\frac{1}{2}f(t_0)f'(t_0)\right|^{1/3} \cdot 2 = (2^{-5/3}) \cdot 2 = 2^{-2/3},$$

so the limit superior is $2^{-2/3}\sup_{g\in\mathcal{G}} T_g = (3/16)^{1/3}$.

Example 3. If we take $f(x) = (\sqrt{2} - x)1_{[0,\sqrt{2}]}(x)$ and $t_0 = \sqrt{2} - 1$, then $f(t_0) = 1$, $-f'(t_0) = 1$, and

$$\left|\frac{1}{2}f(t_0)f'(t_0)\right|^{1/3} \cdot 2 = (2^{-1/3}) \cdot 2 = 2^{+2/3},$$

so the limit superior is $2^{+2/3}\sup_{g\in\mathcal{G}} T_g = 2^{2/3}(3/4)^{1/3} = 3^{1/3}$.

5. Some corollaries

Theorem 1 has a number of corollaries and consequences, since the argument in the proof applies to a number of problems involving nonparametric estimation of a monotone function. Our first corollary, however, involves estimation of the mixing distribution G in the mixture representation of a monotone density: that is,

$$f(x) = \int_0^\infty \frac{1}{y}1_{[0,y)}(x)dG(y) = \int_{\{y>x\}} \frac{1}{y}dG(y), \quad x \in (0,\infty) \tag{11}$$

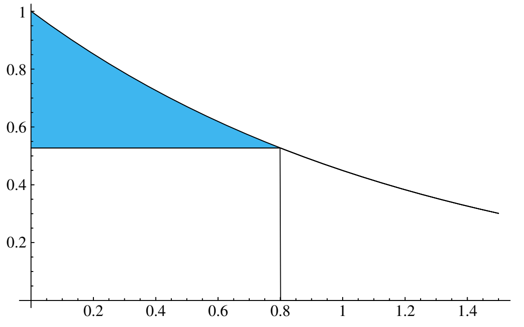

Fig. 3. Graphical view of the inversion formula, monotone density.

for some distribution function G on $(0, \infty)$. This fact apparently goes back at least to Schoenberg [24]; see the introduction of Williamson [29], and Feller [8, page 158]. The relationship (11) implies that the corresponding distribution function F is given by

$$\begin{aligned} F(x) &= \int_0^\infty \frac{x}{y} 1_{[0,y)}(x) dG(y) + \int_0^\infty 1_{[y,\infty)}(x) dG(y) \\ &= x f(x) + G(x), \end{aligned}$$

and this can be "inverted" to yield

$$G(x) = F(x) - x f(x). \tag{12}$$

From Fig. 3 we see that the function on the right side of (12) is non-negative and non-decreasing: the shaded area gives exactly the difference $F(x) - x f(x)$.

The identity (12) implies that the nonparametric maximum likelihood estimator of G is $\widehat{G}_n$ given by

$$\widehat{G}_n(t) = \widehat{F}_n(t) - t \widehat{f}_n(t), \quad \text{for } t \geq 0$$

where $\widehat{F}_n(t) = \int_0^t \widehat{f}_n(x) dx$ is the least concave majorant of $\mathbb{F}_n$ and the MLE of F assuming that f is monotone (and hence F is concave). Thus for $t_0 > 0$ we can write

$$n^{1/3}(\widehat{G}_n(t_0) - G(t_0)) = n^{1/3}(\widehat{F}_n(t_0) - F(t_0)) - t_0 n^{1/3}(\widehat{f}_n(t_0) - f(t_0)).$$

But *Marshall's lemma* yields $\|\widehat{F}_n - F\|_\infty \leq \|\mathbb{F}_n - F\|_\infty$ (see e.g. [19]; [13, Exercise 3.1, page 80]), and since $n^{1/2}\|\mathbb{F}_n - F\|_\infty = O_p(1)$, it follows that $n^{1/3}\|\widehat{F}_n - F\|_\infty = o_p(1)$. Thus if $t_0 > 0$ is a point at which the hypotheses of Theorem 1 hold, then the convergence in (1) implies that

$$n^{1/3}(\widehat{G}_n(t_0) - G(t_0)) \to_d t_0 \left| \frac{1}{2} f(t_0) f'(t_0) \right|^{1/3} 2\mathbb{Z}. \tag{13}$$

Similarly, $\|\widehat{F}_n - F\|_\infty \leq \|\mathbb{F}_n - F\|_\infty$ (Marshall's lemma again) together with Chung's law of the iterated logarithm for $\|\mathbb{F}_n - F\|_\infty$ (see e.g. [25, page 505]), imply that with $b_n \equiv (2 \log\log n)^{1/2}$,

$$\limsup_{n\to\infty} n^{1/2}\|\widehat{F}_n - F\|_\infty / b_n \leq \limsup_{n\to\infty} n^{1/2}\|\mathbb{F}_n - F\|_\infty / b_n = 1/2 \quad \text{a.s.}$$

It follows that if $t_0 > 0$ is a point at which the hypotheses of Theorem 1 hold, then Theorem 1 yields a LIL result for $\widehat{G}_n(t_0)$ as follows:

Corollary 1. *Suppose that $f(t_0) > 0$ and $f'(t_0) < 0$ with f' continuous in a neighborhood of t_0. Then*

$$\limsup_{n\to\infty} \frac{n^{1/3}(\widehat{G}_n(t_0) - G(t_0))}{(2\log\log n)^{1/3}} = t_0 \left| \frac{1}{2} f(t_0) f'(t_0) \right|^{1/3} 2(3/4)^{1/3}$$

almost surely.

6. A further problem

For the problem of estimating a convex decreasing density, Groeneboom et al. [15] described the limiting distribution of the estimator (at a point under a natural curvature condition) in terms of an "invelope" of two-sided integrated Brownian motion plus t^4 which was characterized in [14]. The same distribution has appeared in other nonparametric convex function estimation problems, for example for log-concave density estimation: see [3]. In spite of this description of the limiting distribution for the convex density case in terms of integrated Brownian motion, almost nothing is known concerning a direct analytical description of the limit distribution comparable to the results of Groeneboom [11,12] for Chernoff's distribution. (On the other hand, a preliminary numerical investigation of the distribution is given by Azadbakhsh et al. [1].)

This leads to the following question: can some information concerning the constants involved in the limiting distribution in the convex function case be obtained by establishing LIL results analogous to those established here in the monotone case?

Acknowledgments

The second author owes thanks to Piet Groeneboom for many discussions concerning the Grenander estimator and monotone function estimation more generally. Thanks are due as well to David Mason for references concerning local LIL's for empirical processes.

References

[1] M. Azadbakhsh, H. Jankowski, X. Gao, Computing confidence intervals for log-concave densities, Comput. Statist. Data Anal. 75 (2014) 248–264.
[2] F. Balabdaoui, H. Jankowski, M. Pavlides, A. Seregin, J. Wellner, On the Grenander estimator at zero, Statist. Sinica 21 (2) (2011) 873–899.
[3] F. Balabdaoui, K. Rufibach, J.A. Wellner, Limit distribution theory for maximum likelihood estimation of a log-concave density, Ann. Statist. 37 (3) (2009) 1299–1331.
[4] F. Balabdaoui, J.A. Wellner, Chernoff's density is log-concave, Bernoulli 20 (1) (2014) 231–244.
[5] P. Deheuvels, D.M. Mason, Functional laws of the iterated logarithm for local empirical processes indexed by sets, Ann. Probab. 22 (3) (1994) 1619–1661.
[6] U. Einmahl, D.M. Mason, Gaussian approximation of local empirical processes indexed by functions, Probab. Theory Related Fields 107 (3) (1997) 283–311.
[7] U. Einmahl, D.M. Mason, Strong approximations to the local empirical process, in: High Dimensional Probability, (Oberwolfach, 1996), in: Progr. Probab., vol. 43, Birkhäuser, Basel, 1998, pp. 75–92.
[8] W. Feller, An Introduction to Probability Theory and Its Applications. Vol. II, second ed., John Wiley & Sons, Inc., New York, London, Sydney, 1971.
[9] U. Grenander, On the theory of mortality measurement I, Skand. Aktuarietidskr. 39 (1956) 70–96.
[10] U. Grenander, On the theory of mortality measurement II, Skand. Aktuarietidskr. 39 (1957) (1956) 125–153.
[11] 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), in: Wadsworth Statist./Probab. Ser., Wadsworth, Belmont, CA, 1985, pp. 539–555.

[12] P. Groeneboom, Brownian motion with a parabolic drift and Airy functions, Probab. Theory Related Fields 81 (1) (1989) 79–109.
[13] P. Groeneboom, G. Jongbloed, Nonparametric Estimation Under Shape Constraints, Cambridge University Press, Cambridge, 2014.
[14] P. Groeneboom, G. Jongbloed, J.A. Wellner, A canonical process for estimation of convex functions: the "invelope" of integrated Brownian motion $+t^4$, Ann. Statist. 29 (6) (2001) 1620–1652.
[15] P. Groeneboom, G. Jongbloed, J.A. Wellner, Estimation of a convex function: characterizations and asymptotic theory, Ann. Statist. 29 (6) (2001) 1653–1698.
[16] P. Groeneboom, J.A. Wellner, Computing Chernoff's distribution, J. Comput. Graph. Statist. 10 (2) (2001) 388–400.
[17] Y. Huang, C.-H. Zhang, Estimating a monotone density from censored observations, Ann. Statist. 22 (3) (1994) 1256–1274.
[18] J. Kim, D. Pollard, Cube root asymptotics, Ann. Statist. 18 (1) (1990) 191–219.
[19] A.W. Marshall, Discussion on Barlow and van Zwet's paper, in: M.L. Puri (Ed.), Nonparametric Techniques in Statistical Inference, Proceedings of the First International Symposium on Nonparametric Techniques Held at Indiana University, June, vol. 1969, Cambridge University Press, London, 1970, pp. 174–176.
[20] D.M. Mason, A strong invariance theorem for the tail empirical process, Ann. Inst. H. Poincaré Probab. Statist. 24 (4) (1988) 491–506.
[21] D.M. Mason, A uniform functional law of the logarithm for the local empirical process, Ann. Probab. 32 (2) (2004) 1391–1418.
[22] B.L.S. Prakasa Rao, Estimation of a unimodal density, Sankhyā Ser. A 31 (1969) 23–36.
[23] B.L.S. Prakasa Rao, Estimation for distributions with monotone failure rate, Ann. Math. Statist. 41 (1970) 507–519.
[24] I.J. Schoenberg, On integral representations of completely monotone and related functions (abstract), Bull. Amer. Math. Soc. 47 (1941) 208.
[25] G.R. Shorack, J.A. Wellner, Empirical Processes with Applications to Statistics, in: Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986.
[26] V. Strassen, An invariance principle for the law of the iterated logarithm, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 3 (1964) (1964) 211–226.
[27] A.W. van der Vaart, J.A. Wellner, Weak Convergence and Empirical Processes with Applications to Statistics, in: Springer Series in Statistics, Springer-Verlag, New York, 1996.
[28] M.J. Wichura, Functional laws of the iterated logarithm for the partial sums of I.I.D. random variables in the domain of attraction of a completely asymmetric stable law, Ann. Probab. 2 (1974) 1108–1138.
[29] R.E. Williamson, Multiply monotone functions and their Laplace transforms, Duke Math. J. 23 (1956) 189–207.