Asymptotic normality of the L_k-error of the Grenander estimator

Vladimir N. Kulikov[†]
Hendrik P. Lopuhaä[‡]

01.07.2004

Abstract: We investigate the limit behavior of the L_k-distance between a decreasing density f and its nonparametric maximum likelihood estimator $\hat{f}_n$ for $k \geq 1$. Due to the inconsistency of $\hat{f}_n$ at zero, the case $k = 2.5$ turns out to be some kind of transition point. We extent asymptotic normality of the L_1-distance to the L_k-distance for $1 \leq k < 2.5$, and obtain the analogous limiting result for a modification of the L_k-distance for $k \geq 2.5$. Since the L_1-distance is the area between f and $\hat{f}_n$, which is also the area between the inverse g of f and the more tractable inverse U_n of $\hat{f}_n$, the problem can be reduced immediately to deriving asymptotic normality of the L_1-distance between U_n and g. Although we loose this easy correspondence for $k > 1$, we show that the L_k-distance between f and $\hat{f}_n$ is asymptotically equivalent to the L_k-distance between U_n and g.

[†]EURANDOM, P.O. Box 513 - 5600 MB Eindhoven, the Netherlands, kulikov@eurandom.nl
[‡]Delft University of Technology, Mekelweg 4 2628 CD Delft, the Netherlands, h.p.lopuhaa@its.tudelft.nl

1 Introduction

Let f be a non-increasing density with compact support. Without loss of generality, assume this to be the interval $[0,1]$. The non-parametric maximum likelihood estimator $\hat{f}_n$ for f has been discovered by GRENANDER (1956). It is defined as the left derivative of the least concave majorant (LCM) of the empirical distribution function F_n constructed from a sample $X_1,\dots,X_n$ from f. PRAKASA RAO (1969) obtained the earliest result on the asymptotic pointwise behavior of the Grenander estimator. One immediately striking feature of this result is that the rate of convergence is of the same order as the rate of convergence of histogram estimators, and that the asymptotic distribution is not normal. It took much longer to develop distributional theory for global measures of performance for this estimator. The first distributional result for a global measure of deviation was the convergence to a normal distribution of the L_1-error mentioned in GROENEBOOM (1985) (see GROENEBOOM, HOOGHIEMSTRA AND LOPUHAÄ (1999) for a rigorous proof). A similar result in the regression setting has been obtained by DUROT (2000).

In this paper we extend the result for the L_1-error to the L_k-error, for $k \geq 1$. We will follow the same approach as in GROENEBOOM, HOOGHIEMSTRA AND LOPUHAÄ (1999), which, instead of comparing $\hat{f}_n$ to f, compared both inverses. The corresponding L_1-errors are the same, since they represent the area between the graphs of $\hat{f}_n$ and f and the area between the graphs of the inverses. Clearly, for $k > 1$ we no longer have such an easy correspondence between the two L_k-errors. Nevertheless, we will show that the L_k-error between $\hat{f}_n$ and f can still be approximated by a scaled version of the L_k-error between the two inverses and that this scaled version is asymptotically normal.

The main reason to do a preliminary inversion step, is that we use results from GROENEBOOM, HOOGHIEMSTRA AND LOPUHAÄ (1999) on the inverse process. But apart from this, we believe that working with $\hat{f}_n$ directly will not make life easier. For $a \in [f(1), f(0)]$, the (left continuous) inverse of $\hat{f}_n$ is $U_n(a) = \sup\{x \in [0,1] : \hat{f}_n(x) \geq a\}$. Since $\hat{f}_n(x)$ is the left continuous slope of the LCM of F_n at the point x, a simple picture shows that it has the following more useful representation

$$U_n(a) = \underset{x\in[0,1]}{\operatorname{argmax}}\{F_n(x) - ax\}. \tag{1.1}$$

Here the argmax function is the supremum of the times at which the maximum is attained. Since $U_n(a)$ can be seen as the x-coordinate of the point that is touched first when dropping a line with slope a on F_n, with probability one, $\hat{f}_n(x) \leq a$ if and only if $U_n(a) \leq x$. Asymptotic normality of the L_k error relies on embedding the process $F_n(x) - ax$ into a Brownian motion with drift. The fact that the difference between $F_n(x) - ax$ and the limit process is small, directly implies that the difference of the locations of their maxima is small. However, it does not necessarily imply that the difference of the slopes of the LCM's of both processes is small. Similarly, convergence in distribution of suitably scaled finite dimensional projections of U_n follows immediately from distributional convergence of F_n after suitable scaling, and an argmax type of continuous mapping theorem (see for instance KIM AND POLLARD (1990)). When working with $\hat{f}_n$ directly, similar to Lemma 4.1 in PRAKASA RAO (1969), one needs to bound the probability that the LCM of a gaussian approximation of F_n on $[0,1]$ differs from the one restricted to a shrinking interval, which is somewhat technical and tedious.

Another important difference between the case $k > 1$ and the case $k = 1$, is the fact that for large k, the inconsistency of $\hat{f}_n$ at zero, as shown by WOODROOFE AND SUN (1993), starts to dominate the behavior of the L_k-error. By using results from KULIKOV AND LOPUHAÄ (2002) on the behavior of $\hat{f}_n$ near the boundaries of the support of f, we will show that for $1 \le k < 2.5$ the L_k-error between $\hat{f}_n$ and f is asymptotically normal. This result can be formulated as follows. Define for $c \in \mathbb{R}$,

$$V(c) = \underset{t \in \mathbb{R}}{\text{argmax}}\{W(t) - (t-c)^2\}, \quad (1.2)$$

$$\xi(c) = V(c) - c, \quad (1.3)$$

where $\{W(t) : -\infty < t < \infty\}$ denotes standard two-sided Brownian motion on $\mathbb{R}$ originating from zero (i.e. $W(0) = 0$), and

Theorem 1.1 (Main theorem). *Let f be a decreasing density on $[0,1]$, satisfying:*

(A1) $0 < f(1) \le f(y) \le f(x) \le f(0) < \infty$, *for* $0 \le x \le y \le 1$;

(A2) f *is twice continuously differentiable;*

(A3) $\inf_{x \in (0,1)} |f'(x)| > 0$.

Then for $1 \le k < 2.5$, with $\mu_k = \left\{ E|V(0)|^k \int_0^1 (4f(x)|f'(x)|)^{k/3}\, dx \right\}^{1/k}$, the random variable

$$n^{1/6} \left\{ n^{1/3} \left(\int_0^1 |\hat{f}_n(x) - f(x)|^k\, dx \right)^{1/k} - \mu_k \right\}$$

converges in distribution to a normal random variable with zero mean and variance

$$\frac{\displaystyle\int_0^1 f(x)^{(2k+1)/3} |f'(x)|^{(2k-2)/3}\, dx}{k^2 \left(E|V(0)|^k \displaystyle\int_0^1 \left(f(x)|f'(x)|\right)^{k/3}\, dx \right)^{(2k-2)/k}} \cdot 8 \int_0^\infty \text{cov}(|\xi(0)|^k, |\xi(c)|^k) dc.$$

Note that the theorem holds under the same conditions as in GROENEBOOM ET AL. (1999). For $k \ge 2.5$, Theorem 1.1 is no longer true. However, the results from KULIKOV AND LOPUHAÄ (2002) enable us to show that an analogous limiting result still holds for a modification of the L_k-error.

In Section 2 we introduce a Brownian approximation of U_n and derive asymptotic normality of a scaled version of the L_k-distance between U_n and the inverse g of f. In Section 3 we show that on segments $[s,t]$, where the graph of $\hat{f}_n$ does not cross the graph of f, the difference

$$\left| \int_s^t |\hat{f}_n(x) - f(x)|^k\, dx - \int_{f(t)}^{f(s)} \frac{|U_n(a) - g(a)|^k}{|g'(a)|^{k-1}}\, da \right|$$

is of negligible order. Together with the behavior near the boundaries of the support of f, for $1 \le k < 2.5$, we establish asymptotic normality of the L_k-distance between $\hat{f}_n$ and f in Section 4. In Section 5 we investigate the case $k > 2.5$, and prove a result analogous to Theorem 1.1 for a modified L_k-error.

Remark 1.1 With almost no additional effort one can establish asymptotic normality of a weighted L_k-error $n^{k/3} \int_0^1 |\hat{f}_n(t) - f(t)|^k w(t)\, dt$, where w is continuous differentiable on $[0,1]$. This may be of interest when one wants to use weights proportional to negative powers of the limiting standard deviation $(\frac{1}{2} f(t)|f'(t)|)^{1/3}$ of $\hat{f}_n(t)$. Moreover, when w is estimated at a sufficiently fast rate, one may also replace w by its estimate in the above integral. Similar results are in KULIKOV AND LOPUHAÄ (2004) on a weighted L_k-error.

2 Brownian approximation

In this section we will derive asymptotic normality of the L_k-error of the inverse process of the Grenander estimator. For this we follow the same line of reasoning as in Sections 3 and 4 in Groeneboom et al.(1999). Therefore, we only mention the main steps and transfer all proofs to the appendix.

Let E_n denote the empirical process $\sqrt{n}(F_n - F)$. For $n \geq 1$, let B_n be versions of the Brownian bridge constructed on the same probability space as the uniform empirical process $E_n \circ F^{-1}$ via the Hungarian embedding, and define versions W_n of Brownian motion by

$$W_n(t) = B_n(t) + \xi_n t, \quad t \in [0,1], \tag{2.1}$$

where ξ_n is a standard normal random variable, independent of B_n. For fixed $a \in (f(1), f(0))$ and $J = E, B, W$, define

$$V_n^J(a) = \operatorname*{argmax}_t \left\{ X_n^J(a,t) + n^{2/3} \left[F(g(a) + n^{-1/3}t) - F(g(a)) - n^{-1/3}at \right] \right\}, \tag{2.2}$$

where

$$\begin{aligned} X_n^E(a,t) &= n^{1/6} \left\{ E_n(g(a) + n^{-1/3}t) - E_n(g(a)) \right\}, \\ X_n^B(a,t) &= n^{1/6} \left\{ B_n(F(g(a) + n^{-1/3}t)) - B_n(F(g(a))) \right\}, \\ X_n^W(a,t) &= n^{1/6} \left\{ W_n(F(g(a) + n^{-1/3}t)) - W_n(F(g(a))) \right\}. \end{aligned} \tag{2.3}$$

One can easily check that $V_n^E(a) = n^{1/3}\{U_n(a) - g(a)\}$. A graphical interpretation and basic properties of V_n^J are provided in Groeneboom et al.(1999). For n tending to infinity, properly scaled versions of V_n^J will behave as $\xi(c)$ defined in (1.3).

As a first step we prove asymptotic normality for a Brownian version of the L_k-distance between U_n and g. This is an extension of Theorem 4.1 in Groeneboom et al.(1999).

Theorem 2.1 *Let V_n^W be defined as in (2.2) and ξ by (1.3). Then for $k \geq 1$,*

$$n^{1/6} \int_{f(1)}^{f(0)} \frac{|V_n^W(a)|^k - E|V_n^W(a)|^k}{|g'(a)|^{k-1}} \, da$$

converges in distribution to a normal random variable with zero mean and variance

$$\sigma^2 = 2 \int_0^1 (4f(x))^{(2k+1)/3} |f'(x)|^{(2k-2)/3} \, dx \int_0^\infty \operatorname{cov}(|\xi(0)|^k, |\xi(c)|^k) \, dc.$$

The next lemma shows that the limiting expectation in Theorem 2.1 is equal to

$$\mu_k = \left\{ E|V(0)|^k \int_0^1 \left(4f(x)|f'(x)|\right)^{k/3} dx \right\}^{1/k}. \tag{2.4}$$

Lemma 2.1 *Let V_n^W be defined by (2.2) and let μ_k be defined by (2.4). Then for $k \geq 1$,*

$$\lim_{n\to\infty} n^{1/6} \left\{ \int_{f(1)}^{f(0)} \frac{E|V_n^W(a)|^k}{|g'(a)|^{k-1}} \, da - \mu_k^k \right\} = 0.$$

The next step is to transfer the result of Theorem 2.1 to the L_k-error of V_n^E. This can be done by means of the following lemma.

Lemma 2.2 *For $J = E, B, W$, let V_n^J be defined as in (2.2). Then for $k \geq 1$, we have*

$$n^{1/6} \int_{f(1)}^{f(0)} \left(|V_n^B(a)|^k - |V_n^W(a)|^k \right) da = o_p(1),$$

and

$$\int_{f(1)}^{f(0)} \left| |V_n^E(a)|^k - |V_n^B(a)|^k \right| da = \mathcal{O}_p(n^{-1/3}(\log n)^{k+2}).$$

From Theorem 2.1 and Lemmas 2.1 and 2.2 we immediately have the following corollary.

Corollary 2.1 *Let U_n be defined by (1.1) and let μ_k be defined by (2.4). Then for $k \geq 1$,*

$$n^{1/6} \left(n^{k/3} \int_{f(1)}^{f(0)} \frac{|U_n(a) - g(a)|^k}{|g'(a)|^{k-1}} \, da - \mu_k^k \right)$$

converges in distribution to a normal random variable with zero mean and variance σ^2 defined in Theorem 2.1.

3 Relating both L_k-errors

When $k = 1$, the L_k-error has an easy interpretation as the area between two graphs. In that case $\int |U_n(a) - g(a)| \, da$ is the same as $\int |\hat{f}_n(x) - f(x)| \, dx$, up to some boundaries effects. This is precisely Corollary 2.1 in Groeneboom et al.(1999). In this section we show that a similar approximation holds for $\int_s^t |\hat{f}_n(x) - f(x)|^k \, dx$ on segments $[s, t]$, where the graphs of $\hat{f}_n$ and f do not intersect. In order to avoid boundary problems, we will apply this approximation in subsequent sections to a suitable cut-off version $\tilde{f}_n$ of $\hat{f}_n$.

Lemma 3.1 *Let $\tilde{f}_n$ be a piecewise constant left-continuous non-increasing function on $[0, 1]$ with a finite number of jumps. Suppose that $f(1) \leq \tilde{f}_n \leq f(0)$, and define its inverse function by*

$$\tilde{U}_n(a) = \sup \left\{ x \in [0, 1] : \tilde{f}_n(x) \geq a \right\},$$

for $a \in [f(1), f(0)]$. Suppose that $[s, t] \subseteq [0, 1]$, such that one of the following situations applies:

1. *$\tilde{f}_n(x) \geq f(x)$, for $x \in (s, t)$, such that $\tilde{f}_n(s) = f(s)$ and $\tilde{f}_n(t+) \leq f(t)$,*
2. *$\tilde{f}_n(x) \leq f(x)$, for $x \in (s, t)$, such that $\tilde{f}_n(t) = f(t)$ and $\tilde{f}_n(s) \geq f(s)$.*

If

$$\sup_{x \in [s,t]} \left| \tilde{f}_n(x) - f(x) \right| < \frac{(\inf_{x \in [0,1]} |f'(x)|)^2}{2 \sup_{x \in [0,1]} |f''(x)|}, \tag{3.1}$$

then for $k \geq 1$,

$$\left| \int_s^t \left| \tilde{f}_n(x) - f(x) \right|^k dx - \int_{f(t)}^{f(s)} \frac{\left| \tilde{U}_n(a) - g(a) \right|^k}{\left| g'(a) \right|^{k-1}} da \right| \leq C \int_{f(t)}^{f(s)} \frac{\left| \tilde{U}_n(a) - g(a) \right|^{k+1}}{|g'(a)|^k} \, da,$$

where $C > 0$ only depends on f and k.

Proof: Let us first consider case 1. Let $\tilde{f}_n$ have m points of jump on (s,t). Denote them in increasing order by $\xi_1 < \cdots < \xi_m$, and write $s = \xi_0$ and $\xi_{m+1} = t$. Denote by $\alpha_1 > \cdots > \alpha_m$ the points of jump of $\tilde{U}_n$ on the interval $(f(t), f(s))$ in decreasing order, and write $f(s) = \alpha_0$ and $\alpha_{m+1} = f(t)$ (see Figure 1). We then have

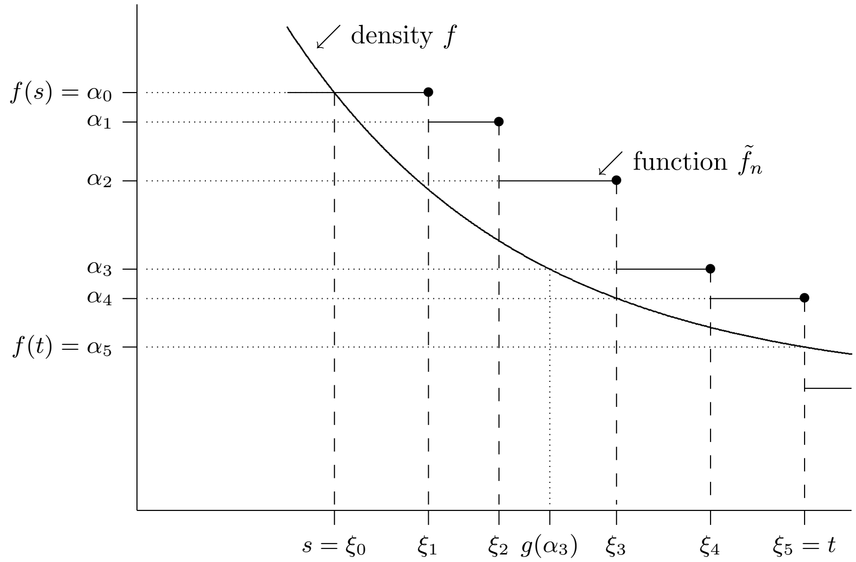

Figure 1: Segment $[s,t]$ where $\tilde{f}_n \geq f$.

$$\int_s^t |\tilde{f}_n(x) - f(x)|^k dx = \sum_{i=0}^{m} \int_{\xi_i}^{\xi_{i+1}} |\tilde{f}_n(\xi_{i+1}) - f(x)|^k \, dx.$$

Apply Taylor expansion to f in the point $g(\alpha_i)$ for each term, and note that $\tilde{f}_n(\xi_{i+1}) = \alpha_i$. Then, if we abbreviate $g_i = g(\alpha_i)$, for $i = 0, 1, \ldots, m$, we can write the right hand side as

$$\sum_{i=0}^{m} \int_{\xi_i}^{\xi_{i+1}} |f'(g_i)|^k (x - g_i)^k \left| 1 + \frac{f''(\theta_i)}{2f'(g_i)}(x - g_i) \right|^k \, dx,$$

for some θ_i between x and g_i, also using the fact that $g_i < \xi_i < x \leq \xi_{i+1}$. Due to condition (3.1) and the fact that $\tilde{f}_n(\xi_{i+1}) = \tilde{f}_n(x)$, for $x \in (\xi_i, \xi_{i+1}]$, we have that

$$\left| \frac{f''(\theta_i)}{f'(g_i)}(x - g_i) \right| \leq \frac{\sup |f''|}{\inf |f'|} \frac{|f(x) - f(g_i)|}{\inf |f'|} \leq \frac{\sup |f''|}{(\inf |f'|)^2} \left| f(x) - \tilde{f}_n(x) \right| \leq \frac{1}{2}. \qquad (3.2)$$

Hence for $x \in (\xi_i, \xi_{i+1}]$,

$$\left| 1 + \frac{f''(\theta_i)(x - g_i)}{2f'(g_i)} \right|^k \leq 1 + \frac{|f''(\theta_i)|(x - g_i)}{2|f'(g_i)|} \sup_{z \in [\frac{1}{2}, \frac{3}{2}]} k z^{k-1} \leq 1 + C_1(x - g_i),$$

where $C_1 = \frac{\sup |f''|}{2 \inf |f'|} k \left(\frac{3}{2}\right)^{k-1}$. Similarly,

$$\left| 1 + \frac{f''(\theta_i)(x - g_i)}{2f'(g_i)} \right|^k \geq 1 - C_1(x - g_i).$$

Therefore we obtain the following inequality

$$\left|\int_s^t |\tilde{f}_n(x) - f(x)|^k\,dx - \sum_{i=0}^m \int_{\xi_i}^{\xi_{i+1}} |f'(g_i)|^k (x-g_i)^k\,dx\right| \leq C_1 \sum_{i=0}^m \int_{\xi_i}^{\xi_{i+1}} (x-g_i)^{k+1}\,dx.$$

After integration, we can rewrite this inequality in the following way:

$$\begin{aligned}&\left|\int_s^t |\tilde{f}_n(x) - f(x)|^k\,dx - \frac{1}{k+1}\sum_{i=0}^m |f'(g_i)|^k \left\{(\xi_{i+1}-g_i)^{k+1} - (\xi_i - g_i)^{k+1}\right\}\right| \\ &\quad \leq \frac{C_1}{k+2}\sum_{i=0}^m \left\{(\xi_{i+1}-g_i)^{k+2} - (\xi_i - g_i)^{k+2}\right\}. \qquad (3.3)\end{aligned}$$

Next, let us consider the corresponding integral for the inverse $\tilde{U}_n$:

$$\int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^k}{|g'(a)|^{k-1}}\,da = \sum_{i=0}^m \int_{\alpha_{i+1}}^{\alpha_i} \frac{|\xi_{i+1} - g(a)|^k}{|g'(a)|^{k-1}}\,da = \sum_{i=0}^m \int_{g_i}^{g_{i+1}} (\xi_{i+1}-x)^k |f'(x)|^k\,dx,$$

now using that $g_i < x < g_{i+1} < \xi_{i+1}$. Apply Taylor expansion to f' in the point g_i. For the right hand side, we then obtain

$$\sum_{i=0}^m \int_{g_i}^{g_{i+1}} (\xi_{i+1}-x)^k |f'(g_i) + f''(\theta_i)(x-g_i)|^k\,dx,$$

for some θ_i between x and g_i. Using (3.2), by means of the same arguments as above we get the following inequality:

$$\begin{aligned}&\left|\int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^k}{|g'(a)|^{k-1}}\,da - \sum_{i=0}^m \int_{g_i}^{g_{i+1}} |f'(g_i)|^k (\xi_{i+1}-x)^k\,dx\right| \\ &\quad \leq C_1 \sum_{i=0}^m \int_{g_i}^{g_{i+1}} (\xi_{i+1}-x)^k (x-g_i)\,dx. \qquad (3.4)\end{aligned}$$

Since $g_i < x < g_{i+1} < \xi_{i+1}$, for each term on the right hand side of (3.4), we have that

$$\begin{aligned}\int_{g_i}^{g_{i+1}} (\xi_{i+1}-x)^k (x-g_i)\,dx &\leq (\xi_{i+1}-g_i)\int_{g_i}^{g_{i+1}} (\xi_{i+1}-x)^k\,dx \\ &= \frac{1}{k+1}\left\{(\xi_{i+1}-g_i)^{k+2} - (\xi_{i+1}-g_{i+1})^{k+1}(\xi_{i+1}-g_i)\right\} \\ &\leq \frac{1}{k+1}\left\{(\xi_{i+1}-g_i)^{k+2} - (\xi_{i+1}-g_{i+1})^{k+2}\right\}.\end{aligned}$$

Hence from (3.4) we find that

$$\begin{aligned}&\left|\int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^k}{|g'(a)|^{k-1}}\,da - \frac{1}{k+1}\sum_{i=0}^m |f'(g_i)|^k \left\{(\xi_{i+1}-g_i)^{k+1} - (\xi_{i+1}-g_{i+1})^{k+1}\right\}\right| \\ &\quad \leq \frac{C_1}{k+1}\sum_{i=0}^m \left\{(\xi_{i+1}-g_i)^{k+2} - (\xi_{i+1}-g_{i+1})^{k+2}\right\}. \qquad (3.5)\end{aligned}$$

For the third integral in the statement of the lemma, similarly as before, we can write

$$\int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^{k+1}}{|g'(a)|^k}\, da = \sum_{i=0}^{m} \int_{g_i}^{g_{i+1}} |f'(g_i)|^{k+1} (\xi_{i+1} - x)^{k+1} \left| 1 + \frac{f''(\theta)}{f'(g_i)}(x - g_i) \right|^{k+1}.$$

According to (3.2) we have that for $x \in (g_i, g_{i+1})$,

$$\left| 1 + \frac{f''(\theta)}{f'(g_i)}(x - g_i) \right| \geq \frac{1}{2},$$

so that after integration we obtain

$$\int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^{k+1}}{|g'(a)|^k}\, da \geq \frac{C_2}{k+2} \sum_{i=0}^{m} \left\{ (\xi_{i+1} - g_i)^{k+2} - (\xi_{i+1} - g_{i+1})^{k+2} \right\}, \tag{3.6}$$

where $C_2 = \left(\frac{1}{2}\right)^{k+1} \inf |f'|^{k+1}$.

Now, let us define Δ as the difference between the first two integrals:

$$\Delta \stackrel{\text{def}}{=} \int_s^t |\tilde{f}_n(x) - f(x)|^k \, dx - \int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^k}{|g'(a)|^{k-1}} da.$$

By (3.3) and (3.5) and the fact that $\xi_0 = g_0$ and $\xi_{m+1} = g_{m+1}$, we find that

$$\begin{aligned} |\Delta| \leq\ & D \sum_{i=0}^{m} (\xi_{i+1} - g_{i+1})^{k+1} \left| |f'(g_i)|^k - |f'(g_{i+1})|^k \right| \\ & + D \sum_{i=0}^{m} \left\{ (\xi_{i+1} - g_i)^{k+2} - (\xi_i - g_i)^{k+2} \right\} \\ & + D \sum_{i=0}^{m} \left\{ (\xi_{i+1} - g_i)^{k+2} - (\xi_{i+1} - g_{i+1})^{k+2} \right\}, \end{aligned} \tag{3.7}$$

where D is some positive constant that depends only on the function f and k. By a Taylor expansion, the first term on the right hand side of (3.7) can be bounded by

$$\begin{aligned} & D \sum_{i=0}^{m} (\xi_{i+1} - g_{i+1})^{k+1} \left| |f'(g_i)|^k - |f'(g_i) + f''(\theta_i)(g_{i+1} - g_i)|^k \right| \\ & \leq D \sum_{i=0}^{m} (\xi_{i+1} - g_{i+1})^{k+1} |f'(g_i)|^k \left| 1 - \left| 1 + \frac{f''(\theta_i)(g_{i+1} - g_i)}{f'(g_i)} \right|^k \right| \\ & \leq D \sum_{i=0}^{m} (\xi_{i+1} - g_{i+1})^{k+1} (g_{i+1} - g_i) \sup |f'|^k \frac{\sup |f''|}{\inf |f'|} \sup_{x \in [\frac{1}{2}, \frac{3}{2}]} k z^{k-1} \\ & \leq C_3 \sum_{i=0}^{m} (\xi_{i+1} - g_{i+1})^{k+1} (g_{i+1} - g_i), \end{aligned}$$

with C_3 only depending on f and k, where we also use (3.2) and the fact that according to (3.1), we have that $(g_{i+1} - g_i) \sup |f''| / \inf |f'| < \frac{1}{2}$. Since $g_i < g_{i+1} < \xi_{i+1}$, this means that the first term on the right hand side of (3.7) can be bounded by

$$\begin{aligned} C_3\sum_{i=0}^{m}(\xi_{i+1}-g_{i+1})^{k+1}(g_{i+1}-g_i) &\le C_3\sum_{i=0}^{m}\left\{(\xi_{i+1}-g_i)-(\xi_{i+1}-g_{i+1})\right\}(\xi_{i+1}-g_{i+1})^{k+1} \\ &\le C_3\sum_{i=0}^{m}\left\{(\xi_{i+1}-g_i)^{k+2}-(\xi_{i+1}-g_{i+1})^{k+2}\right\}. \end{aligned}$$

Because $\xi_0 = g_0$ and $\xi_{m+1} = g_{m+1}$, for the second term on the right hand side of (3.7), we have that

$$\sum_{i=0}^{m}\left\{(\xi_{i+1}-g_i)^{k+2}-(\xi_i-g_i)^{k+2}\right\} = \sum_{i=0}^{m}\left\{(\xi_{i+1}-g_i)^{k+2}-(\xi_{i+1}-g_{i+1})^{k+2}\right\}.$$

Putting things together and using (3.6) we find that

$$|\Delta| \le C_4\sum_{i=0}^{m}\left\{(\xi_{i+1}-g_i)^{k+2}-(\xi_{i+1}-g_{i+1})^{k+2}\right\} \le C_5\int_{f(t)}^{f(s)}\frac{|\tilde{U}_n(a)-g(a)|^{k+1}}{|g'(a)|^k}\,da,$$

where C_5 only depends on f and k. This proves the lemma for case 1.

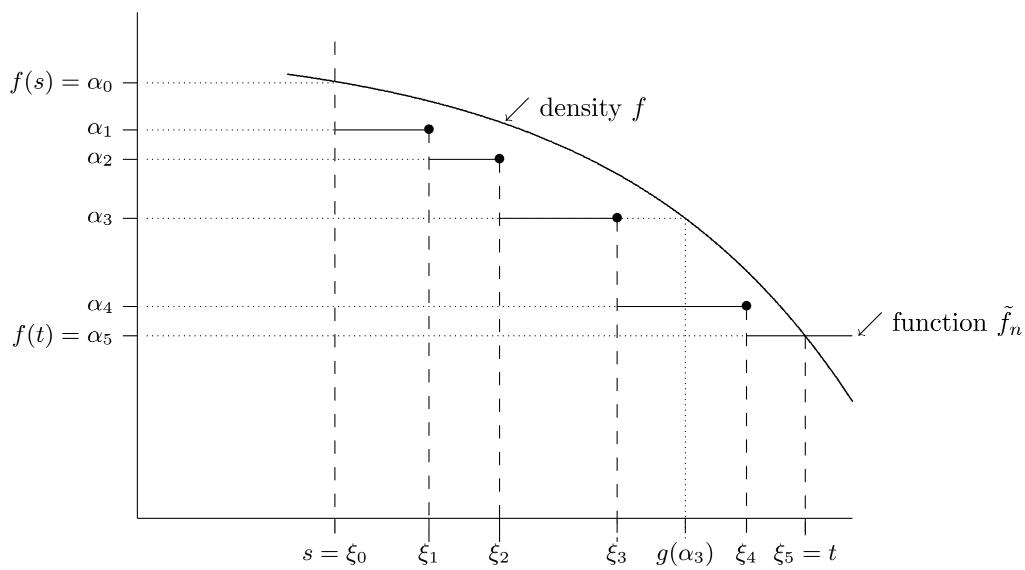

Figure 2: Segment $[s,t]$ where $\tilde{f}_n \le f$.

For case 2, the proof is similar. The main differences are that $\tilde{f}_n(\xi_i) = \alpha_i$ (see Figure 2) and that the Taylor expansions are applied to f in g_{i+1} instead of g_i. Similar to (3.3), now using that $g_{i+1} \ge \xi_{i+1} \ge \xi_i$, we obtain

$$\begin{aligned} &\left|\int_s^t |\tilde{f}_n(x)-f(x)|^k\,dx - \frac{1}{k+1}\sum_{i=0}^{m}|f'(g_{i+1})|^k\left\{(g_{i+1}-\xi_i)^{k+1}-(g_{i+1}-\xi_{i+1})^{k+1}\right\}\right| \\ &\qquad \le \frac{C_1}{k+2}\sum_{i=0}^{m}\left\{(g_{i+1}-\xi_i)^{k+2}-(g_{i+1}-\xi_{i+1})^{k+2}\right\}. \end{aligned} \tag{3.8}$$

Similar to (3.5), now using that $g_{i+1} \geq g_i \geq \xi_i$, we now obtain

$$\left| \int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^k}{|g'(a)|^{k-1}}\, da - \frac{1}{k+1} \sum_{i=0}^{m} |f'(g_{i+1})|^k \left\{ (g_{i+1} - \xi_i)^{k+1} - (g_i - \xi_i)^{k+1} \right\} \right|$$
$$\leq \frac{C_1}{k+1} \sum_{i=0}^{m} \left\{ (g_{i+1} - \xi_i)^{k+2} - (g_i - \xi_i)^{k+2} \right\}, \tag{3.9}$$

and similar to (3.6) we find

$$\int_{f(t)}^{f(s)} \frac{|\tilde{U}_n(a) - g(a)|^{k+1}}{|g'(a)|^k}\, da \geq \frac{C_2}{k+2} \sum_{i=0}^{m} \left\{ (g_{i+1} - \xi_i)^{k+2} - (g_i - \xi_i)^{k+2} \right\}. \tag{3.10}$$

For the difference between the two integrals, again using that $\xi_0 = g_0$ and $\xi_{m+1} = g_{m+1}$, we now find

$$\begin{aligned} |\Delta| \leq\ & D \sum_{i=0}^{m} (g_i - \xi_i)^{k+1} \left| |f'(g_i)|^k - |f'(g_{i+1})|^k \right| \\ & + D \sum_{i=0}^{m} \left\{ (g_{i+1} - \xi_i)^{k+2} - (g_i - \xi_i)^{k+2} \right\} \\ & + D \sum_{i=0}^{m} \left\{ (g_{i+1} - \xi_i)^{k+2} - (g_{i+1} - \xi_{i+1})^{k+2} \right\}, \end{aligned} \tag{3.11}$$

where D is some positive constant that depends only on the function f and k. The first two terms on the right hand side of (3.11) can be bounded similar to the first two terms on the right hand side of (3.7), which yields

$$|\Delta| \leq C_4 \sum_{i=0}^{m} \left\{ (g_{i+1} - \xi_i)^{k+2} - (g_i - \xi_i)^{k+2} \right\} \leq C_5 \int_{f(t)}^{f(s)} \frac{\left| \tilde{U}_n(a) - g(a) \right|^{k+1}}{|g'(a)|^k}\, da,$$

where C_5 only depends on f and k. This proves the lemma for case 2. ■

4 Asymptotic normality of the L_k-error of $\hat{f}_n$

We will apply Lemma 3.1 to the following cut-off version of $\hat{f}_n$:

$$\tilde{f}_n(t) = \begin{cases} f(0) & \text{if } \hat{f}_n(x) \geq f(0), \\ \hat{f}_n(x) & \text{if } f(1) \leq \hat{f}_n(x) < f(0), \\ f(1) & \text{if } \hat{f}_n(x) < f(1). \end{cases} \tag{4.1}$$

The next lemma shows that $\tilde{f}_n$ satisfies condition (3.1) with probability tending to one.

Lemma 4.1 *Define the event*

$$A_n = \left\{ \sup_{x \in [0,1]} \left| \tilde{f}_n(x) - f(x) \right| \leq \frac{\inf_{x \in [0,1]} |f'(x)|^2}{2 \sup_{t \in [0,1]} |f''(x)|} \right\}.$$

Then $P\{A_n^c\} \to 0$.

Proof: It is sufficient to show that $\sup|\tilde{f}_n(x) - f(x)|$ tends to zero. For this we can follow the line of reasoning in Section 5.4 of Groeneboom and Wellner (1992). Similar to their Lemma 5.9 we derive from our Lemma A.1 that for each $a \in (f(1), f(0))$,

$$P\left\{|U_n(a) - g(a)| \geq n^{-1/3}\log n\right\} \leq C_1 \exp\{-C_2(\log n)^3\}.$$

By monotonicity of U_n and the conditions of f, this means that there exists a constant $C_3 > 0$ such that

$$P\left\{\sup_{a\in(f(1),f(0))} |U_n(a) - g(a)| \geq C_3 n^{-1/3}\log n\right\} \leq C_1 \exp\{-\tfrac{1}{2}C_2(\log n)^3\}.$$

This implies that the maximum distance between successive points of jump of $\hat{f}_n$ is of the order $\mathcal{O}(n^{-1/3}\log n)$. Since both $\tilde{f}_n$ and f are monotone and bounded by $f(0)$, this also means that the maximum distance between $\tilde{f}_n$ and f is of the order $\mathcal{O}(n^{-1/3}\log n)$. ∎

The difference between the L_k-errors for $\hat{f}_n$ and $\tilde{f}_n$ is bounded as follows

$$\begin{aligned}&\left|\int_0^1 |\hat{f}_n(x) - f(x)|^k\,dx - \int_0^1 |\tilde{f}_n(x) - f(x)|^k\,dx\right| \\ &\qquad \leq \int_0^{U_n(f(0))} |\hat{f}_n(x) - f(x)|^k\,dx + \int_{U_n(f(1))}^1 |\hat{f}_n(x) - f(x)|^k\,dx.\end{aligned} \tag{4.2}$$

The next lemma shows that the integrals on the right hand side are of negligible order.

Lemma 4.2 *Let U_n be defined in (1.1). Then*

$$\int_0^{U_n(f(0))} |\hat{f}_n(x) - f(x)|^k\,dx = o_p(n^{-(2k+1)/6}),$$

and

$$\int_{U_n(f(1))}^1 |\hat{f}_n(x) - f(x)|^k\,dx = o_p(n^{-(2k+1)/6}).$$

Proof: Consider the first integral, which can be bounded by

$$\begin{aligned}&2^k\int_0^{U_n(f(0))} |\hat{f}_n(x) - f(0)|^k\,dx + 2^k\int_0^{U_n(f(0))} |f(x) - f(0)|^k\,dx \\ &\leq 2^k\int_0^{U_n(f(0))} |\hat{f}_n(x) - f(0)|^k\,dx + \frac{2^k}{k+1}\sup|f'|^k U_n(f(0))^{k+1}.\end{aligned} \tag{4.3}$$

Define the event $B_n = \{U_n(f(0)) \leq n^{-1/3}\log n\}$. Then $U_n(f(0))^{k+1}1_{B_n} = o_p(n^{-(2k+1)/6})$. Moreover, according to Theorem 2.1 in Groeneboom et al.(1999) it follows that $P\{B_n^c\} \to 0$. Since for any $\eta > 0$,

$$P\left(n^{\frac{2k+1}{6}}|U_n(f(0))|^{k+1}1_{B_n^c} > \eta\right) \leq P\{B_n^c\} \to 0,$$

this implies that the second term in (4.3) is of the order $o_p(n^{-(2k+1)/6})$. The first term in (4.3) can be written as

$$2^k\left(\int_0^{U_n(f(0))} |\hat{f}_n(x) - f(0)|^k\,dx\right)1_{B_n} + 2^k\left(\int_0^{U_n(f(0))} |\hat{f}_n(x) - f(0)|^k\,dx\right)1_{B_n^c}, \tag{4.4}$$

where the second integral is of the order $o_p(n^{-(2k+1)/6})$ by the same reasoning as before. To bound the first integral in (4.4), we will construct a suitable sequence $(a_i)_{i=1}^m$, such that the intervals $(0, n^{-a_1}]$ and $(n^{-a_i}, n^{-a_{i+1}}]$, for $i = 1, 2, \dots, m-1$, cover the interval $(0, U_n(f(0))]$, and such that the integrals over these intervals can be bounded appropriately. First of all let

$$1 > a_1 > a_2 > \cdots > a_{m-1} \geq 1/3 > a_m, \tag{4.5}$$

and let $z_0 = 0$ and $z_i = n^{-a_i}$, $i = 1, \dots, m$, so that $0 < z_1 < \cdots < z_{m-1} \leq n^{-1/3} < z_m$. On the event B_n, for n sufficiently large, the intervals $(0, n^{-a_1}]$ and $(n^{-a_i}, n^{-a_{i+1}}]$ cover $(0, U_n(f(0))]$. Hence, when we denote $J_i = [z_i \wedge U_n(f(0)),\, z_{i+1} \wedge U_n(f(0))]$, the first integral in (4.4) can be bounded by

$$\sum_{i=0}^{m-1} \left(\int_{J_i} \left(\hat{f}_n(x) - f(0)\right)^k dx \right) 1_{B_n} \leq \sum_{i=0}^{m-1} (z_{i+1} - z_i) |\hat{f}_n(z_i) - f(0)|^k,$$

using that $\hat{f}_n$ is decreasing and the fact that $J_i \subset (0, U_n(f(0))]$, so that $\hat{f}_n(z_i) - f(0) \geq \hat{f}_n(x) - f(0) \geq 0$, for $x \in J_i$. It remains to show that

$$\sum_{i=0}^{m-1} (z_{i+1} - z_i) |\hat{f}_n(z_i) - f(0)|^k = o_p(n^{-(2k+1)/6}). \tag{4.6}$$

From Woodroofe and Sun (1993), we have that

$$\hat{f}_n(0) \to f(0) \sup_{1 \leq j < \infty} \frac{j}{\Gamma_j}, \tag{4.7}$$

in distribution, where Γ_j are partial sums of standard exponential random variables. Therefore

$$z_1 |\hat{f}_n(0) - f(0)|^k = O_p(n^{-a_1}). \tag{4.8}$$

According to Theorem 3.1 in Kulikov and Lopuhaä (2002), for $1/3 \leq \alpha < 1$,

$$n^{(1-\alpha)/2} \left(\hat{f}_n(n^{-\alpha}) - f(n^{-\alpha}) \right) \to Z, \tag{4.9}$$

in distribution, where Z is a non-degenerate random variable. Since for any $i = 1, \dots, m-1$ we have that $1/3 \leq a_i < 1$, it follows that

$$|\hat{f}_n(z_i) - f(0)| \leq |\hat{f}_n(z_i) - f(z_i)| + \sup |f'| z_i = \mathcal{O}_p(n^{-(1-a_i)/2}) + \mathcal{O}_p(n^{-a_i}) = \mathcal{O}_p(n^{-(1-a_i)/2}).$$

This implies that for $i = 1, \dots, m-1$,

$$(z_{i+1} - z_i) |\hat{f}_n(z_i) - f(0)|^k = \mathcal{O}_p(n^{-a_{i+1} - k(1-a_i)/2}). \tag{4.10}$$

Therefore, if we can construct a sequence (a_i) satisfying (4.5), as well as

$$a_1 > \frac{2k+1}{6}, \tag{4.11}$$

$$a_{i+1} + \frac{k(1-a_i)}{2} > \frac{2k+1}{6}, \quad \text{for all } i = 1, \dots, m-1, \tag{4.12}$$

then (4.6) follows from (4.8) and (4.10). One may take

$$\begin{aligned} a_1 &= \frac{2k+7}{12} \\ a_{i+1} &= \frac{k(a_i-1)}{2} + \frac{2k+3}{8}, \qquad \text{for } i = 1,\ldots,m-1. \end{aligned}$$

Since $k < 2.5$, it immediately follows that (4.11) and (4.12) are satisfied. To show that (4.5) holds, first note that $1 > a_1 > 1/3$, because $k < 2.5$. It remains to show that the described sequence strictly decreases and reaches 1/3 in finitely many steps. As long as $a_i > 1/3$, it follows that

$$a_i - a_{i+1} = \frac{2-k}{2}a_i + \frac{2k-3}{8}.$$

When $k = 2$, this equals 1/8. For $1 \le k < 2$, use that $a_i > 1/3$, to find that $a_i - a_{i+1} > 1/24$, and for $2 \le k < 2.5$, use that $a_i \le a_1 = (2k+1)/7$, to find that $a_i - a_{i+1} \ge (k+1)(2.5-k)/12$. This means that the sequence (a_i) also satisfies (4.5), which proves (4.6).

Similar to (4.3), the second integral can be bounded by

$$2^k \int_{U_n(f(1))}^1 |f(1) - \hat{f}_n(x)|^k \, dx + \frac{2^k}{k+1} \sup |f'|^k (1 - U_n(f(1))^{k+1}.$$

From here the proof is similar as above. We can use the same sequence (a_i) as before, and take $B_n = \{1 - U_n(f(1)) \le n^{-1/3}\log n\}$. If we now define $z_0 = 1$, $z_i = 1 - n^{-a_i}$, for $i = 1, 2, \ldots, m-1$, then similar to the argument above, we are left with considering

$$\left(\int_{U_n(f(1))}^1 |f(1) - \hat{f}_n(x)|^k \, dx \right) 1_{B_n} \le \sum_{i=0}^{m-1} (z_i - z_{i+1}) |f(1) - \hat{f}_n(z_i)|^k. \tag{4.13}$$

The first term is

$$\big(1 - (1 - n^{-a_1})\big) |f(1) - \hat{f}_n(1)|^k = n^{-a_1} f(1)^k.$$

According to Theorem 4.1 in Kulikov and Lopuhaä (2002), for $1/3 < \alpha < 1$,

$$n^{(1-\alpha)/2} \Big(f(1 - n^{-\alpha}) - \hat{f}_n(1 - n^{-\alpha}) \Big) \to \left(f(1) \operatorname*{argmax}_{t \in [0,\infty)} \{W(t) - t^2\} \right)^{1/2}, \tag{4.14}$$

in distribution. Hence, for $i = 1, 2, \ldots, m-1$ each term is of the order $\mathcal{O}_p(n^{-a_{i+1} - k(1-a_i)/2})$. As before the sequence (a_i) chosen above satisfies (4.11) and (4.12), which implies that (4.13) is of the order $o_p(n^{-(2k+1)/6})$. This proves the lemma. ■

We are now able to prove our main result concerning the asymptotic normality of the L_k-error, for $1 \le k < 2.5$.

Proof of Theorem 1.1: First consider the difference

$$\left| \int_0^1 |\hat{f}_n(x) - f(x)|^k \, dx - \int_{f(1)}^{f(0)} \frac{|U_n(a) - g(a)|^k}{|g'(a)|^{k-1}} \, da \right|, \tag{4.15}$$

which can be bounded by

$$\left| \int_0^1 |\hat{f}_n(x) - f(x)|^k \, dx - \int_0^1 |\tilde{f}_n(x) - f(x)|^k \, dx \right| + R_n, \tag{4.16}$$

where

$$R_n = \left| \int_0^1 |\tilde{f}_n(x) - f(x)|^k \, dx - \int_{f(1)}^{f(0)} \frac{|U_n(a) - g(a)|^k}{|g'(t)|^{k-1}} \, da \right|.$$

Let A_n be the event defined in Lemma 4.1, so that $P\{A_n^c\} \to 0$. As in the proof of Lemma 4.2, this means that $R_n 1_{A_n^c} = o_p(n^{-(2k+1)/6})$. Note that on the event A_n, the function $\tilde{f}_n$ satisfies the conditions of Lemma 3.1, and that for any $a \in [f(1), f(0)]$,

$$U_n(a) = \sup\{t \in [0,1] : \; \hat{f}_n(t) > a\} = \sup\{t \in [0,1] : \tilde{f}_n(t) > a\} = \tilde{U}_n(a).$$

Moreover, we can construct a partition $[0, s_1], (s_1, s_2], \ldots, (s_l, 1]$ of $[0,1]$ in such a way that on each element of the partition, $\tilde{f}_n$ satisfies either condition 1 or condition 2 of Lemma 3.1. This means that we can apply Lemma 3.1 to each element of the partition. Putting things together, it follows that $R_n 1_{A_n}$ is bounded from above by

$$C \int_{f(1)}^{f(0)} \frac{|U_n(a) - g(a)|^{k+1}}{|g'(a)|^k} \, da.$$

Corollary 2.1 implies that this integral is of the order $\mathcal{O}_p(n^{-(k+1)/3})$, so that $R_n 1_{A_n} = o_p(n^{-(2k+1)/6})$. Finally, the first difference in (4.16) can be bounded as in (4.2), which means that according to Lemma 4.2 it is of the order $o_p(n^{-(2k+1)/6})$. Together with Corollary 2.1, this implies that

$$n^{1/6} \left(n^{k/3} \int_0^1 |\hat{f}_n(x) - f(x)|^k \, dx - \mu_k^k \right) \to N(0, \sigma^2),$$

where σ^2 is defined in Theorem 2.1. An application of the δ-method then yields that

$$n^{1/6} \left(n^{1/3} \left(\int_0^1 |\hat{f}_n(x) - f(x)|^k dx \right)^{1/k} - \mu_k \right)$$

converges to a normal random variable with mean zero and variance

$$\left\{ \frac{1}{k} \left(\mu_k^k \right)^{1/k-1} \right\}^2 \sigma^2 = \frac{\sigma^2}{k^2 \mu_k^{2k-2}} = \sigma_k^2. \quad \blacksquare$$

5 Asymptotic normality of a modified L_k-error for large k

For large k the inconsistency of $\hat{f}_n$ at zero starts to dominate the behavior of the L_k-error. The following lemma indicates that for $k > 2.5$ the result of Theorem 1.1 does not hold. For $k > 3$, the L_k-error tends to infinity, whereas for $2.5 < k \leq 3$, we are only able to prove that the variance of the integral near zero tends to infinity. In the latter case, it is in principle possible that the behavior of the process $\hat{f}_n - f$ on $[0, z_n]$ depends on the behavior of the process on $[z_n, 1]$ in such a way that the variance of the whole integral stabilizes, but this seems unlikely. The proof of this lemma is transferred to the appendix.

Lemma 5.1 *Let $z_n = 1/(2nf(0))$. Then we have the following.*

(i) If $k > 3$, then $n^{k/3}E\int_0^1 |\hat{f}_n(x) - f(x)|^k\,dx \to \infty$.

(ii) If $k > 2.5$, then $\text{var}\left(n^{(2k+1)/6}\int_0^{z_n} |\hat{f}_n(x) - f(x)|^k\,dx\right) \to \infty$.

Although Lemma 5.1 indicates that for $k > 2.5$ the result Theorem 1.1 will not hold for the usual L_k-error, a similar result can be derived for a modified version. For $k \geq 2.5$ we will consider a modified L_k-error of the form

$$n^{1/6}\left\{n^{1/3}\left(\int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |\hat{f}_n(x) - f(x)|^k\,dx\right)^{1/k} - \mu_k\right\}, \tag{5.1}$$

where μ_k is the constant defined in Theorem 1.1. In this way, for suitable choices of ϵ, we avoid a region where the Grenander estimator is inconsistent in such a way that we are still able to determine its global performance.

We first determine for what values of ϵ we cannot expect asymptotic normality of (5.1). First of all, for $\epsilon > 1$, similar to the proof of Lemma 5.1, it follows that

$$\text{var}\left(n^{(2k+1)/6}\int_{n^{-\epsilon}}^{z_n} |\hat{f}_n(x) - f(x)|^k\,dx\right) \to \infty.$$

For $\epsilon < 1/6$, in view of Lemma 3.1 and the Brownian approximation discussed in Section 2, we have that

$$n^{(2k+1)/6}\int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |\hat{f}_n(x) - f(x)|^k\,dx$$

will behave as

$$n^{1/6}\int_{f(n^{-\epsilon})}^{f(1-n^{-\epsilon})} \frac{n^{k/3}|U_n^W(a) - g(a)|^k}{|g'(a)|^{k-1}}\,da,$$

which, according to Lemma 2.1, is of the order $\mathcal{O}(n^{1/6-\epsilon})$. Hence, we also cannot expect asymptotic normality of (5.1) for $\epsilon < 1/6$. Finally, for $(k-1)/(3k-6) < \epsilon < 1$, a more tedious argument in the same spirit as the proof of Lemma 5.1, yields that

$$\text{var}\left(n^{(2k+1)/6}\int_{n^{-\epsilon}}^{2n^{-\epsilon}} |\hat{f}_n(x) - f(x)|^k\,dx\right) \to \infty.$$

Hence, in order to obtain a proper limit distribution for (5.1) for $k \geq 2.5$, we will choose ϵ between $1/6$ and $(k-1)/(3k-6)$.

To prove a result analogous to Theorem 1.1, we define another cut-off version of the Grenander estimator:

$$f_n^\epsilon(x) = \begin{cases} f(n^{-\epsilon}) & \text{if } \hat{f}_n(x) \geq f(n^{-\epsilon}), \\ \hat{f}_n(x) & \text{if } f(1-n^{-\epsilon}) \leq \hat{f}_n(x) < f(n^{-\epsilon}), \\ f(1-n^{-\epsilon}) & \text{if } \hat{f}_n(x) < f(1-n^{-\epsilon}), \end{cases}$$

and its inverse function

$$U_n^\epsilon(a) = \sup\left\{x \in [n^{-\epsilon}, 1-n^{-\epsilon}] : \hat{f}_n(x) \geq a\right\}, \tag{5.2}$$

for $a \in [f(1-n^{-\epsilon}), f(n^{-\epsilon})]$. The next lemma is the analogue of Lemma 4.1.

Lemma 5.2 *Define the event*

$$A_n^\epsilon = \left\{ \sup_{x\in[0,1]} |f_n^\epsilon(x) - f(x)| \le \frac{\inf_{x\in[0,1]} |f'(x)|^2}{2\sup_{t\in[0,1]} |f''(x)|} \right\}.$$

Then $P\{A_n^\epsilon\} \to 1$.

Proof: It suffices to show that $\sup_{x\in[0,1]} |f_n^\epsilon(x) - f(x)| \to 0$. Using the definition of f_n^ϵ we can bound

$$\sup_{x\in[0,1]} |f_n^\epsilon(x) - f(x)| \le \sup_{x\in[0,1]} \left|f_n^\epsilon(x) - \tilde{f}_n(x)\right| + \sup_{x\in[0,1]} \left|\tilde{f}_n(x) - f(x)\right| \tag{5.3}$$

The first term on the right hand side of (5.3) is smaller than $\sup|f'|n^{-\epsilon}$ which, together with Lemma 4.1, implies that $\sup_{x\in[0,1]} |f_n^\epsilon(x) - f(x)| = o_p(n^{-1/6})$. ■

Similar to (4.2), the difference between the modified L_k-errors for $\hat{f}_n$ and f_n^ϵ is bounded as follows

$$\begin{aligned} &\left| \int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |\hat{f}_n(x) - f(x)|^k \, dx - \int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |f_n^\epsilon(x) - f(x)|^k \, dx \right| \qquad (5.4) \\ &\quad \le \int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(x)|^k \, dx + \int_{U_n^\epsilon(f(1-n^{-\epsilon}))}^{1-n^{-\epsilon}} |\hat{f}_n(x) - f(x)|^k \, dx. \end{aligned}$$

The next lemma is the analogue of Lemma 4.2 and shows that both integrals on the right hand side are of negligible order.

Lemma 5.3 *For* $k \ge 2.5$ *and* $1/6 < \epsilon < (k-1)/(3k-6)$*, let* U_n^ϵ *be defined in (5.2). Then*

$$\int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(x)|^k \, dx = o_p(n^{-(2k+1)/6}),$$

and

$$\int_{U_n^\epsilon(f(1-n^{-\epsilon}))}^{1-n^{-\epsilon}} |\hat{f}_n(x) - f(x)|^k \, dx = o_p(n^{-(2k+1)/6}).$$

Proof: Consider the first integral, then similar to (4.3) we have that

$$\begin{aligned} &2^k \int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(n^{-\epsilon})|^k \, dx + 2^k \int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |f(n^{-\epsilon}) - f(x)|^k \, dx \qquad (5.5) \\ &\quad \le 2^k \int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(n^{-\epsilon})|^k \, dx + \frac{2^k}{k+1} \sup|f'|^k (U_n^\epsilon(f(n^{-\epsilon})) - n^{-\epsilon})^{k+1}. \end{aligned}$$

If we define the event $B_n^\epsilon = \{U_n^\epsilon(f(n^{-\epsilon})) - n^{-\epsilon} \le n^{-1/3}\log n\}$, then by a similar reasoning as in the proof of Lemma 4.2, it follows that $(U_n^\epsilon(f(n^{-\epsilon})) - n^{-\epsilon})^{k+1} = o_p(n^{-(2k+1)/6})$. The first integral on the right hand side of (5.5) can be written as

$$\left(\int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(n^{-\epsilon})|^k \, dx \right) 1_{B_n} + \left(\int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(n^{-\epsilon})|^k \, dx \right) 1_{B_n^c},$$

where the second term is of the order $o_p(n^{-(2k+1)/6})$ by the same reasoning as before. To bound

$$\left(\int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(n^{-\epsilon})|^k \, dx\right) 1_{B_n} \tag{5.6}$$

we distinguish between two cases:

(i) $1/6 < \epsilon \le 1/3$

(ii) $1/3 < \epsilon < (k-1)/(3k-6)$

In case (i), the integral (5.6) can be bounded by $|\hat{f}_n(n^{-\epsilon}) - f(n^{-\epsilon})|^k n^{-1/3} \log n$. According to Theorem 3.1 in Kulikov and Lopuhaä (2002), for $0 < \alpha < 1/3$,

$$n^{1/3}\left(\hat{f}_n(n^{-\alpha}) - f(n^{-\alpha})\right) \to |4f(0)f'(0)|^{1/3} V(0), \tag{5.7}$$

in distribution, where $V(0)$ is defined in (1.2). It follows that $|\hat{f}_n(n^{-\epsilon}) - f(n^{-\epsilon})| = \mathcal{O}_p(n^{-1/3})$ and therefore (5.6) is of the order $o_p(n^{-(2k+1)/6})$.

In case (ii), similar to Lemma 4.2, we will construct a suitable sequence $(a_i)_{i=1}^m$, such that the intervals $(n^{-a_i}, n^{-a_{i+1}}]$, for $i = 1, 2, \ldots, m-1$ cover the interval $(n^{-\epsilon}, U_n(f(n^{-\epsilon}))]$, and such that the integrals over these intervals can be bounded appropriately. First of all let

$$\epsilon = a_1 > a_2 > \cdots > a_{m-1} \ge 1/3 > a_m, \tag{5.8}$$

and let $z_i = n^{-a_i}$, $i = 1, \ldots, m$, so that $0 < z_1 < \cdots < z_{m-1} \le n^{-1/3} < z_m$. Then, similar to the proof of Lemma 4.2, we can bound (5.6) as follows

$$\left(\int_{n^{-\epsilon}}^{U_n^\epsilon(f(n^{-\epsilon}))} |\hat{f}_n(x) - f(n^{-\epsilon})|^k \, dx\right) 1_{B_n} \le \sum_{i=1}^{m-1} (z_{i+1} - z_i)|\hat{f}_n(z_i) - f(n^{-\epsilon})|^k.$$

Since $1/3 \le a_i \le \epsilon < 1$, for $i = 1, \ldots, m-1$, we can apply (4.9) and conclude that each term is of the order $\mathcal{O}_p(n^{-(1-a_i)/2})$. Therefore, it suffices to construct a sequence (a_i) satisfying (5.8), as well as

$$a_{i+1} + \frac{k(1-a_i)}{2} > \frac{2k+1}{6}, \qquad \text{for all } i = 1, \ldots, m-1. \tag{5.9}$$

One may take

$$\begin{aligned} a_1 &= \epsilon \\ a_{i+1} &= \frac{k(a_i - 1)}{2} + \frac{2k+1}{6} + \frac{1}{8}\left(\frac{k-1}{3(k-2)} - \epsilon\right), \qquad \text{for } i = 1, \ldots, m-1. \end{aligned}$$

Then (5.9) is satisfied and it remains to show that the described sequence strictly decreases and reaches 1/3 in finitely many steps. This follows from the fact that $a_i \le \epsilon$ and $k \ge 2.5$, since in that case

$$a_i - a_{i+1} = \frac{k-2}{2}\left(\frac{k-1}{3(k-2)} - a_i\right) - \frac{1}{8}\left(\frac{k-1}{3(k-2)} - \epsilon\right) \ge \frac{4k-9}{8}\left(\frac{k-1}{3(k-2)} - \epsilon\right) > 0.$$

As in the proof of Lemma 4.2, the argument for the second integral is similar. Now take $B_n^\epsilon = \{1 - n^{-\epsilon} - U_n^\epsilon(f(1-n^{-\epsilon})) \le n^{-1/3} \log n\}$. The case $1/6 < \epsilon \le 1/3$ can be treated in the same way as before. For the case $1/3 < \epsilon < (k-1)/(3k-6)$, we can use the same sequence (a_i) as above, but now define $z_i = 1 - n^{-a_i}$, $i = 1, \ldots, m$, so that $1 > z_1 > \ldots > z_{m-1} \geq 1 - n^{-1/3} > z_m$. Then we are left with considering

$$\left(\int_{U_n^\epsilon(f(1-n^{-\epsilon}))}^{1-n^{-\epsilon}} |f(1-n^{-\epsilon}) - \hat{f}_n(x)|^k \, dx \right) 1_{B_n} \leq \sum_{i=1}^{m-1} (z_i - z_{i+1}) |f(1-n^{-\epsilon}) - \hat{f}_n(z_i)|^k.$$

As before, each term in the sum is of the order $\mathcal{O}_p\left(n^{-a_{i+1}-k(1-a_i)/2}\right)$, for $i = 1, \ldots, m-1$. The sequence chosen above satisfies (5.9) and (5.8), which implies that the sum above is of the order $o_p(n^{-(2k+1)/6})$. ■

Apart from (5.4) we also need to bound the difference between integrals for U_n and its cut-off version U_n^ϵ:

$$\left| \int_{f(1)}^{f(0)} \frac{|U_n(a) - g(a)|^k}{|g'(a)|^{k-1}} \, da - \int_{f(1-n^{-\epsilon})}^{f(n^{-\epsilon})} \frac{|U_n^\epsilon(a) - g(a)|^k}{|g'(a)|^{k-1}} \, da \right| \tag{5.10}$$

$$\leq \int_{\tilde{f}_n(n^{-\epsilon})}^{f(0)} \frac{|U_n(a) - g(a)|^k}{|g'(a)|^{k-1}} \, da + \int_{f(1)}^{\tilde{f}_n(1-n^{-\epsilon})} \frac{|U_n(a) - g(a)|^k}{|g'(a)|^{k-1}} \, da.$$

The next lemma shows that both integrals on the right hand side are of negligible order.

Lemma 5.4 *For $k \geq 2.5$, let $1/6 < \epsilon < (k-1)/(3k-6)$. Furthermore let U_n be defined in (1.1) and let $\tilde{f}_n$ be defined in (4.1). Then*

$$\int_{\tilde{f}_n(n^{-\epsilon})}^{f(0)} \frac{\left|U_n(a) - g(a)\right|^k}{|g'(a)|^{k-1}} da = o_p(n^{-(2k+1)/6}),$$

and

$$\int_{f(1)}^{\tilde{f}_n(1-n^{-\epsilon})} \frac{\left|U_n(a) - g(a)\right|^k}{|g'(a)|^{k-1}} da = o_p(n^{-(2k+1)/6}).$$

Proof: Consider the first integral and define the event $A_n = \left\{ f(0) - \tilde{f}_n(n^{-\epsilon}) < n^{-1/6}/\log n \right\}$. For $1/6 < \epsilon \leq 1/3$, according to (5.7) we have that

$$\begin{aligned} f(0) - \tilde{f}_n(n^{-\epsilon}) \leq |\hat{f}_n(n^{-\epsilon}) - f(0)| &\leq |\hat{f}_n(n^{-\epsilon}) - f(n^{-\epsilon})| + \sup |f'| \, n^{-\epsilon} \\ &= \mathcal{O}_p(n^{-1/3}) + \mathcal{O}(n^{-\epsilon}) = o_p(n^{-1/6}/\log n). \end{aligned}$$

This means that if $1/6 < \epsilon \leq 1/3$, the probability $P\{A_n^c\} \to 0$. For $1/3 < \epsilon < 1$,

$$P\{A_n^c\} \leq P\{f(0) - \tilde{f}_n(n^{-\epsilon}) > 0\} \leq P\left\{ \hat{f}_n(n^{-\epsilon}) - f(n^{-\epsilon}) < n^{-\epsilon} \sup |f'| \right\} \to 0,$$

since according to (4.9), $\hat{f}_n(n^{-\epsilon}) - f(n^{-\epsilon})$ is of the order $n^{-(1-\epsilon)/2}$. Next, write the first integral as

$$\left(\int_{\tilde{f}_n(n^{-\epsilon})}^{f(0)} \frac{\left|U_n(a) - g(a)\right|^k}{|g'(a)|^{k-1}} da \right) 1_{A_n} + \left(\int_{\tilde{f}_n(n^{-\epsilon})}^{f(0)} \frac{\left|U_n(a) - g(a)\right|^k}{|g'(a)|^{k-1}} da \right) 1_{A_n^c}. \tag{5.11}$$

Similar to the argument used in Lemma 4.2, the second integral in (5.11) is of the order $o_p(n^{-(2k+1)/6})$. The expectation of the first integral is bounded by

$$\begin{aligned} E\int_{f(0)-n^{-1/6}/\log n}^{f(0)} \frac{\big|U_n(a)-g(a)\big|^k}{|g'(a)|^{k-1}}\,da &\leq n^{-k/3}C_1\int_{f(0)-n^{-1/6}/\log n}^{f(0)} E\big|V_n^E(a)\big|^k\,da \\ &\leq C_2 n^{(2k+1)/6}/\log n, \end{aligned}$$

using Lemma A.1. The Markov inequality implies that the first term in (5.11) is of the order $o_p(n^{-(2k+1)/6})$. For the second integral the proof is similar. ∎

Theorem 5.1 *Suppose conditions (A1) - (A3) of Theorem 1.1 are satisfied. Then for* $k \geq 2.5$ *and for any* ϵ*, such that* $1/6 < \epsilon < (k-1)/(3k-6)$,

$$n^{1/6}\left\{n^{1/3}\left(\int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |\hat{f}_n(x)-f(x)|^k\,dx\right)^{1/k} - \mu_k\right\}$$

converges in distribution to a normal random variable with zero mean and variance σ_k^2*, where* μ_k *and* σ_k^2 *are defined in Theorem 1.1.*

Proof: As in the proof of Theorem 1.1, it suffices to show that the difference

$$\left|\int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |\hat{f}_n(x)-f(x)|^k\,dx - \int_{f(1)}^{f(0)} \frac{|U_n(a)-g(a)|^k}{|g'(a)|^{k-1}}\,da\right|$$

is of the order $o_p(n^{-(2k+1)/6})$. We can bound this difference by

$$\left|\int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |\hat{f}_n(x)-f(x)|^k\,dx - \int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |f_n^\epsilon(x)-f(x)|^k\,dx\right| \tag{5.12}$$

$$+\left|\int_{f(1)}^{f(0)} \frac{|U_n(a)-g(a)|^k}{|g'(a)|^{k-1}}\,da - \int_{f(1-n^{-\epsilon})}^{f(n^{-\epsilon})} \frac{|U_n^\epsilon(a)-g(a)|^k}{|g'(a)|^{k-1}}\,da\right| \tag{5.13}$$

$$+\left|\int_{n^{-\epsilon}}^{1-n^{-\epsilon}} |f_n^\epsilon(x)-f(x)|^k\,dx - \int_{f(1-n^{-\epsilon})}^{f(n^{-\epsilon})} \frac{|U_n^\epsilon(a)-g(t)|^k}{|g'(a)|^{k-1}}\,da\right|. \tag{5.14}$$

Differences (5.12) and (5.13) can be bounded as in (5.4) and (5.10), so that Lemmas 5.3 and 5.4 imply that these terms are of the order $o_p(n^{-(2k+1)/6})$. Finally, Lemma 3.1 implies that (5.14) is bounded by

$$\int_{f(1-n^{-\epsilon})}^{f(n^{-\epsilon})} \frac{|U_n^\epsilon(a)-g(a)|^{k+1}}{|g'(a)|^k}\,da.$$

Write the integral as

$$\int_{f(1)}^{f(0)} \frac{|U_n(a)-g(a)|^{k+1}}{|g'(a)|^k}\,da + \left(\int_{f(1)}^{f(0)} \frac{|U_n(a)-g(a)|^{k+1}}{|g'(a)|^k}da - \int_{f(1-n^{-\epsilon})}^{f(n^{-\epsilon})} \frac{|U_n^\epsilon(a)-g(a)|^{k+1}}{|g'(a)|^k}da\right).$$

Then Corollary 2.1 and Lemma 5.4 imply that both terms are of the order $o_p(n^{-(2k+1)/6})$. This proves the theorem. ∎

A Appendix

The proofs in Section 2 follow the same line of reasoning as in GROENEBOOM ET AL.(1999). Since we will frequently use results from this paper, we state them for easy reference. First, the tail probabilities of V_n^J have a uniform exponential upper bound.

Lemma A.1 *For $J = E, B, W$, let V_n^J be defined by (2.2). Then there exist constants $C_1, C_2 > 0$ only depending on f, such that for all $n \geq 1$, $a \in (f(1), f(0))$ and $x > 0$,*

$$P\left\{\left|V_n^J(a)\right| \geq x\right\} \leq C_1 \exp(-C_2 x^3).$$

Properly normalized versions of $V_n^J(a)$ converge in distribution to $\xi(c)$ defined in (1.3). To be more precise, for $a \in (f(1), f(0))$, define

$$\phi_1(a) = \frac{|f'(g(a))|^{2/3}}{(4a)^{1/3}} > 0, \tag{A.15}$$

$$\phi_2(a) = (4a)^{1/3}|f'(g(a))|^{1/3} > 0, \tag{A.16}$$

and let

$$J_n(a) = \left\{c : a - \phi_2(a)cn^{-1/3} \in (f(1), f(0))\right\}.$$

For $J = E, B, W$ and $c \in J_n(a)$, define,

$$V_{n,a}^J(c) = \phi_1(a)V_n^J(a - \phi_2(a)cn^{-1/3}), \tag{A.17}$$

Then we have the following property.

Lemma A.2 *For $J = E, B, W$, integer $d \geq 1$, $a \in (f(1), f(0))$ and $c \in J_n(a)^d$,we have joint distributional convergence of $(V_{n,a}^J(c_1), \ldots, V_{n,a}^J(c_d))$ to the random vector $(\xi(c_1), \ldots, \xi(c_d))$.*

Due to the fact that Brownian motion has independent increments, the process V_n^W is mixing.

Lemma A.3 *The process $\{V_n^W(a)) : a \in (f(1), f(0))\}$ is strong mixing with mixing function: $\alpha_n(d) = 12e^{-C_3nd^3}$, where the constant $C_3 > 0$ only depends on f.*

As a direct consequence of Lemma A.3 we have the following lemma, which is a slight extension of Lemma 4.1 in GROENEBOOM ET AL.(1990).

Lemma A.4 *Let l and m be fixed such that $l + m > 0$ and let h be a continuous function. Define*

$$c_h = 2\int_0^1 (4f(x))^{(2l+2m+1)/3}|f'(x)|^{(4-4l-4m)/3}h(f(x))^2\,dx.$$

Then,

$$\mathrm{var}\left(n^{1/6}\int_{f(1)}^{f(0)} V_n^W(a)^l|V_n^W(a)|^m h(a)\,da\right) \to c_h\int_0^\infty \mathrm{cov}(\xi(0)^l|\xi(0)|^m, \xi(c)^l|\xi(c)|^m)\,dc,$$

as $n \to \infty$.

Proof: The proof runs along the lines of the proof of Lemma 4.1 in Groeneboom et al.(1999). We first have that

$$\operatorname{var}\left(n^{1/6}\int_{f(1)}^{f(0)} V_n^W(a)^l |V_n^W(a)|^m h(a)\,da\right)$$
$$= -2\int_{f(1)}^{f(0)}\int_0^{n^{1/3}\phi_2(a)^{-1}(a-f(0))} (4a)^{(2l+2m+1)/3}|g'(a)|^{\frac{4(l+m)-1}{3}} h(a)h(a-\phi_2(a)n^{-1/3}c)$$
$$\cdot \operatorname{cov}\left(V_{n,a}^W(0)^l|V_{n,a}^W(0)|^m, V_{n,a}^W(c)^l|V_{n,a}^W(c)|^m\right)\,dc\,da.$$

According to Lemma A.1, for a and c fixed, the sequence $V_{n,a}^W(c)^l|V_{n,a}^W(c)|^m$ is uniformly integrable. Hence by Lemma A.2 the moments of $\left(V_{n,a}^W(0)^l|V_{n,a}^W(0)|^m, V_{n,a}^W(c)^l|V_{n,a}^W(c)|^m\right)$ converge to corresponding moments of $(\xi(0)^l|\xi(0)|^m, \xi(c)^l|\xi(c)|^m)$. Again Lemma A.1 and the fact that $l+m>0$, yields that

$$E|V_{n,a}^W(0)|^{3(l+m)} < C \quad \text{and} \quad E|V_{n,a}^W(c)|^{3(l+m)} < C,$$

where $C>0$ does not depend on n, a and c. Together with Lemma A.3 and Lemma 3.2 in Groeneboom et al.(1999) this yields that

$$\left|\operatorname{cov}\left(V_{n,a}^W(0)^l|V_{n,a}^W(0)|^m, V_{n,a}^W(c)^l|V_{n,a}^W(c)|^m\right)\right| \le D_1 e^{-D_2|c|^3},$$

where D_1 and D_2 do not depend on n, a and c. It follows by dominated convergence that

$$\operatorname{var}\left(n^{1/6}\int_{f(1)}^{f(0)} V_{n,a}^W(0)^l |V_n^W(a)|^m h(a)\,da\right)$$
$$\to -c_h \int_0^{-\infty} \operatorname{cov}\left(\xi(0)^l\,|\xi(0)|^m\,, \xi(c)^l\,|\xi(c)|^m\right)\,dc,$$
$$= c_h \int_0^{\infty} \operatorname{cov}\left(\xi(0)^l\,|\xi(0)|^m\,, \xi(c)^l\,|\xi(c)|^m\right)\,dc,$$

using that the process ξ is stationary, where

$$\begin{aligned} c_h &= 2\int_{f(1)}^{f(0)} (4a)^{(2l+2m+1)/3}|g'(a)|^{(4l+4m-1)/3}h(a)^2\,da \\ &= 2\int_0^1 (4f(x))^{(2l+2m+1)/3}|f'(x)|^{(4-4l-4m)/3}h(f(x))^2\,dx. \end{aligned}$$

This proves the lemma. ■

Proof of Theorem 2.1: Write

$$W_n^k(a) = \frac{|V_n^W(a)|^k - E|V_n^W(a)|^k}{|g'(a)|^{k-1}},$$

and define

$$\begin{aligned} L_n &= (f(0)-f(1))n^{-1/3}(\log n)^3, \\ M_n &= (f(0)-f(1))n^{-1/3}\log n, \\ N_n &= \left[\frac{(f(0)-f(1))}{L_n+M_n}\right] = \left[\frac{n^{1/3}}{\log n + (\log n)^3}\right], \end{aligned}$$

where $[x]$ denotes the integer part of x. We divide the interval $(f(1), f(0))$ into $2N_n+1$ blocks of alternating length

$$\begin{aligned} A_j &= (f(1)+(j-1)(L_n+M_n), f(1)+(j-1)(L_n+M_n)+L_n], \\ B_j &= (f(1)+(j-1)(L_n+M_n)+L_n, f(1)+j(L_n+M_n)], \end{aligned}$$

where $j = 1, \ldots, N_n$. Now write

$$T_{n,k} = S'_{n,k} + S''_{n,k} + R_{n,k},$$

where

$$\begin{aligned} S'_{n,k} &= n^{1/6} \sum_{j=1}^{N_n} \int_{A_j} W_n^k(a)\, da, \\ S''_{n,k} &= n^{1/6} \sum_{j=1}^{N_n} \int_{B_j} W_n^k(a)\, da, \\ R_{n,k} &= n^{1/6} \int_{f(1)+N_n(L_n+M_n)}^{f(0)} W_n^k(a)\, da. \end{aligned}$$

From here on the proof is completely the same as the proof of Theorem 4.1 in Groeneboom et al. (1999). Therefore we omit all specific details and only give a brief outline of the argument. Lemmas A.1 and A.3 imply that all moments of $W_n^k(a)$ are bounded uniformly in a, and that $E|W_n^k(a)W_n^k(b)| \le D_1 \exp\left(-D_2 n|b-a|^3\right)$. This is used to ensure that $ER_n^2 \to 0$ and that the contribution of the small blocks is negligible: $E(S''_{n,k})^2 \to 0$. We then only have to consider the contribution over the big blocks. When we denote

$$Y_j = n^{1/6} \int_{A_j} W_n^k(a)\, da \quad \text{and} \quad \sigma_n^2 = \operatorname{var}\left(\sum_{j=1}^{N_n} Y_j\right),$$

one finds that

$$\left| E\exp\left\{\frac{iu}{\sigma_n}\sum_{j=1}^{N_n} Y_j\right\} - \prod_{j=1}^{N_n} E\exp\left\{\frac{iu}{\sigma_n} Y_j\right\}\right| \le 4(N_n-1)\exp(-C_3 n M_n^3) \to 0,$$

where $C_3 > 0$ only depends on f. This means that we can apply the central limit theorem to independent copies of Y_j. Hence, asymptotic normality of $S'_{n,k}$ follows if we show that the contribution of the big blocks satisfies the Lindeberg condition, i.e., for each $\varepsilon > 0$,

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{N_n} EY_j^2 1_{\{|Y_j|>\varepsilon\sigma_n\}} \to 0, \quad n \to \infty. \tag{A.18}$$

By using the uniform boundedness of the moments of $|W_n^k(a)|$, we have that

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{N_n} EY_j^2 1_{\{|Y_j|>\epsilon\sigma_n\}} \le \frac{1}{\epsilon\sigma_n^3} N_n \sup_{1\le k\le N_n} E|Y_j|^3 = \mathcal{O}\left(\sigma_n^{-3} n^{-1/6} (\log n)^6\right).$$

Similar computations as in the proof of Theorem 4.1 in GROENEBOOM ET AL.(1999), yields that

$$\sigma_n^2 = \mathrm{var}(T_{n,k}) + \mathcal{O}(1).$$

Application of Lemma A.4, with $l = 0$, $m = k$ and $h(a) = 1/|g'(a)|^{k-1}$, implies that $\sigma_n^2 \to \sigma^2$, which proves (A.18). ■

In order to prove Lemma 2.1, we first prove the following lemma.

Lemma A.5 *Let V_n^W be defined by (2.2) and let $V(0)$ be defined by (1.2). Then for $k \geq 1$, and for all a such that*

$$n^{1/3}\left\{F(g(a)) \wedge (1 - F(g(a)))\right\} \geq \log n, \tag{A.19}$$

we have

$$E|V_n^W(a)|^k = E|V(0)|^k \frac{(4a)^{k/3}}{|f'(g(a))|^{2k/3}} + \mathcal{O}(n^{-1/3}(\log n)^{k+3}),$$

where the term $\mathcal{O}(n^{-1/3}(\log n)^{k+3})$ is uniform in all a satisfying (A.19).

Proof: The proof relies on the proof of Corollary 3.2 in GROENEBOOM ET AL.(1999). There it is shown that, if we define

$$H_n(y) = n^{1/3}\left\{H\left(F(g(a)) + n^{-1/3}y\right) - g(a)\right\},$$

with H being the inverse of F, and

$$V_{n,b} = \sup\left\{y \in [-n^{1/3}F(g(a)), n^{1/3}(1 - F(g(a)))] : W(y) - by^2 \text{ is maximal}\right\},$$

with $b = |f'(g(a))|/(2a^2)$, then for the event $A_n = \{|V_n^W(a)| \leq \log n, |H_n(V_{n,b})| \leq \log n\}$, one has that $P\{A_n^c\}$ is of the order $\mathcal{O}(e^{-C(\log n)^3})$, which then implies that

$$\sup_{a \in (f(1), f(0))} E|V_n^W(a) - H_n(V_{n,b})| = \mathcal{O}(n^{-1/3}(\log n)^4).$$

Similarly, together with an application of the mean value theorem, this yields

$$\sup_{a \in (f(1), f(0))} E\left||V_n^W(a)|^k - |H_n(V_{n,b})|^k\right| = \mathcal{O}(n^{-1/3}(\log n)^{3+k}). \tag{A.20}$$

Note that by definition, the argmax $V_{n,b}$ closely resembles the argmax $V_b(0)$, where

$$V_b(c) = \underset{t \in \mathbb{R}}{\mathrm{argmax}}\{W(t) - b(t - c)^2\}. \tag{A.21}$$

Therefore we write

$$E\left|H_n(V_{n,b})\right|^k = E\left|H_n(V_b(0))\right|^k + E\left(|H_n(V_{n,b})|^k - |H_n(V_b(0))|^k\right). \tag{A.22}$$

Since by Brownian scaling $V_b(c)$ has the same distribution as $b^{-2/3}V(cb^{2/3})$, where V is defined in (1.2), together with the conditions on f, we find that

$$E\left|H_n(V_b(0))\right|^k = a^{-k}E\left|V_b(0)\right|^k + \mathcal{O}(n^{-1/3}) = \frac{(4a)^{k/3}}{|f'(g(a))|^{2k/3}} E|V(0)|^k + \mathcal{O}(n^{-1/3}).$$

As in the proof of Corollary 3.2 in Groeneboom et al.(1999), $V_{n,b}$ can only be different from $V_b(0)$ with probability of the order $\mathcal{O}(e^{-\frac{2}{3}(\log n)^3})$. Hence, from (A.22) we conclude that

$$E\,|H_n(V_{n,b})|^k = \frac{(4a)^{k/3}}{|f'(g(a))|^{2k/3}} E|V(0)|^k + \mathcal{O}(n^{-1/3}).$$

Together with (A.20) this proves the lemma. ■

Proof of Lemma 2.1: The result immediately follows from Lemma A.5. The values of a for which condition (A.19) does not hold, gives a contribution of the order $\mathcal{O}(n^{-1/3}\log n)$ to the integral $\int E|V_n^W(a)|^k\,da$, and finally,

$$\int_{f(1)}^{f(0)} \frac{(4a)^{k/3}}{|f'(g(a))|^{2k/3}|g'(a)|^{k-1}}\,da = \int_0^1 (4f(x))^{k/3}|f'(x)|^{k/3}dx. \quad ■$$

Proof of Lemma 2.2: The proof of the first statement relies on the proof of the Corollary 3.3 in Groeneboom et al.(1990). Here it is shown, that if for a belonging to the set

$$J_n = \{a : \text{both } a \text{ and } a(1-\xi_n n^{-1/2}) \ \in (f(1), f(0))\},$$

we define

$$V_n^B(a,\xi_n) = V_n^B(a(1-n^{-1/2}\xi_n)) + n^{1/3}\left\{g(a(1-n^{-1/2}\xi_n)) - g(a)\right\},$$

then for the event $A_n = \{|\xi_n| \le n^{1/6}, |V_n^W(a)| \le \log n, |V_n^B(a,\xi_n)| \le \log n\}$, one has that $P\{A_n^c\}$ is of the order $\mathcal{O}(e^{-C(\log n)^3})$, which then implies that

$$\int_{a\in J_n} E\left|V_n^B(a,\xi_n) - V_n^W(a)\right|\,da = \mathcal{O}(n^{-1/3}(\log n)^3).$$

Hence, by using the same method as in proof of Lemma 2.1, we obtain:

$$\int_{a\in J_n} E\left||V_n^B(a,\xi_n)|^k - |V_n^W(a)|^k\right|\,da = \mathcal{O}(n^{-1/3}(\log n)^{k+2}).$$

From Lemma A.1 it also follows that $E|V_n^B(a)|^k = \mathcal{O}(1)$ and $E|V_n^W(a)|^k = \mathcal{O}(1)$, uniformly with respect to n and $a \in (f(1), f(0))$. Hence the contribution of the integrals over $[f(1), f(0)] \setminus J_n$ is negligible, and it remains to show that

$$n^{1/6}\int_{a\in J_n} \left\{|V_n^B(a,\xi_n)|^k - |V_n^B(a)|^k\right\}\,da = o_p(1). \tag{A.23}$$

For $k=1$, this is shown in the proof of Corollary 3.3 in Groeneboom et al.(1999), so we may assume that $k>1$. Completely similar to the proof in the case $k=1$, we first obtain

$$\begin{aligned} &n^{1/6}\int_{a\in J_n} \left\{|V_n^B(a,\xi_n)|^k - |V_n^B(a)|^k\right\}\,da \\ &\quad = n^{1/6}\int_{f(1)}^{f(0)} \left\{\left|V_n^B(a) - ag'(a)\xi_n n^{-1/6}\right|^k - |V_n^B(a)|^k\right\}\,da + \mathcal{O}_p(n^{-1/3}). \end{aligned}$$

Let $\epsilon > 0$ and write $\Delta_n(a) = ag'(a)\xi_n n^{-1/6}$. Then we can write

$$\begin{aligned} & n^{1/6}\int_{f(1)}^{f(0)} \left\{\left|V_n^B(a)-\Delta_n(a)\right|^k - |V_n^B(a)|^k\right\}\,da \\ &= n^{1/6}\int_{f(1)}^{f(0)} \left\{\left|V_n^B(a)-\Delta_n(a)\right|^k - |V_n^B(a)|^k\right\} 1_{[0,\epsilon]}(|V_n^B(a)|)\,da \qquad \text{(A.24)} \\ &+ n^{1/6}\int_{f(1)}^{f(0)} \left\{\left|V_n^B(a)-\Delta_n(a)\right|^k - |V_n^B(a)|^k\right\} 1_{(\epsilon,\infty)}(|V_n^B(a)|)\,da. \qquad \text{(A.25)} \end{aligned}$$

First consider the term (A.24) and distinguish between

1. $|V_n^B(a)| < 2|\Delta_n(a)|$,
2. $|V_n^B(a)| \geq 2|\Delta_n(a)|$,

In case 1,

$$\left||V_n^B(a)-\Delta_n(a)|^k - |V_n^B(a)|^k\right| \leq 3^k|\Delta_n(a)|^k + 2^k|\Delta_n(a)|^k \leq (3^k+2^k)|ag'(a)\xi_n|^k n^{-k/6}.$$

In case 2, note that

$$\left||V_n^B(a)-\Delta_n(a)|^k - |V_n^B(a)|^k\right| = k|\theta|^{k-1}|\Delta_n(a)|,$$

where θ is between $|V_n^B(a)| \leq \epsilon$ and $|V_n^B(a)-\Delta_n(a)| \leq \frac{3}{2}\epsilon$. Using that ξ_n and V_n^B are independent, the expectation of (A.24) is bounded from above by

$$C_1\epsilon^{k-1}E|\xi_n|\int_{f(1)}^{f(0)} |ag'(a)|P\{|V_n^B(a)| \leq \epsilon\}\,da + \mathcal{O}_p(n^{-(k-1)/6}),$$

where $C_1 > 0$ only depends on f and k. Hence, since $k > 1$, we find that

$$\limsup_{n\to\infty} n^{1/6}\int_{f(1)}^{f(0)} \left\{\left|V_n^B(a) - ag'(a)\xi_n n^{-1/6}\right|^k - |V_n^B(a)|^k\right\} 1_{[0,\epsilon]}(|V_n^B(a)|)\,da \qquad \text{(A.26)}$$

is bounded from above by $C_2\epsilon^{k-1}$, where $C_2 > 0$ only depends on f and k. Letting $\epsilon \downarrow 0$ and using that $k > 1$, then yields that (A.24) tends to zero.

The term (A.25) is equal to

$$\int_{f(1)}^{f(0)} \frac{-2\xi_n ag'(a)V_n^B(a) + (ag'(a)\xi_n)^2 n^{-1/6}}{|V_n^B(a)-\Delta_n(a)| + |V_n^B(a)|} \cdot k\theta(a)^{k-1} 1_{(\epsilon,\infty)}(|V_n^B(a)|)\,da, \qquad \text{(A.27)}$$

where $\theta(a)$ is between $|V_n^B(a)-\Delta_n(a)|$ and $|V_n^B(a)|$. Note that for $|V_n^B(a)| > \epsilon$,

$$\left|\frac{2V_n^B(a)}{|V_n^B(a)-\Delta_n(a)| + |V_n^B(a)|} - \frac{V_n^B(a)}{|V_n^B(a)|}\right| \leq \frac{|ag'(a)n^{-1/6}\xi_n|}{\epsilon} = \mathcal{O}_p(n^{-1/6}),$$

uniformly in $a \in (f(1), f(0))$, so that (A.27) is equal to

$$\begin{aligned} & -k\xi_n\int_{f(1)}^{f(0)} ag'(a)V_n^B(a)|V_n^B(a)|^{k-2} 1_{(\epsilon,\infty)}(|V_n^B(a)|)\,da \\ &+ k\xi_n\int_{f(1)}^{f(0)} ag'(a)\frac{V_n^B(a)}{|V_n^B(a)|}\left(|V_n^B(a)|^{k-1} - \theta(a)^{k-1}\right) 1_{(\epsilon,\infty)}(|V_n^B(a)|)\,da + \mathcal{O}_p(n^{-1/6}). \end{aligned}$$

We have that

$$\left||V_n^B(a)|^{k-1} - \theta(a)^{k-1}\right| \leq |V_n^B(a)|^{k-1} \left|\left|1 - \frac{\Delta_n(a)}{V_n^B(a)}\right|^{k-1} - 1\right| = \mathcal{O}_p(n^{-1/6}),$$

where the big $\mathcal{O}$-term in uniform in a. This means that (A.27) is equal to

$$-k\xi_n \int_{f(1)}^{f(0)} ag'(a)V_n^B(a)|V_n^B(a)|^{k-2}\,da \tag{A.28}$$

$$+k\xi_n \int_{f(1)}^{f(0)} ag'(a)\text{sign}(V_n^B(a))|V_n^B(a)|^{k-1}1_{[0,\epsilon)}(|V_n^B(a)|)\,da + \mathcal{O}_p(n^{-1/6}). \tag{A.29}$$

The integral in (A.29) is of the order $\mathcal{O}(\epsilon^{k-1})$, whereas $E\xi_n^2 = 1$. Since $k > 1$, this means that after letting $\epsilon \downarrow 0$, (A.29) tends to zero. Finally, let $S_n^B(a) = ag'(a)V_n^B(a)\left|V_n^B(a)\right|^{k-2}$ and write

$$E\left(\xi_n \int_{f(1)}^{f(0)} S_n^B(a)\,da\right)^2 = \text{var}\left(\int_{f(1)}^{f(0)} S_n^B(a)\,da\right) + \left(E\int_{f(1)}^{f(0)} S_n^B(a)\,da\right)^2.$$

Then, since according to Lemma A.1, all moments of $|S_n^B(a)|$ are bounded uniformly in a, we find by dominated convergence and Lemma A.2 that

$$\lim_{n\to\infty} E\int_{f(1)}^{f(0)} S_n^B(a)\,da = \int_{f(1)}^{f(0)} \frac{a|g'(a)|}{(\phi_1(a))^k}\left(E\xi(0)\,|\xi(0)|^{k-2}\right)\,da = 0,$$

because the distribution of $\xi(0)$ is symmetric. Applying Lemma A.4 with $l = 1$, $m = k-2$ and $h(a) = ag'(a)$ we obtain

$$\text{var}\left(\int_{f(1)}^{f(0)} ag'(a)V_n^B(a)\left|V_n^B(a)\right|^{k-2}\,da\right) = O(n^{-1/3}).$$

We conclude that (A.27) tends to zero in probability. This proves the first statement of the lemma.

The proof of the second statement relies on the proof of Corollary 3.1 in GROENEBOOM ET AL.(1999). There it is shown that for the event $A_n = \left\{|V_n^B(a)| < \log n\,, |V_n^E(a)| < \log n\right\}$ one has that $P\{A_n^c\}$ is of the order $\mathcal{O}(e^{-C(\log n)^3})$. Furthermore, if $K_n = \{\sup_t |E_n(t) - B_n(F(t))| \leq n^{-1/2}(\log n)^2\}$, then $P\{K_n\} \to 1$ and

$$E\left||V_n^E(a)| - |V_n^B(a)|\right| 1_{A_n\cap K_n} = \mathcal{O}(n^{-1/3}(\log n)^3), \tag{A.30}$$

uniformly in $a \in (f(1), f(0))$. By the mean value theorem, together with (A.30), we now have that

$$\begin{aligned} E\left||V_n^E(a)|^k - |V_n^B(a)|^k\right| 1_{K_n} &\leq k(\log n)^{k-1}E\left||V_n^E(a)| - |V_n^B(a)|\right| 1_{A_n\cap K_n} + 2n^{k/3}P\{A_n^c\} \\ &= \mathcal{O}(n^{-1/3}(\log n)^{k+2}) + \mathcal{O}(n^{k/3}e^{-C(\log n)^3}). \end{aligned}$$

This proves the lemma. ■

Before proving Lemma 5.1, we first prove the following lemma.

Lemma A.6 *Let $k \geq 2.5$ and $z_n = 1/(2nf(0))$. Then there exist $0 < a_1 < b_1 < a_2 < b_2 < \infty$, such that for $i = 1, 2$*

$$\liminf_{n\to\infty} P\left\{ n \int_0^{z_n} |\hat{f}_n(x) - f(x)|^k \, dx \in [a_i, b_i] \right\} > 0.$$

Proof: Consider the event $A_n = \{X_i \geq z_n, \text{ for all } i = 1, 2, \ldots, n\}$. Then it follows that $P\{A_n\} \to 1/\sqrt{e} > 1/2$. Since on the event A_n, the estimator $\hat{f}_n$ is constant on the interval $[0, z_n]$, for any $a_i > 0$ we have

$$\begin{aligned} P\left\{ n \int_0^{z_n} |\hat{f}_n(x) - f(x)|^k \, dx \in [a_i, b_i] \right\} &\geq P\left\{ \left(n \int_0^{z_n} |\hat{f}_n(0) - f(x)|^k \, dx \right) 1_{A_n} \in [a_i, b_i] \right\} \\ &= P\left\{ \left(\frac{|\hat{f}_n(0) - f(0)|^k}{2f(0)} + R_n \right) 1_{A_n} \in [a_i, b_i] \right\}, \end{aligned}$$

where

$$R_n = n \int_0^{z_n} k\theta_n(x)^{k-1} \left(|\hat{f}_n(0) - f(x)| - |\hat{f}_n(0) - f(0)| \right) dx,$$

with $\theta_n(x)$ between $|\hat{f}_n(0) - f(x)|$ and $|\hat{f}_n(0) - f(0)|$. Using (4.7), we obtain that R_n is of the order $\mathcal{O}_p(n^{-1})$ and therefore

$$\frac{|\hat{f}_n(0) - f(0)|^k}{2f(0)} + R_n \to \frac{f(0)^{k-1}}{2} \left| \sup_{1\leq j<\infty} \frac{j}{\Gamma_j} - 1 \right|^k$$

in distribution. Now, choose $0 < a_1 < b_1 < a_2 < b_2 < \infty$, such that for $i = 1, 2$

$$P\left\{ \frac{f(0)^{k-1}}{2} \left| \sup_{1\leq j<\infty} \frac{j}{\Gamma_j} - 1 \right|^k \in [a_i, b_i] \right\} > 1 - 1/\sqrt{e}.$$

Then, for $i = 1, 2$ we find

$$P\left\{ n \int_0^{z_n} |\hat{f}_n(x) - f(x)|^k \, dx \in [a_i, b_i] \right\} \geq P\left\{ \left(\frac{|\hat{f}_n(0) - f(0)|^k}{2f(0)} + R_n \right) \in [a_i, b_i] \right\} - P\{A_n^c\},$$

which converges to a positive value. ■

Proof of Lemma 5.1: Take $0 < a_1 < b_1 < a_2 < b_2 < \infty$ as in Lemma A.6, and let A_{ni} be the event

$$A_{ni} = \left\{ n \int_0^{z_n} |\hat{f}_n(x) - f(x)|^k \, dx \in [a_i, b_i] \right\}.$$

Then

$$n^{k/3} E \int_0^1 |\hat{f}_n(x) - f(x)|^k \, dx \geq n^{k/3} E \int_0^{z_n} |\hat{f}_n(x) - f(x)|^k \, dx 1_{A_{n1}} \geq a_1 n^{(k-3)/3} P\{A_{n1}\}.$$

Since, according to Lemma A.6, $P\{A_{n1}\}$ tends to a positive constant, this proves (i).

For (ii), write $X_n = n\int_0^{z_n} |\hat{f}_n(x) - f(x)|^k\,dx$, and define $B = \{EX_n \geq (a_2 + b_1)/2\}$. Then

$$\begin{aligned}\mathrm{var}(X_n) &\geq E(X_n - EX_n)^2 1_{A_{n1}\cap B} + E(X_n - EX_n)^2 1_{A_{n2}\cap B^c} \\ &\geq \frac{1}{4}(a_2 - b_1)^2 P\{A_{n1}\} 1_B + \frac{1}{4}(a_2 - b_1)^2 P\{A_{n2}\} 1_{B^c} \\ &\geq \frac{1}{4}(a_2 - b_1)^2 \min(P\{A_{n1}\}, P\{A_{n2}\}).\end{aligned}$$

Hence, according to Lemma A.6,

$$\begin{aligned}&\liminf_{n\to\infty} \mathrm{var}\left(n^{(2k+1)/6}\int_0^{z_n} |\hat{f}_n(x) - f(x)|^k\,dx\right) \\ &\quad \geq \liminf_{n\to\infty} n^{(2k-5)/3}\frac{1}{4}(a_2 - b_1)^2 \min(P\{A_{n1}\}, P\{A_{n2}\}) = \infty. \qquad \blacksquare\end{aligned}$$

References

Durot, C.(2000) Sharp asymptotics for isotonic regression, *Probab. Theory and Related Fields* **122**, no. 2, p.222–240.

Grenander, U. (1956). On the theory of mortality measurements, Part II, *Skand. Akt. Tid.* **39**, p.125–153.

Groeneboom, P. (1985). Estimating a monotone density, In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (Edited by L.LeCam and R.Olshen) **2**, p.539–555.

Groeneboom, P., Hooghiemstra, G. and Lopuhaä, H.P. (1999). Asymptotic normality of the L_1-error of the Grenander estimator, *Ann. Statist.* **27**, No.4, p.1316–1347.

P.Groeneboom and J.A.Wellner (1992) "Information bounds and nonparametric maximum likelihood estimation", *Birkhäuser Verlag.*

Kim, J. and Pollard, D. (1990). Cube root asymptotics, *Ann. Statist.* **18**, No.1, p.191–219.

Komlos, J., Major, P. and Tusnady, G. (1975). An approximation of partial sums of independent rv'-s, and the sample DF. I, *Z. Wahrsch. Verw. Gebiete* **32**, p.111–131.

Kulikov, V.N. and Lopuhaä, H.P. (2002). The behavior of the NPMLE of a decreasing density near the boundaries of the support. Tentatively accepted by Annals of Statistics.

Kulikov, V.N. and Lopuhaä, H.P. (2004). Testing for a monotone density using L_k-distances between the empirical distribution function and its concave majorant. Report 2004-028 Eurandom. Submitted.
(`http://www.eurandom.nl/reports/reports%202004/028VKreport.pdf`)

Pollard, D. (1984). Convergence of stochastical processes, *Springer-Verlag.*

Prakasa Rao, B.L.S. (1969). Estimation of a unimodal density, *Sankhya Ser. A* **31**, p.23–36.

Sun, J. and Woodroofe, M. (1996). Adaptive smoothing for a penalized NPMLE of a non-increasing density, *J. Statist. Planning and Inference* **52**, p.143–159.

Woodroofe, M. and Sun, J. (1993). A penalized maximum likelihood estimate of $f(0+)$ when f is non-increasing, *Statist. Sinica* **3**, p.501–515.