Journal of Statistical Planning and Inference 98 (2001) 107–117

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Asymptotic analysis of isotonic estimation for grouped data

Rong Zhang[a, 1], Jeankyung Kim[b], Michael Woodroofe[a, *, 2]

[a] *Department of Statistics, University of Michigan, 105 South State St., Ann Arbor, MI 48109-1285, USA*

[b] *Department of Statistics, Inha University, Incheon, 402-751, South Korea*

Received 25 February 2000; received in revised form 20 October 2000; accepted 23 October 2000

Abstract

A non-parametric estimator of a non-increasing density is found in a class of piecewise linear functions when the data consist only counts. The estimator is shown to be consistent, and the limiting distribution of the estimator is found under different assumptions on the width of the class intervals. © 2001 Elsevier Science B.V. All rights reserved.

MSC: 62G07; 62G20

Keywords: Asymptotic distribution; Brownian motion; Central limit theorem; Counts; Distance sampling; EM-algorithm; Empirical process and maximum likelihood estimation

1. Introduction

Let $X_1, \ldots, X_n$ be independent and identically distributed non-negative random variables with an unknown density f, but suppose that f is known to be non-increasing on $[0, \infty)$. It was shown by Grenander (1985) that the maximum likelihood estimate of f is the left derivative of the least concave majorant (LCM) of the empirical function F_n. Its asymptotic properties can be found in Prakasa Rao (1969), Groeneboom (1985), and Devroye (1987). Grenander's estimator is a step function with jumps at some of the sample points. A continuous non-parametric maximum likelihood estimate (NPMLE) of a monotone density for grouped data is studied in Woodroofe and Zhang (1999). Let $h > 0$ and suppose that only the counts

$$n_k = \#\{i : (k-1)h < X_i \leqslant kh\}, \quad k = 1, 2, \ldots \tag{1}$$

are observed. The NPMLE of f is found via EM-algorithm within the class of continuous, non-increasing, piecewise linear densities on $[0, \infty)$ with knots at $h, 2h, \ldots$, along

* Corresponding author.

E-mail address: michaelw@umich.edu (M. Woodroofe).

[1] Research supported by the National Science Foundation under DMS9626347.

[2] Research supported by the Army under DAAG55-98-1-0482.

PII: S0378-3758(00)00301-3

with an approximate NPMLE which requires no iteration. The NPMLE for grouped data is partly motivated by distance sampling models, as described by Buckland et al. (1992). There an observer searches for hidden objects and records the distance X from the object to his/her closest point of approach. Let $g(x)$ be the probability of finding an object x distance away. The conditional density function of X, given an object is detected is $f(x)=g(x)/\mu$, where $\mu=\int_0^\infty g(x)\,\mathrm{d}x$ is the unconditional probability of finding an object. Grouped data are common in such studies, and there is interest in the density of X, especially at $x=0$. It is reasonable to suppose that the density of X (given observation) is non-increasing here and, therefore, natural to consider isotonic methods. Further motivation is provided by Bickel and Fan (1996) who suggest grouping data before applying isotonic methods in the context of unimodal density estimation. See Robertson et al. (1989) for background on isotonic methods.

In this paper, the behavior of the approximate NPMLE as $n\to\infty$ is studied. Consistency is established in Section 2 and asymptotic distributions are studied in Section 4. For the limiting distribution, the width of the class intervals h, called bandwidth henceforth, plays an important role. For large h, say $nh^3\to\infty$, the estimator follows a normal distribution with convergent rate $(nh)^{1/2}$. For smaller h, such as $h\sim cn^{-1/3}$ or $nh^3=\mathrm{o}(1)$, the convergent rate is $n^{1/3}$ and the limiting distribution is more complicated.

2. Consistency

Let $\{n_1,\dots,n_m\}$ be as in (1). Here $m=\lfloor\max(X_1,\dots,X_n)/h\rfloor+1$, where $\lfloor x\rfloor$ is the greatest integer less than x, and $n=\sum_{k=1}^m n_k$. Denote kh by x_k, $k=0,1,\dots,m$, and define $y_{-1}=0$, $y_k=(x_k+x_{k+1})/2$ for $k=0,\dots,m$, and $x_{m+1}=x_m$. The piecewise linear NPMLE of f at $x=x_k$ is the left derivative of the LCM of $G_{\hat{f}_n}$ at $y=y_k$, where

$$G_{\hat{f}}(y_k)=\frac{1}{n}\left[\sum_{i=1}^{k}n_i+\left(\frac{\hat{f}(x_k)}{\hat{f}(x_k)+\hat{f}(x_{k+1})}\right)n_{k+1}\right].$$

Woodroofe and Zhang (1999) show that the NPMLE can be calculated via EM-algorithm and that the iteration converges. If we replace $\hat{f}(x_k)/(\hat{f}(x_k)+\hat{f}(x_{k+1}))$ with $1/2$ in $G_{\hat{f}}(y_k)$, $G_{\hat{f}}$ no longer depends on $\hat{f}$. Denote it by $\bar{G}_n(y_k)$; the approximate NPMLE is the left derivative of LCM($\bar{G}_n$). Hence,

$$\bar{f}_n(x_k)=\min_{-1\leqslant i<k}\ \max_{k\leqslant j\leqslant m}\frac{\bar{G}_n(y_j)-\bar{G}_n(y_i)}{y_j-y_i}, \tag{2}$$

which requires no iteration. Let $F_n(x_k)$ be the empirical distribution function at x_k, $k=0,\dots,m$, then

$$\bar{G}_n(y_k)=\frac{F_n(x_k)+F_n(x_{k+1})}{2}. \tag{3}$$

Theorem 2.1. *If $h=h_n\to 0$ as $n\to\infty$, then with probability one, $\bar{f}_n(x)\to f(x)$ at all $x>0$, at which f is continuous.*

Proof. Let F denote the distribution function of f. Then $\sup_{x>0}|\bar{G}_n(x) - F(x)| \to 0$ $w.p.1$ by the Glivenko–Cantelli theorem and the continuity of F. Let $\tilde{G}_n$ be the LCM of $\bar{G}_n$ on $[0, y_m]$ and $\tilde{G}_n(x) = 1$ for $x > y_m$. Then

$$\sup_{x>0} |\tilde{G}_n(x) - F(x)| \leqslant \sup_{0<x\leqslant y_m} |\bar{G}_n(x) - F(x)| + (1 - F(y_m)) \to 0 \ w.p.1$$

using Marshall's lemma and $F(y_m) \to 1$. The theorem follows easily since the convergence of concave functions implies the convergence of their derivatives wherever the derivative of the limit is continuous, see Rockafellar (1970, p. 248). □

As noted by Woodroofe and Sun (1993), Grenander's estimate is not consistent at $x = 0$. For the approximate NPMLE, $\bar{f}_n(0)$ is consistent under mild conditions on h. In the proof, essential use is made of the strong approximation theorem of Komlós et al. (1975). On a sufficiently rich probability space, there are Brownian bridges B_n for which

$$F_n(x) = F(x) + \frac{1}{\sqrt{n}} B_n(F(x)) + \mathrm{O_P}\left(\frac{\log n}{n}\right). \tag{4}$$

Further, B_n may be written as $B_n(t) = W_n(t) - tW_n(1)$, where W_n is a standard Brownian motion.

Theorem 2.2. *If $h \to 0$ as $n \to \infty$ and $nh/\log n \to \infty$, then $\bar{f}_n(0)$ is consistent.*

Proof. It is easy to see that

$$|\bar{f}_n(0) - f(0)| \leqslant \max_{k\geqslant 0} \left| \frac{\bar{G}_n(y_k) - F(y_k)}{y_k} \right| + \left| \max_{k\geqslant 0} \frac{F(y_k)}{y_k} - f(0) \right|.$$

Since F is concave,

$$f(0+) = \lim_{h\to 0} \max_{k\geqslant 0} \frac{F(y_k)}{y_k}.$$

For the first term, write $\bar{G}_n(y_k)$ in terms of F_n as in (3), and apply the strong approximation theorem (4) and a Taylor expansion of $F(x)$ at y_k. After some algebra,

$$\left| \frac{\bar{G}_n(y_k) - F(y_k)}{y_k} \right| \leqslant \left| \frac{W_n(F(x_k)) + W_n(F(x_{k+1}))}{2y_k\sqrt{n}} \right| + \mathrm{O_P}\left(\frac{\log n}{hn}\right) + \mathrm{o_P}(1).$$

It is easily seen that

$$\left| \frac{W_n(F(x_k))}{2y_k\sqrt{n}} \right| \leqslant \frac{f(0)}{\sqrt{n}} \left| \frac{W_n(F(x_k))}{F(x_k)} \right|.$$

Here

$$\max_{k\geqslant 1} \left| \frac{W_n(F(x_k))}{F(x_k)} \right| \leqslant \sup_{t>F(h)} \left| \frac{W_n(t)}{t} \right| = \mathrm{O_p}\left(\sqrt{\frac{1}{h} \log\log \frac{1}{h}} \right),$$

where the last step uses the law of the iterated logarithm for Brownian motion. Since $nh/\log(n) \to \infty$, it follows that

$$\max_{k\geqslant 0}\left|\frac{\bar{G}_n(y_k)-F(y_k)}{y_k}\right| \leqslant \mathrm{O_P}\left(\sqrt{\frac{\log n}{hn}}\right)+\mathrm{o_p}(1),$$

completing the proof. □

3. Localization

For the rest of the paper, suppose that the underlying non-increasing density function f is differentiable in a small neighborhood of $x_0>0$, at which point the density is estimated, and $f'\leqslant -b<0$ in the neighborhood. Next, we establish a local property of $\bar{f}_n(x_0)$: With large probability, $\bar{f}_n(x_0)$ can be determined by the LCM of $\bar{G}_n$ restricted to a small interval around x_0. For each n, let k_n be an integer satisfying $x_{k_n}<x_0\leqslant x_{k_n+1}$.

Lemma 3.1. *Let*

$$\bar{g}_n(x_{k_n})=\frac{\bar{G}_n(y_{k_n})-\bar{G}_n(y_{k_n-1})}{y_{k_n}-y_{k_n-1}}.$$

If $nh^3\to\infty$, then

$$\lim_{n\to\infty} P\{\bar{f}_n(x_{k_n})=\bar{g}_n(x_{k_n})\}=1. \tag{5}$$

Proof. By algorithm (2), $\bar{f}_n(x_{k_n})=\bar{g}_n(x_{k_n})$ is equivalent to

$$\bar{G}_n(y_j)-\bar{G}_n(y_{k_n-1})\leqslant(j-k_n+1)(\bar{G}_n(y_{k_n})-\bar{G}_n(y_{k_n-1}))\quad\text{for all } j\geqslant k_n$$

and

$$\bar{G}_n(y_{k_n})-\bar{G}_n(y_i)\geqslant(k_n-i)(\bar{G}_n(y_{k_n})-\bar{G}_n(y_{k_n-1}))\quad\text{for all } i<k_n.$$

Let

$$A_{jk_n}=\bar{G}_n(y_j)-\bar{G}_n(y_{k_n})-(j-k_n)(\bar{G}_n(y_{k_n})-\bar{G}_n(y_{k_n-1})),$$

where $j=-1,0,1,\ldots,$ then,

$$\bar{f}_n(x_{k_n})=\bar{g}_n(x_{k_n})\ \Leftrightarrow\ A_{jk_n}\leqslant 0\quad\text{for all } j. \tag{6}$$

It is enough to show that

$$\lim_{n\to\infty} P\left\{\max_{j\geqslant -1} A_{jk_n}\leqslant 0\right\}=1.$$

First we consider $j>k_n$. Let δ be so small that f is continuously differentiable on $(x_{k_n}-2\delta, x_{k_n}+2\delta)$ and $f'(x)\leqslant -b<0$ for $|x-x_{k_n}|<2\delta$. For $0<y_j-y_{k_n}<\delta$, a Taylor series expansion of $F(x)$ at $x=x_{k_n}$, shows the existence of positive constants c_0 and c_1 for which

$$E(A_{jk_n})\leqslant -c_0 b(j-k_n)^2h^2$$

and

$$Var(A_{jk_n}) \leqslant \frac{c_1(j-k_n)^2 h}{n}.$$

Let $c = c_1/c_0^2$. Then, by Chebyshev's inequality,

$$\sum_{\{j:0<y_j-y_{k_n}<\delta\}} P\{A_{jk_n} \geqslant 0\} \leqslant \frac{c}{b^2 n h^3} \sum_{\{j:0<y_j-y_{k_n}<\delta\}} \frac{1}{(j-k_n)^2},$$

which approaches zero as $n \to \infty$.

For $y_j - y_{k_n} \geqslant \delta$, $\bar{G}_n(x) \to F(x)$ almost surely for all $x > 0$, as $n \to \infty$. It follows that

$$\max_{\{j:y_j-y_{k_n}>\delta\}} \frac{A_{jk_n}}{y_j - y_{k_n}} \to \sup_{\{x':x' \geqslant x_0+\delta\}} \frac{F(x') - F(x_0)}{x' - x_0} - f(x_0) \quad \text{a.s.},$$

which is negative since F is concave and strictly concave near x_0. So,

$$P\left\{ \max_{\{j:y_j-y_{k_n}>\delta\}} A_{jk_n} > 0 \right\} \to 0.$$

We have shown that $\lim_{n\to\infty} P\{\max_{j>k_n} A_{jk_n} \leqslant 0\} = 1$. The same result holds for $j < k_n$ by similar arguments, and the Lemma follows by (6). □

For $h = \mathrm{O}(n^{-1/3})$, we show a similar local property: With high probability, $\bar{f}_n(x_{k_n})$ can be determined by the LCM of $\bar{G}_n$ on a finite interval $[x_{k_n} - Mn^{-1/3}, x_{k_n} + Mn^{-1/3}]$ for some integer M. Let $\gamma_n = \bar{f}_n(x_{k_n}) - f(x_{k_n})$ and

$$Z_n^*(x) = \bar{G}_n(y_{k_n} + x) - \bar{G}_n(y_{k_n}) - x f(x_{k_n}), \tag{7}$$

then

$$\gamma_n = \min_{0<r} \max_{0 \leqslant s} \frac{Z_n^*(s) - Z_n^*(-r)}{s+r},$$

where r and s are multiples of h. The following lemma gives the convergent rate of γ_n when $h = \mathrm{O}(n^{-1/3})$, and allows us to consider the process Z_n^* for the values x in an $\mathrm{O_P}(n^{-1/3})$ neighborhood of 0. The proof of the lemma is virtually the same as the proof of the assertion in p. 217 of Kim and Pollard (1990).

Lemma 3.2. *If $h = \mathrm{O}(n^{-1/3})$, then $\gamma_n = \mathrm{O_P}(n^{-1/3})$ and*

$$\lim_{M\to\infty} \lim_{n\to\infty} P\left\{ \gamma_n = \min_{0<r\leqslant Mn^{-1/3}} \max_{0\leqslant s\leqslant Mn^{-1/3}} \frac{Z_n^*(s) - Z_n^*(-r)}{s+r} \right\} = 1, \tag{8}$$

where M is some integer.

Proof. Let

$$\Gamma_n(x) = \bar{G}_n(y_{k_n} + x) - \bar{G}_n(y_{k_n}) - x\bar{f}_n(x_{k_n}).$$

Then $\Gamma_n(x)$ is the distance between $\bar{G}_n(y_{k_n} + x)$ and $\bar{G}_n(y_{k_n}) + x\bar{f}(x_{k_n})$. It is easy to see that $\Gamma_n(x)$ is maximized at two points, say $x = -L_n$ and $x = R_n$; $y_{k_n} - L_n$ and

$y_{k_n} + R_n$ are the nearest points to y_{k_n} such that LCM($\bar{G}_n$) and $\bar{G}_n$ coincide. Moreover, $\Gamma_n(-L_n) = \Gamma_n(R_n)$. At $x = -L_n$ and R_n, Z_n^* and the LCM of Z_n^* also coincide. So γ_n is determined by Z_n^* for the values of $x \in [-L_n, R_n]$, and it suffices to show that $L_n = \mathrm{O_P}(n^{-1/3})$ and $R_n = \mathrm{O_P}(n^{-1/3})$.

Let

$$\begin{aligned} g_n(y,x) = {} & \tfrac{1}{2} I(x_{k_n} < y \leqslant x_{k_n+1}) + I(x_{k_n+1} < y \leqslant x_{k_n+\lfloor x/h \rfloor}) \\ & + \tfrac{1}{2} I(x_{k_n+\lfloor x/h \rfloor} < y \leqslant x_{k_n+1+\lfloor x/h \rfloor}) - x f(x_{k_n}), \end{aligned} \tag{9}$$

where I is the indicator function. Then

$$\frac{1}{n} \sum_{i=1}^{n} (g_n(X_i,x) - Eg_n(\cdot,x)) = \Gamma_n(x) + x\gamma_n - Eg_n(\cdot,x). \tag{10}$$

Since $g_n(y,x)$ satisfies the conditions on Lemma 4.1 of Kim and Pollard (1990), for each $\varepsilon > 0$,

$$|\Gamma_n(x) + x\gamma_n - Eg_n(\cdot,x)| = \left| \frac{1}{n} \sum_{i=1}^{n} (g_n(X_i,x) - Eg_n(\cdot,x)) \right| \leqslant \varepsilon |x|^2 + \mathrm{O_P}(n^{-2/3}).$$

By the Taylor expansion of $Eg_n(\cdot,x)$ for x near 0,

$$Eg_n(\cdot,x) = f(x_{k_n}) \left(\left\lfloor \frac{x}{h} \right\rfloor h - x \right) + \frac{1}{2} f'(x_{k_n}) \left(\left\lfloor \frac{x}{h} \right\rfloor + \frac{1}{2} \right)^2 h^2 + \mathrm{O_P}(n^{-2/3}) + \mathrm{o}(x^2).$$

Since the derivative of f is non-zero in a small neighborhood of $x = x_{k_n}$, there exist constants $0 < c_2 < c_1$, such that

$$\begin{aligned} & -c_1 x^2 - x\gamma_n + f(x_{k_n}) \left(\left\lfloor \frac{x}{h} \right\rfloor h - x \right) - \mathrm{O_P}(n^{-2/3}) \\ & \quad \leqslant \Gamma_n(x) \leqslant -c_2 x^2 - x\gamma_n + f(x_{k_n}) \left(\left\lfloor \frac{x}{h} \right\rfloor h - x \right) + \mathrm{O_P}(n^{-2/3}). \end{aligned}$$

Note both L_n and R_n are positive, so, either $-c_2 L_n^2 + L_n\gamma_n < 0$ or $-c_2 R_n^2 - R_n\gamma_n < 0$, and $\lfloor x/h \rfloor h - x < 0$, hence, $\max_x \Gamma_n(x) \leqslant \mathrm{O_P}(n^{-2/3})$ and for some $c_0 > 0$,

$$\max_x \Gamma_n(x) \geqslant \Gamma_n \left(- \left\lfloor \frac{\gamma_n}{2c_1 h} \right\rfloor h \right) \geqslant c_0 \gamma_n^2 - \mathrm{O_P}(n^{-2/3}).$$

So $\gamma_n = \mathrm{O_P}(n^{-1/3})$. It is easy to see that $L_n = \mathrm{O_P}(n^{-1/3})$ follows from

$$0 = \Gamma_n(0) \leqslant \Gamma_n(-L_n) \leqslant -c_2 \left(L_n - \frac{\gamma_n}{2c_2} \right)^2 + \mathrm{O_P}(n^{-2/3}).$$

Similarly, we can show $R_n = \mathrm{O_P}(n^{-1/3})$. □

4. Asymptotic distributions

With the preparation in Sections 2 and 3, we are able to study the asymptotic distribution of $\bar{f}_n$ under different bandwidths h. The bandwidth h varies with n, and $h \to 0$ as $n \to \infty$, but it is denoted by h rather than $h(n)$ or h_n for simplicity.

4.1. $n^{1/3}h \to \infty$

As before, let $x_0 \in (0,\infty)$ and k_n be such that $x_{k_n} < x_0 \leqslant x_{k_n+1}$ and recall the standing assumptions on f. Also let $\alpha_n = (x_{k_n+1} - x_0)/h$.

Theorem 4.1. *If $nh^3 \to \infty$, $nh^5 \to 0$, and f is twice differentiable near x_0, then*

$$(c_n nh)^{1/2}(\bar{f}_n(x_0) - f(x_0)) \Rightarrow N(0, f(x_0)),$$

where $c_n = 2/(1-\alpha_n+\alpha_n^2)$.

Proof. Define

$$\bar{g}_n(x_0) = \alpha_n \bar{g}_n(x_{k_n}) + (1-\alpha_n)\bar{g}_n(x_{k_n+1}).$$

Then $E(\bar{g}_n(x_0)) = g_n(x_0)$, where

$$g_n(x_0) = \alpha_n g_n(x_{k_n}) + (1-\alpha_n) g_n(x_{k_n+1})$$

and

$$g_n(x_{k_n}) = \frac{F(x_{k_n+1}) - F(x_{k_n-1})}{2h}.$$

By a Taylor series expansion, $g_n(x_0) = f(x_0) + \mathrm{O}(h^2)$. So, by the assumption $nh^5 \to 0$,

$$(c_n nh)^{1/2}(\bar{g}_n(x_0) - f(x_0)) = (c_n nh)^{1/2}(\bar{g}_n(x_0) - g_n(x_0)) + \mathrm{o}(1).$$

Since

$$\begin{aligned}\bar{g}_n(x_0) = \frac{1}{2h}[&\alpha_n(F_n(x_{k_n}) - F_n(x_{k_n-1})) + (F_n(x_{k_n+1}) - F_n(x_{k_n}))\\ &+(1-\alpha_n)(F_n(x_{k_n+2}) - F_n(x_{k_n+1}))],\end{aligned}$$

variance of $\bar{g}_n(x_0)$ is $f(x_0)/(c_n nh) + \mathrm{o}((nh)^{-1})$. It is clear that $\bar{g}_n(x_0)$ is asymptotically normal, and the theorem follows by Lemma 3.1. □

4.2. $h \sim cn^{-1/3}$

In this sub-section, suppose that $h = cn^{-1/3}$ for some constant c and that f is twice differentiable near a fixed point $x_0 \in (0,\infty)$. Define k_n by $x_{k_n} < x_0 \leqslant x_{k_n+1}$ as above. Then

$$\begin{aligned}\bar{f}_n(x_0) - f(x_0) = {}&\alpha_n(\bar{f}_n(x_{k_n}) - f(x_{k_n}))\\ &+(1-\alpha_n)(\bar{f}_n(x_{k_n+1}) - f(x_{k_n+1})) + \mathrm{O}(h^2),\end{aligned} \tag{11}$$

where $\alpha_n = (x_{k_n+1} - x_0)/h$. Let $\bar{g}_{n,M}(x_{k_n})$ be the left derivative of $\mathrm{LCM}(\bar{G}_n)$ at y_{k_n}, where the LCM is restricted to the interval $(y_{k_n} - Mh, y_{k_n} + Mh)$, and $y_{k_n} - Mh > 0$. Then

$$\bar{g}_{n,M}(x_{k_n}) = \min_{0<i\leqslant M} \max_{0\leqslant j\leqslant M} \frac{\bar{G}_n(y_{k_n+j}) - \bar{G}_n(y_{k_n-i})}{(j+i)h},$$

and by Lemma 3.2,

$$\lim_{M\to\infty} \lim_{n\to\infty} P\{\bar{f}_n(x_{k_n}) = \bar{g}_{n,M}(x_{k_n})\} = 1. \tag{12}$$

For any integer ℓ, define

$$Z_n(\ell) = n^{2/3}(\bar{G}_n(y_{k_n+\ell}) - \bar{G}_n(y_{k_n}) - f(x_{k_n})\ell h). \tag{13}$$

Then

$$n^{1/3}(\bar{f}_n(x_{k_n}) - f(x_{k_n})) = \min_{i>0} \max_{j\geqslant 0} \frac{Z_n(j) - Z_n(-i)}{c(j+i)}.$$

Let

$$V_{nM} = \min_{0<i\leqslant M} \max_{0\leqslant j\leqslant M} \frac{Z_n(j) - Z_n(-i)}{c(j+i)}.$$

We prove the following results.

Lemma 4.2. *If $h = cn^{-1/3}$, then for any integer M,*

$$V_{nM} \Rightarrow V_M$$

as $n \to \infty$, where

$$V_M = \min_{0<i\leqslant M} \max_{0\leqslant j\leqslant M} \frac{Z(j) - Z(-i)}{c(j+i)}$$

and

$$Z(\ell) = \tfrac{1}{2}\sqrt{cf(x_0)}\,(W(\ell+1) + W(\ell) - W(1)) + \tfrac{1}{2}f'(x_0)c^2\ell^2 + \tfrac{1}{2}f'(x_0)c^2\ell$$

for integer ℓ. Here W is a standard two-sided Brownian motion originating from zero, i.e. $W(0) = 0$ and $E[(W(t) - W(s))^2] = |t-s|$.

Proof. Eqs. (3) and (13) lead to

$$Z_n(\ell) = \tfrac{1}{2}n^{2/3}[F_n(x_{k_n+\ell+1}) + F_n(x_{k_n+\ell}) - F_n(x_{k_n+1}) - F_n(x_{k_n}) - 2f(x_{k_n})\ell h].$$

Here

$$F(x_{k_n+\ell+1}) + F(x_{k_n+\ell}) = F(x_{k_n+1}) + F(x_{k_n}) + 2f(x_{k_n})\ell h + f'(x_{k_n})h^2(\ell^2+\ell) + o(h^2) \tag{14}$$

by a Taylor series expansions of F. So, by the strong approximation (4), $Z_n(\ell)$ can be written as

$$Z_n(\ell) = A_1(\ell) + A_2(\ell),$$

where

$$A_2(\ell) = \tfrac{1}{2}f'(x_{k_n})c^2(\ell^2+\ell) + o_P(1)$$

and

$$A_1(\ell) = \tfrac{1}{2}n^{1/6}[B_n(F(x_{k_n+\ell+1})) + B_n(F(x_{k_n+\ell})) - B_n(F(x_{k_n+1})) - B_n(F(x_{k_n}))].$$

Write $B_n(F(x)) = W_n(F(x)) - F(x)W_n(1)$, where W_n is a standard Brownian motion, and let

$$W_n^*(t) = (f(x_{k_n})h)^{-1/2}(W_n(F(x_{k_n+t})) - W_n(F(x_{k_n}))),$$

then

$$A_1(\ell) = \tfrac{1}{2}\sqrt{cf(x_{k_n})}(W_n^*(\ell+1) + W_n^*(\ell) - W_n^*(1)) - R_n(\ell),$$

where

$$R_n(\ell) = \tfrac{1}{2}n^{1/6}[(2f(x_{k_n})\ell h + f'(x_{k_n})h^2(\ell^2+\ell) + \mathrm{o}(h^2))W_n(1)] \to 0.$$

Therefore

$$Z_n(\ell) = \tfrac{1}{2}\sqrt{cf(x_{k_n})}(W_n^*(\ell+1) + W_n^*(\ell) - W_n^*(1)) + \tfrac{1}{2}f'(x_{k_n})c^2(\ell^2+\ell) + \mathrm{o}_{\mathrm{P}}(1)$$

for $-M \leqslant \ell \leqslant M$. It follows easily that the joint distribution of $(Z_n(-M),\ldots,Z_n(M))$ converges to that of $(Z(-M),\ldots,Z(M))$. Define a function $\psi : \mathscr{R}^{2M+1} \to \mathscr{R}$, by

$$\psi(X) = \min_{0<i\leqslant M}\ \max_{0\leqslant j\leqslant M} \frac{x(j)-x(-i)}{c(j+i)}$$

for $X = (x(-M),\ldots,x(M)) \in \Re^{2M+1}$. Then ψ is a continuous mapping. Therefore, $\psi(Z_n(-M),\ldots,Z_n(M)) \Rightarrow \psi(Z(-M),\ldots,Z(M))$. That is, $V_{nM} \Rightarrow V_M$, as $n\to\infty$. □

Theorem 4.3. *Let*

$$V = \min_{i>0}\ \max_{j\geqslant 0} \frac{Z(j)-Z(-i)}{c(j+i)},$$

where $Z(\ell)$ as defined in Lemma 4.2.

If $h = cn^{-1/3}$, then

$$n^{1/3}(\bar{f}_n(x_{k_n}) - f(x_{k_n})) \Rightarrow V.$$

Proof. Let

$$U_n(x_{k_n}) = n^{1/3}(\bar{f}_n(x_{k_n}) - f(x_{k_n})).$$

It is easy to see that $V_{nM} = n^{1/3}(\bar{g}_{n,M}(x_{k_n}) - f(x_{k_n}))$. Then (12) leads to

$$\lim_{M\to\infty}\ \lim_{n\to\infty} P\{U_n(x_{k_n}) - V_{nM} = 0\} = 1.$$

Since $V_{nM} \Rightarrow V_M$ by Lemma 4.2 and $\lim_{M\to\infty} V_M = V$, the result follows from Theorem 4.2 of Billingsley (1968). □

If f is twice differentiable near x_0, by Eq. (11), we have

$$U_n(x_0) = \alpha_n U_n(x_{k_n}) + (1-\alpha_n)U_n(x_{k_n+1}) + \mathrm{O}(h).$$

So, for the limiting distribution of $U_n(x_0)$, we need to explore more about that of $U_n(x_{k_n+1})$ and its relationship to $U_n(x_{k_n})$.

Corollary 4.4. *Let*

$$V' = \min_{i\geqslant 0}\ \max_{j>0} \frac{Z(j)-Z(-i)}{c(j+i)}$$

and α_{n_r} be a subsequence of α_n such that $\alpha_{n_r} \to \alpha$ for some $\alpha \in (0,1]$, then

$$U_{n_r}(x_0) \Rightarrow \alpha V + (1-\alpha)(V' - cf'(x_0)).$$

Proof. It is obvious that $\bar{f}_n(x_{k_n+1})$ is the right derivative of LCM$(\bar{G}_n)$ at $y = y_{k_n}$, therefore,

$$\bar{f}_n(x_{k_n+1}) = \min_{i \geqslant 0} \max_{j > 0} \frac{\bar{G}_n(y_{k_n+j}) - \bar{G}_n(y_{k_n-i})}{(i+j)h}.$$

Hence,

$$U_n(x_{k_n+1}) = n^{1/3}(\bar{f}_n(x_{k_n+1}) - f(x_{k_n})) + n^{1/3}(f(x_{k_n}) - f(x_{k_n+1}))$$

$$= \min_{i \geqslant 0} \max_{j > 0} \frac{Z_n(j) - Z_n(-i)}{c(j+i)} - cf'(x_{k_n}),$$

where $Z_n(\ell)$ is defined by (13). Same arguments for the proof of Theorem 4.3 lead to $U_n(x_{k_n+1}) \Rightarrow V' - cf'(x_0)$, where cV' is the right derivative of LCM(Z) at $\ell = 0$. The result follows by continuous mapping theorem. □

4.3. $h = \mathrm{o}(n^{-1/3})$

For the case of $h = \mathrm{o}(n^{-1/3})$, the approximate NPMLE has the same distribution as that of Grenander's estimator. Recall that $x_{k_n} < x_0 \leqslant x_{k_n+1}$. Since $\bar{f}_n$ and f are continuous,

$$\bar{f}_n(x_0) - f(x_0) = \bar{f}_n(x_{k_n}) - f(x_{k_n}) + \mathrm{O_P}(h).$$

We need to establish the limit distribution of $\gamma_n = \bar{f}_n(x_{k_n}) - f(x_{k_n})$.

Let

$$Z_n(t) = n^{2/3} Z_n^*(tn^{-1/3}),$$

where Z_n^* is defined in (7). The following lemma gives the weak convergence of $\{Z_n(t)\colon |t| \leqslant M\}$ for each integer $M < \infty$.

Lemma 4.5. *If* $h = \mathrm{o}(n^{-1/3})$, *then*

$$\{Z_n(t)\colon |t| \leqslant M\} \Rightarrow \{Z(t)\colon |t| \leqslant M\},$$

where $Z(t) = f'(x_0)t^2/2 + \sqrt{f(x_0)}W(t)$, *$W$ is a standard Brownian motion.*

Proof. This follows from Theorem 4.7 of Kim and Pollard (1990). For $|t| \leqslant M$, $Z_n(t) = n^{-1/3} \sum g_n(X_i, tn^{-1/3})$, where $g_n(y,x)$ is defined in (9). Since $g_n(y,x)$ is bounded for $|x| \leqslant Mn^{-1/3}$, most of the conditions are easy to check. Here we check only condition (ii) of Lemma 4.5 and (iii) of Lemma 4.6 of Kim and Pollard (1990). For $s < t$,

$$n^{1/3} Eg_n(\cdot, sn^{-1/3}) g_n(\cdot, tn^{-1/3})$$

$$= n^{1/3} \left[F\left(x_{k_n} + \left\lfloor \frac{sn^{-1/3}}{h} \right\rfloor h \right) - F(x_{k_n}) \right] + \mathrm{o}(1)$$

$$= f(x_0)s + \mathrm{o}(1).$$

This implies (ii) of Lemma 4.5 of Kim and Pollard (1990), and gives $\sqrt{f(x_0)}W(t)$ in $Z(t)$. For $\theta_1 < \theta_2$ near 0,

$$\begin{aligned} &E|g_n(y,\theta_1) - g_n(y,\theta_2)| \\ &\leqslant \left[F\left(x_{k_n} + \left\lfloor \frac{\theta_2}{h} \right\rfloor h\right) - F\left(x_{k_n} + \left\lfloor \frac{\theta_1}{h} \right\rfloor h\right)\right] + (\theta_2 - \theta_1)f(x_{k_n}) + \mathrm{o}(1) \\ &= \mathrm{O}(|\theta_2 - \theta_1|). \end{aligned}$$

So, condition (iii) of Lemma 4.6 of Kim and Pollard (1990) follows. □

By the same argument given in the proof of Theorem 4.3, we can get the following theorem.

Theorem 4.6. *If $h = \mathrm{o}(n^{-1/3})$, then*

$$n^{-1/3}(\bar{f}_n(x_0) - f(x_0)) \Rightarrow \min_{r>0} \max_{s \geqslant 0} \frac{Z(s) - Z(-r)}{s + r}.$$

This limiting distribution is same as that of the Grenander's MLE. It is studied in Groneboom (1989) and Groeneboom et al. (1999). Comparing with the result of Bickel and Fan (1996), this result requires weaker assumptions on the class width h for grouped MLE to have the same limiting distribution as the Grenander's MLE.

References

Bickel, P.J., Fan, J., 1996. Some problems on the estimation of unimodal densities. Statist. Sinica 6, 23–45.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., 1992. Distance Sampling. Chapman & Hall, London.

Devroye, L., 1987. A Course in Density Estimation. Birkhauser, Basel.

Grenander, U., 1985. On the theory of mortality measurement. part ii. Skand. Akt 39, 125–153.

Groeneboom, P., 1985. Estimating a monotone density. In: Le Cam, L.M., Olshen, R.A. (Eds.), Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, Vol. II, pp. 539–555.

Groneboom, P., 1989. Brownian motion with a parabolic drift and airy functions. Probab. Theory Related Fields 81, 79–109.

Groeneboom, P., Houghiemstra, G., Lopuhau, H.P., 1999. Asymptotic normality of the l^1 error of the Grenander estimator. Ann. Statist. 27, 1316–1347.

Kim, J., Pollard, D., 1990. Cube root asymptotics. Ann. Statist. 18, 191–219.

Komlós, J., Major, P., Tusnády, G., 1975. An approximation of partial sums of independent r.v.'s and the sample d.f. Z. Wahrsch. Verw. Gebiete 32, 111–131.

Prakasa Rao, B.L.S., 1969. Estimation of a unimodal density. Sankhya A 31, 23–36.

Robertson, T., Wright, F.T., Dykstra, R.L., 1989. Order Restricted Statistical Inference. Wiley, New York.

Rockafellar, R.T., 1970. Convex Analysis. Princeton University Press, Princeton.

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

Woodroofe, M., Zhang, R., 1999. Isotonic estimation for grouped data. Statist. Probab. Lett. 45, 41–47.