Contributions to the Theory of Estimation of Monotone and Unimodal Densities

by

Yazhen Wang

Technical Report No. 325
September 1991

Department of Statistics
University of California
Berkeley, California 94720

Contributions to the Theory of Estimation of Monotone and Unimodal Densities

by

Yazhen Wang

Department of Statistics,
University of California at Berkeley,
Berkeley, CA 94720

September, 1991

Abstract

The motivation of this paper is to investigate the asymptotic behavior of Grenander estimator locally as well as globally when the underlying densities have both flat and nonflat parts. This paper consists of two parts. In the first part I consider the problem of estimating a monotone density. In the second part I use the Plug-in MLE to estimate the a unimodal density with unknown mode.

For the monotone density the limit distribution of the estimator at a point has been derived. Under L_1-norm the asymptotic distribution is the sum of a normal distribution and maxima of Brownian bridge.

For the unimodal density with unknown mode, I show that, except for the mode, the Plug-in MLE will eventually agree with the estimator when the mode is known. However it blows up at the mode, which causes the problem of "spiking" near the mode. However, from limit distribution of the L_1 -norm, it seems that this spiked estimator behaves very well globally. Moreover, whether the mode is known or not the L_1-norm remains the same up to the first order.

1 Introduction

Let $\mathcal{F}$ be the class of nonincreasing right continuous densities on the interval $[0, \infty)$. It was shown by Grenander (1956) that the nonparametric maximum likelihood estimator f_n of a density f under the order restriction that it belongs to $\mathcal{F}$ is given by the right slope of the concave majorant $\hat{F}_n$ of the empirical distribution function F_n. Since a unimodal density f with mode m is nondecreasing on the left side of the mode and nonincreasing on the right side of the mode, by use of results about the monotone density, we can derive that the maximum likelihood estimator is the left slope of the greatest convex minorant of empirical distribution F_n in $(-\infty, m)$ and the right slope of the least concave majorant of F_n in (m, ∞). For a discussion of this result and more genral results in isotonic regression, see Robertson et al. (1988).

In 1969 Prakasa Rao first studied the asymptotic distribution of these estimators at a point, see [17]. By use of a jump process Groeneboom simplied Prakasa Rao's proof and gave the limit distribution of the L_1-norm, see [10]. However these results are derived under some kind of conditions like strict monotonicity. They don't have much to say about the limit distribution of f_n when f is not strictly monotone, especially when f has both flat and nonflat parts. The motivation of this paper is to try to investigate the asymptotic behavior of f_n locally as well as globally when f has both flat and nonflat parts.

This paper consists of two parts. In the first part, including section 2, 3 and 4, we deal with the problem of estimating a monotone density. In section 1 and 2, I'll consider the behavior of f_n at a point t_0. The rate depends on the smoothness of f at this point. If f has the kth nonzero derivative, the rate is $\frac{k}{2k+1}$ and the limit distribution relates to Brownian motion. When f is flat near this point, the rate is $\frac{1}{2}$ and the limit distribution relates to Brownian bridge, see theorem 1 and 2 in section 1. For the boundary points of flat ranges or support of f, the asymptotic rates are given in theorem 3 and 4 in section 2.

Groeneboom proved that L_1-norm is asymptotic normal when f is strictly monotone and maximum of Brownian bridge if f is uniform , see [10]. In section 3, I'll show that these are two extreme cases. Generally the asymptotic distribution of L_1-norm is sum of this two kinds of distributions, which are presented in theorem 5.

In the second part, including section 4, I'll try to estimate a unimodal density with unknown mode by use of the Plug-in MLE. If the estimator of the mode is consistent, it will eventually agree with the estimator with known mode except for the mode, which has the same asymptotic distribution as the monotone case. At mode the plug-in MLE is spiked. However it seems that the problem of "spiking" at the mode does not affect the L_1-norm very much. The asymptotic distribution of L_1-norm has been derived. Moreover, it turns out that whether the mode is known or not the L_1-norm is the same up to the first order.

2 Limit Distribution

Let $X_1, \cdots, X_n \sim f$, f be the nonincreasing in $[0,\infty)$, f_n is Grenander estimator. The asymptotic distribution of $f_n(t_0)$ has been studied by Prakasa Rao(1969) and Groeneboom(1984) when $f'(t_0) \neq 0$. In this section I'll study how it behaves asymptotically when f is flat near t_0 or $f^{(k)}(t_0) \neq 0$ for some k.

If f is smooth at $t_0 \in (0,\infty)$, then

(A) f is flat in a neighborhood of t_0, let $[a,b]$ be the flat part containning t_0, i.e. $[a,b] = \{t : f(t) = f(t_0)\}$.

(B) $f(t) - f(t_0) \sim d(t-t_0)^k$ near t_0 for some $k > 0$ and $d < 0$. Set $f^{(k)}(t_0) = d$, $\alpha = \frac{k}{2k+1}$, $\beta = \frac{1}{2k+1}$.

First I reduced the problem of finding the distribution of f_n to that of the locations of maxima of the process $(F_n(t) - at, t \geq 0)$, $a > 0$. Let

$$U_n(a) = sup\{t : F_n(t) - at \text{ is maximal }\} \tag{1}$$

Then we have, with probability one,

$$f_n(t) \leq a \Longleftrightarrow U_n(a) \leq t \tag{2}$$

Theorem 1 *For the case (A), set* $[a,b] = \{t : f_n(t) = f_n(t_0)\}$. *Then for* $a < t_0 < b$,

$$\sqrt{n}(f_n(t_0) - f(t_0)) \Longrightarrow \hat{S}_{a,b}(t_0) \tag{3}$$

Here $\hat{S}_{a,b}(t)$ *is the slope at* $F(t)$ *of the least concave majorant in* $[F(a), F(b)]$ *of a standard Brownian Bridge in* $[0,1]$.

Corollary 1 *(Groeneboom) If f is the uniform density on $[0,1]$, then*

$$\sqrt{n}(f_n(t_0) - f(t_0))$$

converges in distribution to the slope of the least concave majorant of Brownian Bridge at t_0.

Theorem 2 *For case (B), set $\delta = [\frac{f^k(t_0)|f^{(k)}(t_0)|}{(k+1)!}]^{\frac{1}{2k+1}}$, then*

$$n^{\alpha}\delta^{-1}(f_n(t_0) - f(t_0)) \Longrightarrow V_k(0) \tag{4}$$

Where $V_k(t)$ is the slope at t of the least concave majorant of $(W(z) - |z|^{k+1}, z \in (-\infty, \infty))$, and W is standard two-sided Brownian motion on $(-\infty, \infty)$ with $W(0) = 0$.

Corollary 2 *(Prakasa Rao) If $f'(t_0) \neq 0$, then*

$$n^{\frac{1}{3}}|f'(t_0)f(t_0)|^{-\frac{1}{3}}(f_n(t_0) - f(t_0))$$

converges in distribution to $V(0)$, where $V(a)$ is the location of the maximum of the process $(W(z) - (z-a)^2, z \in R)$, W is standard two-sided Brownian motion with $W(0) = 0$.

Remark. 1. The rate of f_n at a point depends on the smoothness of f at this point, if f has a kth derivative the rate is $\frac{k}{2k+1}$ which tends to $\frac{1}{2}$ corresponding to the case of f is flat at this point. However the limit distribution of $V_k(0)$ does not tend to that of $\hat{S}$.

2. Since the distribution of $V_k(0)$ is symmetric and $\hat{S}$ is a.s positive, so f_n is asymptotically unbiased at point where f is strictly monotone and overestimates on the flat range asymptotically.

3 Asymptotic Behaviors at Boundary Points

From section 1 we know the asymptotic distribution of $f_n(t_0)$ when t_0 is a 'regular' point, i.e t_0 the interior point of a flat range or has nonzero derivative of some order, but this still leave out some points, like the boundary points of the flat ranges or the support of f, in this section I'll study the behaviors of f_n at those points.

Theorem 3 *Let $[a,b]$ be a flat range of f as in the case (A). If a and b are not the boundary of the support of f and are points at which f has left or right nonzero derivatives, then $f_n(a)$ always overestimates $f(a)$ and $f_n(b)$ underestimates $f(b)$ asymptotically. Moreover*

(a) $n^{\frac{1}{3}}(f_n(a) - f(a)) = O_p(1)$, $n^{\frac{1}{3}}(f_n(b) - f(b)) = O_p(1)$.

(b) *For* $\mu > \frac{1}{3}$, $n^{\mu}(f_n(a) - f(a)) \xrightarrow{P} \infty$, $n^{\mu}(f_n(b) - f(b)) \xrightarrow{P} \infty$.

Remark. 1. At points in the left or right of a, f_n has rate $\frac{1}{3}$ or $\frac{1}{2}$ respectively, so it seem resonable that the rate of f_n at the boundary of the flat range is $\frac{1}{3}$.

2. The hehavior of f_n in an $n^{-\frac{1}{3}}$-shinking neighborhood of a or b are similar.

3. If $f_-^{(k)}(a) \neq 0$ for some $k > 1$ and $f_-^{(i)}(a) = 0$ for $1 \leq i < k$, then results similar to those of section 2 can be formulated.

Theorem 4 *$f_n(0)$ always overestimates $f(0)$ and is not consistent. Moreover*

$$f_n(0) - f(0) \xrightarrow{P} \infty$$

4 Asymptotic Distribution Under the L_1-Norm

Groeneboom [10] has shown that the L_1-norm $\|f_n - f\|_1$ is asymptotically normal if f is strictly monotone, but didn't know how it behaves if f has both flat and nonflat parts. In this section I'll answer this question.

Let f be nonincreasing, concentrated on a bounded interval and have bounded piecewise continuous second derivatives. Then we have

Theorem 5 *Let $0 \leq a_1 < b_1 < \cdots < a_r < b_r \leq B$. If f is flat in each $[a_j, b_j]$, $1 \leq j \leq r$, and strictly decreasing in $\Theta = [0,B] - \bigcup_{j=1}^{r}[a_j, b_j]$. Then*

$$\sqrt{n}(\|f_n - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0, \sigma^2 P(X \in \Theta)) + \sum_{j=1}^{r} Y_j \tag{5}$$

where $N(0, \sigma^2 P(X \in \Theta))$ and $(Y_i, 1 \leq i \leq r)$ are independent and

$$Y_j \sim 2 \sup_{a_j \leq s \leq b_j} B(F(s)) - B(F(a_j)) - B(F(b_j))$$

B is standard Brownian Bridge in $[0,1]$, *and*

$$\theta = 2E|V(0)| \int_{\Theta} |\frac{1}{2}f'(t)f(t)|^{\frac{1}{3}} dt \approx .82 \int |\frac{1}{2}f'(t)f(t)|^{\frac{1}{3}} dt$$
$$\sigma^2 = 8 \int_0^{\infty} Covar(|V(0)|, |V(\xi) - \xi|) d\xi \approx .17$$

Corollary 3 *(Groeneboom) If f is strictly monotone then*

$$\sqrt{n}(\|f_n - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0, \sigma^2) \tag{6}$$

where θ and σ^2 are defined in theorem 5.

Corollary 4 *If f is a uniform distribution on* $(0,1)$ *then*

$$\sqrt{n}\|f_n - f\|_1 \Longrightarrow 2 \sup_{0 \le s \le 1} B(s) \tag{7}$$

Remark. 1. For the two extreme cases, the limit distributions are respectively normal and maximum of Brownian bridge. Combining thess two distributions together we get the asymptotic distributions for general case.

2. The flat parts don't make any contribution to θ. The normal variance is that σ^2 time the proportion $P(X \in \Theta)$ corresponding to the strict monotone parts.

4. The normal mean is zero, but all Y's are positive, this agree with the results in section 2 that f_n is unbiased at the points where f is strictly monotone and overestimates on the flat ranges.

5 Estimating the Unimodal Density

Now suppose $X_1, \cdots, X_n \sim f$, where f is unimodal with mode m. We know that one of the difficuities of estimating of a unimodal density with known or unknown mode is the problem of 'spiking' near mode. In this section I'll study asymptotic behavior under $L_1 - norm$. Thoughout this section I'll assume that $d_1|t-m|^q \le f(m) - f(t) \le d_2|t-m|^q$ as $t \to m$ for some $q > 0$ and $d_2 \ge d_1 > 0$. First let consider the case that the mode m is known.

Let $\hat{F}_{n,m}$ be the greatest convex minorant of F_n in $(-\infty, m)$ and the least concave majorant of F_n in $[m, \infty)$, $f_{n,m}^-$ is the left slope of $\hat{F}_{n,m}$ in $(-\infty, m)$,

$f^+_{n,m}$ is the right slope of $\hat{F}_{n,m}$ in $[m,\infty)$. Set $f_{n,m} = f^-_{n,m}$ on $(-\infty,m)$, $f_{n,m} = f^+_{n,m}$ on $[m,\infty)$, which is the MLE subject to the restriction that f is a unimodal density with known mode m, see [19, 17]. For $t \neq m$, the asymptotic distribution of $f_{n,m}(t)$ is similar to that when the density is monotone, see Robertson (1967) and Prakasa Rao(1969). But at the mode by theorem 4 we have

Theorem 6 *$f_{n,m}$ is not consistent and always overestimate. Moreover*

$$f_{n,m}(m) - f(m) \to \infty$$

Remark. This is the source of the problem of "spiking".

Theorem 7 *Suppose f is concentrated on $[A,B]$ and has piecewise continuous second derivative. Let $A \leq a_1 < b_1 < \cdots < a_r < b_r \leq B$, $r \leq \infty$. If f is flat on $[a_j,b_j]$ for all $1 \leq j \leq r$ and strictly monotone on $\Theta = [A,B] - \bigcup_{j=1}^r [a_j,b_j]$, then*

$$\sqrt{n}(\|f_{n,m} - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0,\sigma^2 P(X \in \Theta)) + \sum_{i=1}^{r} Y_j \tag{8}$$

where $N(0,\sigma^2 P(X \in \Theta))$ and $(Y_i, 1 \leq i \leq r)$ are independent.

$$Y_j \sim 2 \sup_{a_j \leq s \leq b_j} B(F(s)) - B(F(a_j)) - B(F(b_j))$$

where B is standard Brownian Bridge in $[0,1]$, and

$$\theta = 2E|V(0)| \int |\frac{1}{2} f'(t) f(t)|^{\frac{1}{3}} dt \approx .82 \int |\frac{1}{2} f'(t) f(t)|^{\frac{1}{3}} dt$$
$$\sigma^2 = 8 \int_0^\infty Covar(|V(0)|, |V(\xi) - \xi|) d\xi \approx .17$$

Corollary 5 *If f is strictly monotone in $[A,m)$ and $(m,B]$, then*

$$\sqrt{n}(\|f_{n,m} - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0,\sigma^2) \tag{9}$$

where θ and σ^2 are defined in theorem 7.

A more interesting problem, both from a mathematical and practical point of view, is the problem of estimating a unimodal density with unknown mode. Here our approach to this problem is to find a direct estimate of the mode location first and then an indirect estimate of the density, say using the techniques described in the known mode case.

Let m_n be an estimate of the mode m, $f_{n,m}$ be the Grenander estimator with known mode m. Then replace m by m_n in $f_{n,m}$ to get the plug-in estimate f_{n,m_n}, I'll study the asymptotic behavior of f_{n,m_n} globally as well as locally.

Theorem 8 *If m_n is consistent then for all $\varepsilon > 0$, eventually f_{n,m_n} and $f_{n,m}$ will agree on $(m-\varepsilon, m+\varepsilon)^c$. Hence for $t \neq m$, f_{n,m_n} has the same asymptotic distribution as that of $f_{n,m}$.*

Remark. From theorem 8 we can easily get theorem 1 and the conjecture in Bickel and Fan [2]

$$\sup_{|t-m|\geq\varepsilon} |f_{n,m_n}(t) - f_{n,m}(t)| = o_p(n^{-\frac{1}{3}})$$

Theorem 9 *Under the conditions of theorem 7, if $m_n = m + O_p(n^{-\frac{1}{2q+1}})$, then*

$$\sqrt{n}(\|f_{n,m_n} - f\|_1 - \|f_{n,m} - f\|_1) \xrightarrow{P} 0 \tag{10}$$

That is, the L_1 − norm of f_{n,m_n} (the estimator with unknown mode) is equivalent to that of $f_{n,m}$ (the estimator with known mode) up to order $n^{-\frac{1}{2}}$. Hence the asymptotic distribution of the L_1 − norm of f_{n,m_n} is the same as that of $f_{n,m}$ in theorem 7, i.e,

$$\sqrt{n}(\|f_{n,m} - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0, \sigma^2 P(X \in \Theta)) + \sum_{i=1}^{r} Y_j$$

Corollary 6 *If $f''(m) \neq 0$ and m_n is $n^{\frac{1}{5}}$-consisent, then*

$$\sqrt{n}(\|f_{n,m} - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0, \sigma^2 P(X \in \Theta)) + \sum_{i=1}^{r} Y_j$$

Corollary 7 *If f is strictly monotone in $[A, m)$ and $(m, B]$, then*

$$\sqrt{n}(\|f_{n,m_n} - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0, \sigma^2)$$

where θ and σ^2 are defined in theorem 9.

Remark. 1. There are many methods to estimate the mode. For example the Chernoff's mode estimator can achieve the rate required by theorem 9.

2. Intuitively we may think that the problem of "spiking" becomes worse for estimating a unimodal density with unknown mode, but globally, from the L_1-norm point view, it makes no diffence up to order $n^{-\frac{1}{2}}$ whether the mode is known or not.

6 Proofs

6.1 Proof of theorem 1

By the relation (2) we have

$$P(\sqrt{n}(f_n(t_0) - f(t_0)) \leq x) = P(U_n(f(t_0) + n^{-\frac{1}{2}}x) \leq t_0) \tag{11}$$

From the definition of U_n in (1)

$$\begin{aligned} U_n(f(t_0) + n^{-\frac{1}{2}}x) &= sup\{s : F_n(s) - (f(t_0) + n^{-\frac{1}{2}}x)s \text{ is maximal }\} \\ &= sup\{s : \sqrt{n}(F_n(s) - F(s)) + \sqrt{n}(F(s) - f(t_0)s) - xs \text{ is maximal }\} \end{aligned}$$

By Komlós et al. (1975),

$$\sqrt{n}(F_n(s) - F(s)) = B_n(F(s)) + O_p(n^{-\frac{1}{2}} \log n),$$

where $(B_n, n \in N)$ is a sequence of Brownian Bridges, constructed on the same space as the F_n. So the limit distribution of $U_n(f(t_0) + n^{-\frac{1}{2}}x)$ is the same as that of the location of the maximum of the process $(B_n(F(s)) + \sqrt{n}(F(s) - f(t_0)s) - xs, s \geq 0)$. Noting $F(s)$ is concave and linear in $[a, b]$, then

$$F(s) = F(a) + f(t_0)(s - a) \text{ for } s \in [a, b]$$

and

$$F(s) - f(t_0)(s - a) < F(a) \text{ for } s \notin [a, b]$$

Hence the location of the maximum of $(B_n(F(s)) + \sqrt{n}(F(s) - f(t_0)s) - xs, s \geq 0)$ behaves asymptotically as that of

$$\{B(F(s)) - xs, a \leq s \leq b\} = \{B(F(a) + f(t_0)(s - a)) - xs, a \leq s \leq b\}$$

where B is a standard Brownian bridge in $[0,1]$. Combining the equation (11)

$$P(\sqrt{n}(f_n(t_0) - f(t_0)) \le x) \to$$
$$P(\text{ the location of the maximum of } \{B(F(s)) - xs, a \le s \le b\} \le t_0)$$
$$= P(\hat{S}_{a,b}(t_0) \le x)$$

by the definition of $\hat{S}$. □

6.2 Proof of theorem 2

By relation (2) we have

$$P(n^{\alpha}\delta^{-1}(f_n(t_0) - f(t_0) \le x) = P(U_n(f(t_0) + n^{-\alpha}\delta x) \le t_0) \qquad (12)$$

From the definition of U_n in (1)

$$U_n(f(t_0) + n^{-\alpha}\delta x) = sup\{s : F_n(s) - (f(t_0) + n^{-\alpha}\delta x)s \text{ is maximal }\}$$
$$= sup\{s : \sqrt{n}(F_n(s) - F(s)) + \sqrt{n}(F(s) - (f(t_0) + n^{-\alpha}\delta x)s \text{ is maximal }\}$$

by Komlós et al. (1975),

$$\sqrt{n}(F_n(s) - F(s)) = B_n(F(s)) + O_p(n^{-\frac{1}{2}}\log n),$$

where $(B_n, n \in N)$ is a sequence of Brownian Bridges, constructed on the same space as the F_n. So the limit distribution of $n^{\beta}(U_n(f(t_0) + n^{-\alpha}\delta x) - t_0)$ is the same as that of $n^{\beta}(U(f(t_0) + n^{-\alpha}\delta x) - t_0)$, where $U(u)$ is the locations of the maximum of the process $(B(F(s)) + \sqrt{n}(F(s) - us), s \ge 0)$, and B is a standard Brownian Bridge on $[0,1]$. Since

$$B(F(s)) + \sqrt{n}(F(s) - (f(t_0) + n^{-\alpha}\delta x)s)$$
$$= B(F(t_0) + f(t_0)(s - t_0) + O((s - t_0)^{k+1})) + \sqrt{n}\frac{f^{(k)}(t_0)}{(k+1)!}(s - t_0)^{k+1} +$$
$$\sqrt{n}(F(t_0) - n^{-\alpha}\delta x t_0) - n^{\frac{1}{2}-\alpha}\delta x(s - t_0) + o(\sqrt{n}(s - t_0)^{k+1})$$

then $B(F(s)) + \sqrt{n}(F(s) - (f(t_0) + n^{-\alpha}\delta x)s)$ achieves its maximum within an $O_p(n^{-\frac{1}{2(2k+1)}})$-neighborhood of t_0, so

$$B(F(s)) + \sqrt{n}(F(s) - (f(t_0) + n^{-\alpha}\delta x)s)$$
$$= B(F(t_0) + f(t_0)(s - t_0)) + \sqrt{n}\frac{f^{(k)}(t_0)}{(k+1)!}(s - t_0)^{k+1} +$$
$$\sqrt{n}(F(t_0) - n^{-\alpha}\delta x t_0) - n^{\frac{1}{2}-\alpha}\delta x(s - t_0) + o_p(n^{-\frac{1}{2(2k+1)}})$$

Hence the location of maximum of the process

$$B(F(s)) + \sqrt{n}(F(s) - (f(t_0) + n^{-\alpha}\delta x))$$

behaves asymptotically like that of the process

$$\begin{aligned}
&B(F(t_0) + f(t_0)(s - t_0)) - B(F(t_0)) + n^{\frac{1}{2}} \frac{f^{(k)}(t_0)}{(k+1)!}(s - t_0)^{k+1} - \\
&n^{\frac{1}{2}-\alpha}\delta x(s - t_0) \\
&= n^{-\frac{1}{2(2k+1)}} \sqrt{f(t_0)c}(\tilde{B}(z) - z^{k+1} - xz)
\end{aligned} \tag{13}$$

here

$$\begin{aligned}
c &= [\frac{\sqrt{f(t_0)}(k+1)}{|f^{(k)}(t_0)|}]^{\frac{2}{2k+1}} \\
z &= \frac{n^{\frac{1}{2k+1}}(s - t_0)}{c} \\
\tilde{B}(z) &= [B(F(t_0) + f(t_0)(s - t_0)) - B(F(t_0))]/[n^{-\frac{1}{2(2k+1)}}\sqrt{f(t_0)c}]
\end{aligned}$$

By the fact that Brownian bridge behaves locally as Brownian motion, so the location of the maximum of process in (13) behaves asymptotically the same way as that of $(W(z) - z^{k+1} - xz, z \in R)$, where $(W(z), z \in R)$ is two-sided Brownian motion with $W(0) = 0$. Therefore $n^{\beta}(U_n(f(t_0) + n^{-\alpha}\delta x) - t_0)$ converges in distribution to the location of the maximum of the process $W(z) - z^{k+1} - xz$. Combining with equation (12) we get

$$\begin{aligned}
&P(n^{\alpha}\delta^{-1}(f_n(t_0) - f(t_0)) \le x) = P(U_n(f(t_0) + n^{\alpha}\delta x) \le t_0) \\
&= P(n^{\beta}(U_n(f(t_0) + n^{\alpha}\delta x) - t_0) \le 0) \\
&\to P(\text{ the location of maximum of } W(z) - z^{k+1} - xz \le 0) \\
&= P(V_k(0) \le x)
\end{aligned}$$

by the definition of V_k. □

6.3 Proof of corollary 2

Similarly to (2) we can relate of least concave majorant and the location of maximum of a process. Note that the process $(V(a) - a, a \in R)$ is stationary, see Groeneboom [10]. Then the distribution of the slope $V_2(0)$ at 0 of the least concave majorant of process $(W(z) - z^2, z \in R)$ is the same as that of $V(0)$. □

6.4 Proof of theorem 3

Since the proofs are similar. I'll only give the results for a. From the proofs of theorem 1 and 2 we have

$$P(n^{\mu}(f_n(a)-f(a))\le x)=P(U_n(f(a)+n^{-\mu}x)\le a)$$

and the limit distribution of $U_n(f(a)+n^{-\mu}x)$ is the same as that of the location of maximum of the process $(B(F(s))+\sqrt{n}(F(s)-f(a)s)-n^{\frac{1}{2}-\mu}xs, s\ge 0)$, where B is standard Brownian bridge. Noting that $F(s)$ is concave and linear in $[a,b]$, then

$$F(s)=F(a)+f(a)(s-a) \text{ for } s\in[a,b]$$

and

$$F(s)-f(a)(s-a)<F(a) \text{ for } s\notin[a,b]$$

So the process must achieves its maximum near a or b according to $x>0$ or $x<0$.

(1) If $x<0$. For $0\le\mu<\frac{1}{2}$. Then the process $(B(F(s))+\sqrt{n}(F(s)-f(a)s)-n^{\frac{1}{2}-\mu}xs, s\ge 0)$ must achieve its maximum near b asymptotically, so $P(n^{\mu}(f_n(a)-f(a))\le x)\to 0$, i.e $f_n(a)$ always overestimates $f(a)$ asymptotically.

(2) If $x>0$, for $\mu=\frac{1}{3}$, $a\le s\le b$,

$$\begin{aligned}&B(F(s))+\sqrt{n}(F(s)-f(a)s)-n^{\frac{1}{2}-\mu}xs\\&=B(F(a)+f(a)(s-a))-n^{\frac{1}{2}-\mu}xs\end{aligned}$$

we see that asymptotically the maximum of the process in $[a,b]$ is no more than

$$B(F(a))-n^{\frac{1}{6}}xa+O_p(n^{-\frac{1}{6}})$$

For $s=a-n^{-\frac{1}{3}}$,

$$\begin{aligned}&B(F(s))+\sqrt{n}(F(s)-f(a)s)-n^{\frac{1}{6}}xs\\&=B(F(a)+f(a)(s-a))+\frac{1}{2}n^{\frac{1}{2}}f'_{-}(a)(s-a)^2-\\&n^{\frac{1}{6}}x(s-a)-n^{\frac{1}{6}}xa\\&=B(F(a))-n^{\frac{1}{6}}xa+O_p(n^{-\frac{1}{6}})+O(n^{-\frac{1}{6}})+xn^{-\frac{1}{6}}\\&=B(F(a))-n^{-\frac{1}{6}}xa+n^{-\frac{1}{6}}(x+O_p(1))\end{aligned}$$

So for sufficiently large x, the process $(B(F(s))+\sqrt{n}(F(s)-f(a)s)-n^{\frac{1}{6}}xs, s \geq 0)$ attains its maximum at points on the left side of a asymptotically, therefore

$$\lim_{x\to\infty} \liminf_{n\to\infty} P(n^{\mu}(f_n(a) - f(a)) \leq x) = 1$$

(3) If $x > 0$, for $\frac{1}{3} < \mu < \frac{1}{2}$, for $s \leq a$,

$$\begin{aligned}
&B(F(s)) + \sqrt{n}(F(s) - f(a)s) - n^{\frac{1}{2}-\mu}xs \\
&= B(F(a) + f(a)(s-a)) + \frac{1}{2}n^{\frac{1}{2}}f'_{-}(a)(s-a)^2 - \\
&n^{\frac{1}{2}-\mu}x(s-a) - n^{\frac{1}{2}-\mu}xa \\
&\leq B(F(a)) - n^{\frac{1}{2}-\mu}xa + O_p(n^{-\frac{1}{6}})
\end{aligned}$$

Let $1 - 2\mu - \eta \leq \nu \leq \frac{1}{3} - \eta$ for some small $\eta > 0$, $s = a + n^{-\nu}$, then

$$\begin{aligned}
&B(F(s)) + \sqrt{n}(F(s) - f(a)s) - n^{\frac{1}{2}-\mu}xs \\
&= B(F(a) + f(a)(s-a)) - n^{\frac{1}{2}-\mu}xs \\
&= B(F(a) + f(a)n^{-\nu}) - n^{\frac{1}{2}-\mu}xa + xn^{\frac{1}{2}-\mu-\nu} \\
&= B(F(a)) - n^{\frac{1}{2}-\mu}xa + n^{-\frac{\nu}{2}}\tilde{B}(a)(1 + o_p(1))
\end{aligned}$$

where $\tilde{B}(F(a)) = n^{\frac{\nu}{2}}[B(F(a) + f(a)n^{-\nu}) - B(F(a))]$ behaves like Brownian motion near $F(a)$. However the probability that a Brownian motion ever takes a positive value in $n^{-\frac{1}{3}+\eta}(F(a+n^{-\frac{1}{3}-\eta})-F(a), F(a+n^{2\mu-1-\eta})-F(a)) = (f(a), n^{2\mu-\frac{2}{3}}f(a))$ tends to 1, so, with probability one, the process can at least achieve positive value with magnitude $O_p(n^{-\frac{1}{3}-\eta})$ at the right side of a, therefore the process must take its maximum on the right of a asymptotically, hence for any $x > 0$,

$$P(n^{\mu}(f_n(a) - f(a)) \leq x) \to 0$$

Combining (1) and (2), (1) and (3) we get (a) and (b) respectively. □

6.5 Proof of theorem 4

From the proof of theorem 2 we can see

$$P(f_n(0) - f(0) \leq x) = P(U_n(f(0) + x) \leq 0)$$

and the limit distribution of $U_n(f(0)+x)$ is the same as that of the location of the maximum of the process $B(F(s))+\sqrt{n}(F(s)-(f(0)+x)s)$. Set $\nu \geq \frac{1}{2}-\eta$ for some small $\eta > 0$, $s = n^{-\nu}$, then

$$\begin{aligned}
&B(F(s)) + \sqrt{n}(F(s) - (f(0) + x)s) \\
&= B(f(0)s) + \sqrt{n}[\frac{f_+^{(k)}(0)}{(k+1)!}s^{k+1} - xs] + o(\sqrt{n}s^{k+1}) \\
&= B(f(0)n^{-\nu}) + O_p(n^{\frac{1}{2}-\nu})
\end{aligned}$$

However $n^{\frac{\nu}{2}}B(f(0)n^{-\nu})$ behaves like Brownian motion near 0, and a Brownian motion must take a positive value in $(0, f(0))$, so the process must take its maximum on the right of a asymptotically, hence for any $x > 0$, $P(f_n(0) - f(0) \leq x) \to 0$ i.e $f_n(0) - f(0) \to \infty$. □

6.6 Proof of theorem 5

Since the proofs are same except for details. I'll only give the proof for the case $r = 1$ which is stated in the following theorem.

Theorem 10 *If f has support $[0, a_2]$, and f is flat in $[0, a_1]$ and strictly decreasing in $[a_1, a_2]$, then*

$$\sqrt{n}(\|f_n - f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0, \sigma^2(F(a_2) - F(a_1))) + Y \tag{14}$$

Where $N(0, \sigma^2(F(a_2) - F(a_1)))$ and Y are independent.

$$Y \sim 2 \sup_{0 \leq s \leq a_1} B(F(s)) - B(F(a_1))$$

B is standard Brownian Bridge in $[0, 1]$, and

$$\begin{aligned}
&\theta = 2E|V(0)| \int_{a_1}^{a_2} |\frac{1}{2}f'(t)f(t)|^{\frac{1}{3}}dt \approx .82 \int_{a_1}^{a_2} |\frac{1}{2}f'(t)f(t)|^{\frac{1}{3}}dt \\
&\sigma^2 = 8\int_0^\infty Covar(|V(0)|, |V(\xi) - \xi|)d\xi \\
&\approx .17
\end{aligned}$$

Proof of theorem 10. In order to simplify the proof we suppose $f'_+(a_1) \neq 0$. (Otherwise we can find the first k such that $f_+^{(k)}(a_1) \neq 0$.) Let $f_{n,1}$,

$f_{n,c,1}$ and $f_{n,c,2}$ be the slope of the least concave majorant of F_n in $[a_1, a_2]$, $[a_1 + 3cn^{-\frac{1}{3}}, a_2]$ and $[0, a_1 + 2cn^{-\frac{1}{3}}]$ respectively. The proof of theorem 10 is divided into several lemmas.

Lemma 1

$$\lim_{c\to\infty}\liminf_{n\to\infty} P(f_n = f_{n,c,1} = f_{n,1} \ on \ [a_1 + 9cn^{-\frac{1}{3}}, a_2]) = 1 \qquad (15)$$

$$\lim_{c\to\infty}\liminf_{n\to\infty} P(f_n = f_{n,c,2} \ on \ [0, a_1]) = 1 \qquad (16)$$

Proof. The proofs are similar and so I only give the arguments for (15).

$$p(n,c) = P\{F_n(z) \le F_n(a_1 + 8cn^{-\frac{1}{3}}) + (z - a_1 - 8cn^{-\frac{1}{3}})(D - An^{-\frac{1}{3}})$$
$$\text{for } z \le a_1 + 3cn^{-\frac{1}{3}}\}$$

where $A = -8cf'_+(a_1) > 0$, $D = f(a_1)$. Similar to the proof of lemma 4.1 in Prakasa Rao [17], it is enough to show

$$\lim_{c\to\infty}\liminf_{n\to\infty} p(n,c) = 1 \qquad (17)$$

But

$$p(n,c) = P\{F_n(y + 3cn^{-\frac{1}{3}}) \le F_n(a_1 + 8cn^{-\frac{1}{3}}) + (y - a_1 - 5cn^{-\frac{1}{3}})$$
$$(D - An^{-\frac{1}{3}}) \text{ for } y \le a_1\}$$
$$= P\{n[F_n(y) - F_n(a_1)] \le n\{[F_n(a_1 + 8cn^{-\frac{1}{3}}) - F_n(a_1)]$$
$$-[F_n(y + 3cn^{-\frac{1}{3}}) - F_n(y)] + (y - a_1 - 5cn^{-\frac{1}{3}})(D - An^{-\frac{1}{3}})\} \text{ for } y \le a_1\}$$
$$\ge P\{n[F_n(y) - F_n(a_1)] \le n\{[F_n(a_1 + 8cn^{-\frac{1}{3}}) - F_n(a_1)]$$
$$-[F_n(y + 3cn^{-\frac{1}{3}}) - F_n(y)] + (y - a_1 - 5cn^{-\frac{1}{3}})(D - An^{-\frac{1}{3}})\} \text{ for } y \le a_1\}$$
$$\ge P\{n[F_n(y) - F_n(a_1)] \le n\{[F_n(a_1 + 8cn^{-\frac{1}{3}}) - F_n(a_1)]$$
$$-F_n(3cn^{-\frac{1}{3}}) + (y - a_1 - 5cn^{-\frac{1}{3}})(D - An^{-\frac{1}{3}})\} \text{ for } y \le a_1\}$$

Set

$$I_n = n\{[F_n(a_1 + 8cn^{-\frac{1}{3}}) - F_n(a_1)] - F_n(3cn^{-\frac{1}{3}}) - 5cn^{-\frac{1}{3}}(D - An^{-\frac{1}{3}})\}$$
$$= \sum_{i=1}^{n}\{1(a_1 < X_i \le a_1 + 8cn^{-\frac{1}{3}}) - 1(0 < X_i \le 3cn^{-\frac{1}{3}})\}$$
$$-5cn^{-\frac{1}{3}}(D - An^{-\frac{1}{3}})$$
$$= n^{\frac{1}{3}}(cA + \sqrt{5cD}V_n)$$

where $EV_n = 0$, $VarV_n = 1 + o(1)$. Then by Chebyshev's inequality,

$$\liminf_{n\to\infty} P(I_n > \frac{1}{2}cAn^{\frac{1}{3}}) \geq \liminf_{n\to\infty} P(V_n > -\frac{cA}{2\sqrt{5cD}})$$
$$\geq 1 - \limsup_{n\to\infty} P(|V_n| > \frac{cA}{2\sqrt{5cD}}) \geq 1 - \limsup_{n\to\infty} \frac{2\sqrt{5cD}\, VarV_n}{cA}$$
$$\geq 1 - \frac{2\sqrt{5cD}}{cA}$$

hence

$$p(n,c) \geq P\{n[F_n(y) - F_n(a_1) \leq \frac{1}{2}cAn^{\frac{1}{3}} + n(y-a_1)(D - An^{-\frac{1}{3}})$$
$$\text{for } y \leq a_1\} - P(I_n \leq \frac{1}{2}cAn^{\frac{1}{3}})$$
$$\equiv p'(n,c) - P(I_n \leq \frac{1}{2}cAn^{\frac{1}{3}})$$

Now we have

$$p'(n,c) \geq P\{n[F_n(y) - F_n(a_1)] \leq \frac{1}{2}cAn^{\frac{1}{3}} + n(y-a_1)(D - An^{-\frac{1}{3}})$$
$$\text{for } y \leq a_1\}$$
$$\geq P\{\sqrt{n}[F_n(y) - F_n(a_1) - F(y) + F(a_1)] \leq \frac{1}{2}cAn^{-\frac{1}{6}}$$
$$+\sqrt{n}[F(a_1) + (y-a_1)(D - An^{-\frac{1}{3}} - F(y))] \text{ for } y \leq a_1\}$$
$$\geq P\{B_n([F(a_1) - F(y)]/F(a_1)) + O_p(n^{-\frac{1}{2}}logn) \leq \frac{1}{2}cAn^{-\frac{1}{6}}$$
$$+\sqrt{n}[F(a_1) - F(y) + (y-a_1)(D - An^{-\frac{1}{3}})] \text{ for } y \leq a_1\}$$
$$\geq P\{B_n([F(a_1) - F(y)]/F(a_1)) + O_p(n^{-\frac{1}{2}}logn) \leq \frac{1}{2}cAn^{-\frac{1}{6}}$$
$$+A(F(a_1) - F(y))n^{\frac{1}{6}}/D + \sqrt{n}[F(a_1) - F(y) + (y-a_1)D](1 - An^{-\frac{1}{3}}/D)$$
$$\text{for } y \leq a_1\}$$
$$\geq P\{B_n([F(a_1) - F(y)]/F(a_1)) + O_p(n^{-\frac{1}{2}}logn) \leq \frac{1}{2}cAn^{-\frac{1}{6}}$$
$$+A(F(a_1) - F(y))n^{\frac{1}{6}}/D \text{ for } y \leq a_1\}$$
$$\geq P\{B_n(u) + O_p(n^{-\frac{1}{2}}logn) \leq \frac{1}{2}cAn^{-\frac{1}{6}} + AF(a_1)un^{\frac{1}{6}}/D \text{ for } 0 \leq u \leq 1\}$$
$$\geq P\{B_n(u) \leq \frac{1}{2}cAn^{-\frac{1}{6}} + AF(a_1)un^{\frac{1}{6}}/D \text{ for } 0 \leq u \leq 1\}$$

Here I have used the fact that

$$\sqrt{n}\{F_n(y) - F_n(a_1) - F(y) + F(a_1)\} = B_n([F(a_1) - F(y)]/F(a_1))$$
$$+O_p(n^{-\frac{1}{2}}\log n)$$

where B_n is Brownian Bridge, by the the Hungarian embedding of Komlós et al. (1975) and

$$(y - a_1)D - F(y) + F(a_1) \geq 0 \text{ for } y \geq a_1$$

by the concavity of F.

It can be shown by the Cameron-Martin-Girsanov formula that the probability that a Brownian Bridge ever crosses the line between the points $(0, x)$ and $(1, y)$ is e^{-2xy}, therefore

$$p'(n,c) \geq 1 - exp\{-cAn^{-\frac{1}{6}}(\frac{1}{2}cAn^{-\frac{1}{6}} + \frac{1}{D}AF(a_1)n^{\frac{1}{6}}) - o(1)$$
$$\geq 1 - exp\{-cA^2F(a_1)/D\} - o(1)$$

hence

$$\liminf_{n\to\infty} p(n,c) \geq \liminf_{n\to\infty} p'(n,c) - \limsup_{n\to\infty} P(I_n \leq \frac{1}{2}cAn^{\frac{1}{3}})$$
$$\geq 1 - exp\{-cA^2F(a_1)/D\} - \frac{2\sqrt{5cD}}{cA}$$
$$\to 1$$

as $c \to \infty$. □

Remark. If we use the Hungarian embedding of Komlós et al.(1975) and the results about the probability that a Brownian Bridge crosses a line as I did in proving the lemma, we can simplify the proof of the key lemma 4.1 in Prakasa Rao [17], which showed that $f_n(t_0)$ depend on only the data in the $n^{-\frac{1}{3}}$-shrinking neighborhood.

Lemma 2

$$\sqrt{n}(\int_{a_1}^{a_2} |f_{n,1}(t) - f(t)|\,dt - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0, \sigma^2(F(a_2) - F(a_1))) \quad (18)$$

Proof. It comes directly from the theorem 2 in Groeneboom[10]. □

Lemma 3 *Suppose a concave function G has a right continuous right derivative g. Then for $t_1 < t_2$,*

$$\int_{t_1}^{t_2} |g(t)|\, dt = 2 \sup_{t_1 \le t \le t_2} G(t) - G(t_1) - G(t_2)$$

Proof. Set $t_0 = inf\{t_1 \le t \le t_2 : g(t) \le 0\}$, and $t_0 = t_2$ if $g > 0$ on $[t_1, t_2]$. then at t_0 G achieves its maximum in $[t_1, t_2]$. Hence

$$\begin{aligned}
&\int_{t_1}^{t_2} |g(t)|\, dt = \int_{t_1}^{t_0} g(t)\, dt - \int_{t_0}^{t_2} g(t)\, dt \\
&= 2G(t_0) - G(t_1) - G(t_2) \\
&= 2 \sup_{t_1 \le t \le t_2} G(t) - G(t_1) - G(t_2) \ \square
\end{aligned}$$

Lemma 4 *For each $c > 0$,*

$$\sqrt{n} \int_0^{a_1} |f_{n,c,2}(t) - f(t)|\, dt \Longrightarrow Y \tag{19}$$

Proof. Let $\hat{F}_{n,c,1}$ be the least concave majorant of F_n in $[0, a_1 + 2cn^{-\frac{1}{3}}]$ and F_n outside $[0, a_1 + 2cn^{-\frac{1}{3}}]$. Set $F(t,n) = F(a_1) + D(t - a_1)$ for $t \in [a_1, a_1 + 2cn^{-\frac{1}{3}}]$, and $F(t,n) = F(t)$ for $t \notin [0, a_1 + 2cn^{-\frac{1}{3}}]$. Then $\sqrt{n}(\hat{F}_{n,c,1}(\cdot) - F(\cdot, n))$ is the least concave majorant of $\sqrt{n}(F_n(\cdot) - F(\cdot, n))$ in $[0, a_1 + 2cn^{-\frac{1}{3}}]$ and $\sqrt{n}(F_n(\cdot) - F(\cdot, n))$ outside $[0, a_1 + 2cn^{-\frac{1}{3}}]$, and

$$\begin{aligned}
&\sqrt{n} \sup_{0 \le t \le a_2} |F(t,n) - F(t)| = \sqrt{n} \sup_{a_1 \le t \le a_1 + 2cn^{-\frac{1}{3}}} |F(t,n) - F(t)| \\
&= \sqrt{n} \sup_{0 \le s \le 2c} |F(a_1 + sn^{-\frac{1}{3}}) - F(a_1) - Dsn^{-\frac{1}{3}}| \\
&= O(n^{-\frac{1}{6}})
\end{aligned}$$

By lemma 3 we have

$$\begin{aligned}
&\sqrt{n} \int_0^{a_1 + 2cn^{-\frac{1}{3}}} |f_{n,c,2}(t) - D|\, dt \\
&= 2\sqrt{n} \sup_{0 \le t \le a_1 + 2cn^{-\frac{1}{3}}} [\hat{F}_{n,c,1}(t) - F(t,n)] - \sqrt{n}[\hat{F}_{n,c,1}(0) - F(0,n)] \\
&- \sqrt{n}[\hat{F}_{n,c,1}(a_1 + 2cn^{-\frac{1}{3}}) - F(a_1 + 2cn^{-\frac{1}{3}}, n)]
\end{aligned}$$

$$
\begin{aligned}
&= 2\sqrt{n} \sup_{0\le t\le a_1+2cn^{-\frac{1}{3}}} [F_n(t) - F(t)] \\
&-\sqrt{n}[F_n(a_1 + 2cn^{-\frac{1}{3}}) - F(a_1 + 2cn^{-\frac{1}{3}})] + o(n^{-\frac{1}{6}}) \\
&\Longrightarrow 2 \sup_{0\le t\le a_1} B(F(t)) - B(F(a_1))
\end{aligned}
$$

Here I have used

$$\sqrt{n}(F_n - F) = B_n(F(\cdot)) + O_p(n^{-\frac{1}{2}}\log n)$$

by Komlós et al. (1975). Again by lemma 3,

$$
\begin{aligned}
&\sqrt{n}\int_{a_1}^{a_1+2cn^{-\frac{1}{3}}} |f_{n,c,2}(t) - D|\,dt \\
&= 2\sqrt{n} \sup_{a_1\le t\le a_1+2cn^{-\frac{1}{3}}} [\hat{F}_{n,c,1}(t) - F(t,n)] - \sqrt{n}[\hat{F}_{n,c,1}(a_1) - F(a_1,n)] \\
&-\sqrt{n}[\hat{F}_{n,c,1}(a_1 + 2cn^{-\frac{1}{3}}) - F(a_1 + 2cn^{-\frac{1}{3}}, n)] \\
&= 2\sqrt{n} \sup_{a_1\le t\le a_1+2cn^{-\frac{1}{3}}} [\hat{F}_{n,c,1}(t) - F(t)] - \sqrt{n}[\hat{F}_{n,c,1}(a_1) - F(a_1)] \\
&-\sqrt{n}[F_n(a_1 + 2cn^{-\frac{1}{3}}) - F(a_1 + 2cn^{-\frac{1}{3}})] + O(n^{-\frac{1}{6}}) \\
&\Longrightarrow 0
\end{aligned}
$$

since

$$\sqrt{n}[\hat{F}_{n,c,1}(\cdot) - F(\cdot)] \Longrightarrow \hat{B}_{0,a_1}(\cdot)$$

where $\hat{B}_{0,a_1}$ is the least concave majorant in $[F(a_1), F(a_2)]$ of Brownian bridge B and B outside $[F(a_1), F(a_2)]$, which is continuous process. This complete the proof of this lemma. □

Lemma 5 *For each $c > 0$,*

$$\sqrt{n}\int_{a_1}^{a_1+9cn^{-\frac{1}{3}}} |f_n(t) - f(t)|\,dt \xrightarrow{P} 0$$

$$\sqrt{n}\int_{a_1}^{a_1+9cn^{-\frac{1}{3}}} |f_{n,1}(t) - f(t)|\,dt \xrightarrow{P} 0$$

Proof. Since the proofs are similar I'll only prove the first result. One has

$$\begin{aligned}
&\sqrt{n}\int_{a_1}^{a_1+9cn^{-\frac{1}{3}}} |f_n(t)-f(t)|\,dt \\
&\leq 9cn^{\frac{1}{6}}|f_n(a_1)-f(a_1+9cn^{-\frac{1}{3}})| \vee |f_n(a_1+9cn^{-\frac{1}{3}})-f(a_1)|.
\end{aligned}$$

So the lemma results directly from

$$\begin{aligned}
f_n(a_1+9cn^{-\frac{1}{3}})-f(a_1) &= O_p(n^{-\frac{1}{3}}) \\
f_n(a_1)-f(a_1+9cn^{-\frac{1}{3}}) &= O_p(n^{-\frac{1}{3}}),
\end{aligned}$$

which is true by the theorem 3 and its remark. □

Now theorem 10 can be proved easily: set

$$\Gamma_{n,c} = \{f_n = f_{n,c,1} = f_{n,1} \text{ on } [a_1+9cn^{-\frac{1}{3}}, a_2], f_n = f_{n,c,2} \text{ on } [0,a_1]\},$$

then on $\Gamma_{n,c}$,

$$\begin{aligned}
&\sqrt{n}\|f_n-f\|_1 - \int_{a_1+9cn^{-\frac{1}{3}}}^{a_2} |f_{n,c,1}(t)-f(t)|\,dt - \int_0^{a_1} |f_{n,c,2}(t)-f(t)|\,dt \\
&= \sqrt{n}\int_{a_1}^{a_1+9cn^{-\frac{1}{3}}} |f_n(t)-f(t)|\,dt
\end{aligned}$$

and

$$\begin{aligned}
&\sqrt{n}\{\int_{a_1}^{a_2} |f_{n,1}(t)-f(t)|\,dt - \int_{a_1+9cn^{-\frac{1}{3}}}^{a_2} |f_{n,c,1}(t)-f(t)|\,dt\} \\
&= \sqrt{n}\int_{a_1}^{a_1+9cn^{-\frac{1}{3}}} |f_{n,1}(t)-f(t)|\,dt
\end{aligned}$$

by lemma 1 and lemma 5 we get

$$\begin{aligned}
&|\sqrt{n}\|f_n-f\|_1 - \int_{a_1+9cn^{-\frac{1}{3}}}^{a_2} |f_{n,c,1}(t)-f(t)|\,dt - \int_0^{a_1} |f_{n,c,2}(t)-f(t)|\,dt| \\
&\xrightarrow{P} 0
\end{aligned}$$

Noting that $f_{n,c,1}$ and $f_{n,c,2}$ are independent, and $\|f_n-f\|_1$ don't depend on c, then by lemma 2 and lemma 4 we can get

$$\sqrt{n}(\|f_n-f\|_1 - \theta n^{-\frac{1}{3}}) \Longrightarrow N(0,\sigma^2)+Y$$

□

6.7 Proof of theorem 7

Noting

$$\|f_{n,m} - f\|_1 = \int_A^m |f^-_{n,m}(t) - f(t)|\,dt + \int_m^B |f^+_{n,m}(t) - f(t)|\,dt$$

since f is nonincreasing and nondecreasing in (m, ∞) and $(-\infty, m)$, then by the results for the monotone density, e.g. theorem 10 and theorem 5, we can get the asymptotic distribution of $\|f_{n,m} - f\|_1$. □

6.8 Proof of theorem 8

For $m_1 \geq m_2$, if $\hat{F}_{n,m_1}$ and $\hat{F}_{n,m_2}$ meet at a point after m_2 or before m_1, by the construction of the least concave majorant, this point must be a observation point and they will be equal after this point. Since m_n is consistent, it is enough to prove that F_{n,m_n} and $F_{n,m}$ will meet between $m+\delta$ and $m+\varepsilon$ for some $\delta < \varepsilon$. Since F is strictly concave in $[m, m+\varepsilon]$, this is a consequence of the following lemma. □

Lemma 6 *If F is strictly concave in (a, b), then*

$$P(\exists a < t < b \ni \hat{F}_{n,m}(t) = F_n(t) \textit{ for all } m < a) \to 1$$

Proof. If $\hat{F}_{n,m}(t) > F_n(t)$ for all $a < t < b$, then $\exists X_i \leq a$ and $X_j \geq b$ such that $\hat{F}_{n,m}$ is linear in (X_i, X_j) and equal to F_n at X_i and X_j. Since $F_n - F = O_p(n^{-\frac{1}{2}})$ uniformly, then $\hat{F}_{n,m} + O_p(n^{-\frac{1}{2}})$ is below the line joining point $(a, F(a))$ and $(b, F(b))$. So does $F_n + O_p(n^{-\frac{1}{2}})$. This implies that F is below the line joining point $(a, F(a))$ and $(b, F(b))$, which contradicts the strict concavity of F in (a, b). □

6.9 Proof of theorem 9

Let Let $\hat{F}_{n,m_n}$ be the greatest convex minorant of F_n in $(-\infty, m_n)$ and the least concave majorant of F_n in $[m_n, \infty)$, $f_{n,c}$ be the right slope of the least concave majorant of F_n in $[m + 2cn^{-\frac{1}{2q+1}}, B)$ and the left slope of the greatest convex minorant of F_n in $(A, m - 2cn^{-\frac{1}{2q+1}})$. Noting that m_n is $n^{\frac{1}{2q+1}}$-consistent, as in the proof of lemma 1 we can show

$$\lim_{c\to\infty} \liminf_{n\to\infty} P\{f_{n,m_n} = f_{n,m} = f_{n,c}$$

$$\text{on } (A, m - 2cn^{-\frac{1}{2q+1}}] \cup [m + 2cn^{-\frac{1}{2q+1}}, B)\} = 1.$$

Now it is enough to show that

$$\sqrt{n}\int_{m-2cn^{-\frac{1}{2q+1}}}^{m+2cn^{-\frac{1}{2q+1}}} (|f_{n,m}(t) - f(t)| + |f_{n,m_n}(t) - f(t)|\, dt \xrightarrow{P} 0.$$

By lemma 3

$$\begin{aligned}
&\sqrt{n}\int_{m}^{m+2cn^{-\frac{1}{2q+1}}} |f_{n,m}(t) - f(t)|\, dt \\
&= \sqrt{n}\int_{m}^{m+2cn^{-\frac{1}{2q+1}}} |f_{n,m}(t) - f(m)|\, dt + O(n^{-\frac{1}{2(2q+1)}}) \\
&= 2\sqrt{n} \sup_{m \le t \le m+2cn^{-\frac{1}{2q+1}}} [\hat{F}_{n,m}(t) - F(t)] - \sqrt{n}[\hat{F}_{n,m}(m) - F(m)] - \\
&\sqrt{n}[\hat{F}_{n,m}(m + 2cn^{-\frac{1}{2q+1}}) - F(m + 2cn^{-\frac{1}{2q+1}})] + O(n^{-\frac{1}{2(2q+1)}})
\end{aligned} \tag{20}$$

and

$$\begin{aligned}
&\sqrt{n}\int_{m_n}^{m+2cn^{-\frac{1}{2q+1}}} |f_{n,m_n}(t) - f(t)|\, dt \\
&= \sqrt{n}\int_{m_n}^{m+2cn^{-\frac{1}{2q+1}}} |f_{n,m_n}(t) - f(m)|\, dt + O(n^{-\frac{1}{2(2q+1)}}) \\
&= 2\sqrt{n} \sup_{m_n \le t \le m+2cn^{-\frac{1}{2q+1}}} [\hat{F}_{n,m_n}(t) - F(t)] - \sqrt{n}[\hat{F}_{n,m_n}(m_n) - F(m_n)] - \\
&\sqrt{n}[\hat{F}_{n,m_n}(m + 2cn^{-\frac{1}{2q+1}}) - F(m + 2cn^{-\frac{1}{2q+1}})] + O(n^{-\frac{1}{2(2q+1)}}).
\end{aligned} \tag{21}$$

By the strong approximation of Komlós et al. (1975), we have that $\sqrt{n}[\hat{F}_{n,m}(s) - F(s)] + O_p(n^{-\frac{1}{2}}\log n)$ is that the least concave majorant of $\sqrt{n}F(s) + B_n(F(s))$ in $[m, \infty)$ minus $\sqrt{n}F(s)$, which is between a Brownian bridge B and its least least concave majorant $\hat{B}_m$ in $[F(m), 1]$, so the sum of (20) and (21) is no more than

$$\begin{aligned}
&2 \sup_{m \le t \le m+2cn^{-\frac{1}{2q+1}}} [\hat{B}_m(F(s)) + \hat{B}_{m_n}(F(s))] - B(F(m)) - B(F(m_n)) - \\
&B(F(m + 2cn^{-\frac{1}{2q+1}})) - B(F(m_n + +2cn^{-\frac{1}{2q+1}})) + o_p(1) \\
&= 4\hat{B}_m(F(m)) - 4B(F(m)) + o_p(1) \\
&= o_p(1).
\end{aligned}$$

Here I have used that $\hat{B}_m$ is continuous process and equal to B outside $[F(m), 1]$.

Similarly I can prove

$$\sqrt{n}\left(\int_{m-2cn^{-\frac{1}{2q+1}}}^{m} |f_{n,m}(t) - f(t)| + \int_{m-2cn^{-\frac{1}{2q+1}}}^{m_n} |f_{n,m_n}(t) - f(t)|\right) dt \xrightarrow{P} 0$$

□

Acknowledgments. I wish to express my gratitute to my thesis advisor Professor Peter Bickel for his guidance and encouragement.

References

[1] Barlow, R., E., Bartholomew, D., J., Bremmer, J., M. and Brunk, H., D.(1972). Statistical Inference under Order Restrictions. New York: Wiley.

[2] Bickel, P.J and Fan, J.(1990). Some Problems on the Eatimation of Densities under Shape Restrictions. Technical Report No. 258, Dept of Statistics, University of California at Berkeley.

[3] Billingsley, P.(1968). Weak Convergence of Probability Measures. New York: Wiley.

[4] Birgé, L.(1989). The Grenander estimator: a Nonasymptoic Approach. Ann. Statist. Vol.17, 1532-1549.

[5] Birgé, L.(1987). Robust Estimation of Unimodal Densities. To appear.

[6] Brunk, H., D.(1970). Estimation of Isotonic Regression. Nonparametric Techniques in Statistical Inference. 177-195 Cambridge Univ. Press.

[7] Grenander,U.(1956). On the Theory of Mortality Measurement, Part II. Skand. Akt. 39, 125-153.

[8] Groeneboom, P.(1983). Concave Majorant of Brownian Motion. Ann. Probab. 11, 1016-1027.

[9] Groeneboom, P and Pyke, R.(1983). Asymptotical Normality of the Statistics Based on the Convex Minorant of Empirical Distribution Functions. Ann. Probab. 11, 328-345.

[10] Groeneboom, P.(1985). Estimating Monotone Density. Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer. 539-555.

[11] Groeneboom, P.(1989). Brownian Motion with a Parabolic Drift and Airy Functions. Z.Wahr.verw.Geb. 81, 79-109.

[12] Ikeda, N. and Watanabe, S(1991). Stochastic Differential Equations and Diffusion Processes. Second edition. Springer-Verlag.

[13] Itô, K and Mckean, H. P., Jr(1965). Diffusion Processes and Their Sample Paths. Springer-Verlag.

[14] Kiefer,J and Wolfowitz, J(1976). Asymptotically Minimax Estimation of Concave and Convex Distribution Functions. Z.Wahr.verw.Geb. 34, 73-85.

[15] Komlós, J., Major, P. and Tusnády, G.(1975). An Approximation of Partial Sums of Independent r.v.'s and the Sample d.f I. Z.Wahr.verw.Geb. 32, 111-131.

[16] Komlós, J., Major, P. and Tusnády, G.(1976). An Approximation of Partial Sums of Independent r.v.'s and the Sample d.f II. Z.Wahr.verw.Geb. 34, 33-58.

[17] Prakasa Rao.(1969). Estimation of a Unimodal Density. Sankhȳa Ser. A 31, 23-36.

[18] Prakasa Rao.(1983). Nonparametric Functional Estimation. New York: Academic Press.

[19] Robertson, T., Wright, F., T. and Dykstra, R., L.(1988). Order Restricted Statistical Inferences. New York: Wiley.

[20] Strook, D. W. and Varadhan, S.R.S(1979). Multidimensional Diffusion Processes. Springer-Verlag.

[21] Wright, F., T.(1981). The Asymptotical Behavior of Monotone Regression Estimates. Ann. Statist. Vol 9, No. 2, 443-448.