Title
Limiting Behavior of Extrema of Certain Conditional Sample Functions

Author
Zhenwei Zhou

Publication Date
2011-10-25

LIMITING BEHAVIOR OF EXTREMA OF CERTAIN CONDITIONAL SAMPLE FUNCTIONS

ZHENWEI ZHOU AND PRANAB K. SEN

ABSTRACT. Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be independent identically distributed as (X, Y) ($X \in \mathcal{R}^+$ and $Y \in \mathcal{R}^q$). Given $Y = y$, X has the conditional distribution $G(x|y)$ (unknown). We investigate the behavior of the extrema of sample functions of the form $\Psi(x, G_{nk}(x|y))$, where $G_{nk}(x|y)$ is the k nearest neighbor (k-NN) estimator of $G(x|y)$. Specifically, we first study the properties of a multiparameter process $W_n = k^{1/2}(G_{nk}(x|y) - G(x|y))$, $(x, y) \in \mathcal{R}^+ \times \mathcal{C}$, where $\mathcal{C}$ is a compact set in $\mathcal{R}^q$. Then under some regular conditions we show that there exists a Gaussian process $W(.)$, such that $\{k^{1/2}(\sup_x \Psi(x, G_{nk}(x|y)) - \sup_x \Psi(x, G(x|y))),\ ; y \in \mathcal{C}\}$ convergences weakly to W. The above problem is motivated by the study of the strength of a bundle of parallel filaments in reliability theory when there are concomitant variates.

1. INTRODUCTION

In reliability theory, there is a considerable literature on the statistical theory of the strength of a bundle of n parallel filaments or fibres. Suppose that a bundle is clamped at each end and elongated by increasing the distance between the clamps, the problem is to determine the probabilistic characteristic of the maximum tensile load that the bundle will sustain in terms of the probabilistic and mechanical characteristics of the individual fibre and the clamping mechanism.

Let $X_{n1} \leq \cdots \leq X_{nn}$ be the n ordered values of $X_1, \cdots, X_n$, representing the strengths (non-negative random variables) of individual filaments in a bundle of n parallel filaments of equal length. If we assume that the force of a fre e load on the bundle is distributed equally on each filament and the strength of an individual filament is independent of the number of filaments in a bundle, then the minimum load B_n beyond which all the filaments of the bundle give away is defined to be the *strength of the bundle.*

1991 *Mathematics Subject Classification.* Primary 62E20, 62G20; secondary 62G05 62N99 .

Key words and phrases. bundle strength of filaments; conditional distribution; conditional sample function; gaussian process; conditional multiparameter process; shrinking neighborhood; extrema.

If a bundle breaks under a load L, then the inequalities $nX_{n1} \leq L, (n-1)X_{n2} \leq L, \cdots, X_{nn} \leq L$ are simultaneously satisfied. Consequently, the bundle strength can be represented as: $B_n = \max\{nX_{n1}, (n-1)X_{n2}, \cdots, X_{nn}\}$ Note that if $\hat{F}_n(x)$ is the empirical distribution function for $X_1, \cdots, X_n$, then $B_n^* \equiv n^{-1}B_n = \sup_{x\geq 0} x[1-\hat{F}_n(x)]$ (see Sue et al.(1970)). The above formulation of B_n is due to Daniels (1945) w ho investigated the probability distribution of B_n^* (B_n) and established its asymptotic normality by very elaborate and complicated analysis. Later on, Sen et al.(1973), Phoenix and Taylor (1973), Smith (1982),Daniels (1985, 1989), and oth ers studied some extensions of Daniels' model. We will propose another model by allowing some mechanism variations and investigate its probabilistic properties. Specifically, since the strength of a bundle of fibres depends on the materials, the individu al length, diameter and so on, it is natural to take these kind of variations into account. Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be independently distributed as $(X, Y) \in \mathcal{R}^{p+1}$, where $X \in \mathcal{R}^1$, $Y \in \mathcal{R}^p$, and X_i $(i = 1, \cdots, n)$ represents the strength of the i-th fibre, Y_i represents the corresponding individual variations. Instead of $\hat{F}_n(x)$, we will use the conditional empirical distribution $\hat{G}_n(x|Y)$—the estimator of conditional distributi on $G(x|Y)$, and want to study the behavior of the functionals of the form

$$\theta(\mathbf{y}) = \sup_{x\in\mathcal{R}^+} \Psi(x, G(x|\mathbf{y})\,),\ \mathbf{y} \in \mathcal{C}\ (compact) \subset \mathcal{R}^q.$$

Obviously $\theta(\mathbf{y})$ has a natural estimator

$$\hat{\theta}_n(\mathbf{y}) = \sup_{x\geq 0} \Psi(x, G_n(x|\mathbf{y})),\ \mathbf{y} \in \mathcal{C}, \tag{1.1}$$

where $G_n(x|\mathbf{y})$ is a estimator of conditional distribution $G(x|\mathbf{y})$.
Remark: if $\Psi(x,z) = x(1-z)$ and $\hat{F}_n(x|Y)$ is replaced by $\hat{F}_n(x)$, then (1.1) has the form $\sup_{x\geq 0} x(1-\hat{F}_n(x))$ which is the same as B_n^*.

We will explore the behavior of stochastic process

$$V_n(\mathbf{y}) = k_n^{1/2}[\hat{\theta}_n(\mathbf{y}) - \theta(\mathbf{y})],$$

for $\mathbf{y}$ belonging to some neighborhood of $\mathbf{Y_0} \in \mathcal{C}$ where k_n, tending to infinity, is a constant sequence. The study of $V_n(\mathbf{y})$, in turn, may require to investigate the properties of the multi-dimensional parameter conditional empirical process

$$W_n(x, \mathbf{y}) = k_n^{1/2}[G_{nk_n}(x|\mathbf{y}) - G(x|\mathbf{y})],\ (x, \mathbf{y}) \in \mathcal{R}^+ \times \mathcal{C}.$$

The next section deals with the empirical process. First we define the estimator $G_{nk_n}(x|\mathbf{y})$ of $G(x|\mathbf{y})$ by the nearest neighbor method. If, say, k_n observations are used in estimating $G(x|\mathbf{y})$, and $k_n = [n^{4/5}t]$, $0 < a \leq t \leq b < \infty$, then in section 2.1 we show that (in one dimensional case) for fixed $\mathbf{y}$, $W_n(x, \mathbf{y})$ converges weakly to a Gaussian process jointly in x and t as its parameters. In section 2.2 we show that when $\mathbf{y}$ varies in a neighborhood of a fixed point $\mathbf{y_0}$, process $W_n(x, \mathbf{y})$ converges weakly to a Gaussian process in $\mathbf{y}$ (as its parameter). The behavior of process $V_n(\mathbf{y})$, as $\mathbf{y}$ varies in a neighborhood of a fixed point $\mathbf{y_0}$ is investigated in section 3. Under some conditions such defined $V_n(\mathbf{y})$ converges weakly to a Gaussian process too. All the proofs are given in section 4.

2. Asymptotic Properties of the Conditional Empirical Process.

In this section we mainly investigate the properties of the stochastic process

$$W_n(x, \mathbf{y}) = k_n^{1/2}[G_{nk_n}(x|\mathbf{y}) - G(x|\mathbf{y})],\ (x, \mathbf{y}) \in \mathcal{R}^+ \times \mathcal{C}.$$

First we define the estimator $G_{nk_n}(x|\mathbf{y})$ of $G(x|\mathbf{y})$ by the nearest neighbor method.

Suppose we have a collection of observations $\{(X_i, \mathbf{Y}_i), i \geq 1\}$ which are independent identically distributed as $(X, \mathbf{Y})$ ($X \in \mathcal{R}^1$ representing the strengths and $\mathbf{Y} \subset \mathcal{R}^q$ representing the corresponding variations). Given $\mathbf{Y} = \mathbf{y}$, we set $Z_i = \|\mathbf{Y}_i - \mathbf{y}\|$ $i = 1, \cdots, n$, where $\|.\|$ is a norm on $\mathcal{R}^q$ (e.g., Euclidean) so that we have the usual topology. Therefore, we have the data transform: $(X_i, \mathbf{Y}_i) \to (X_i, Z_i), i = 1, \cdots, n$. Let $0 \leq Z_{n1} \leq Z_{n2} \leq \cdots \leq Z_{nn}$ be the order statistics corresponding to $Z_1, Z_2, \cdots, Z_n$, and let $X_{n1}, X_{n2}, \cdots, X_{nn}$ be the *induced order statistics*, i.e.

$$X_{ni} = X_j \ if \ Z_{ni} = Z_j,\ for\ i, j = 1, \cdots, n.$$

For every positive integer $k_n(\leq n)$, the k_n nearest neighbor (k-NN) empirical distribution estimator of $G(x|\mathbf{y})$ is defined by

$$G_{nk_n}(x|\mathbf{y}) = k_n^{-1} \sum_{i=1}^{k_n} I(X_{ni} \leq x), \quad x \in \mathcal{R}^1,$$

where $I(A)$ stands for the indicator function of the set A.
Remarks: (1) The induced order statistics $X_{n1}, X_{n2}, \cdots, X_{nn}$ are not i.i.d. (however, they are conditionally independent as stated in the proposition 1). This brings some difficulties in the study of related statistics.
(2) For any two distinct points $\mathbf{y}_1$ and $\mathbf{y}_2$, the number of observations $k_n(\mathbf{y}_i)$ $(i = 1, 2)$ being used to estimate $G(x|\mathbf{y}_i)$ $(i = 1, 2)$ are small compared to n (since usually $k_n/n \to 0$);
(3) As n increases, the intersection of the induced order statistics sets for these two neighbors becomes asymptotically a null set (by Bahadur theorem), thus the two estimators $G_{nk}(x|\mathbf{y}_1)$ and $G_{nk}(x|\mathbf{y}_2)$ are asymptotically independent.

The induced order statistics have many nice properties. For our purpose we state one of them which is due to Bhattacharya (1974). Let $G^*(\cdot|z)$ be the conditional distribution of X given $Z = z$, then

Proposition 1. *For every n, the induced order statistics $X_{n1}, X_{n2}, \cdots, X_{nn}$ are conditionally independent given $Z_1, Z_2, \cdots, Z_n$ with conditional distributions $G^*(\cdot|Z_{n1}), G^*(\cdot|Z_{n2}), \cdots, G^*(\cdot|Z_{nn})$ r espectively.*

In the following, for the sake of simplicity, we only consider the situation where y is on the (positive) real line.

2.1 An Asymptotic Property of the Conditional Empirical Process. Let $k_n = [n^{4/5}t]$, $0 < a \le t \le b < \infty$, and

$$W_n(t, x, y) = k_n^{1/2}(G_{nk_n}(x|y) - G(x|y)).$$

Note that for any fixed two points y_1 and y_2, the processes $W_n(t, x, y_1)$ and $W_n(t, x, y_2)$ are asymptotically independent. This leads us to first consider the case when y is fixed. In this case it is expected that the process $W_n(t, x, y)$ possesses some usual properties as the ordinary empirical process has, namely, as k_n tends infinity, under proper conditions $W_n(t, x, y)$ converges weakly to a Gaussian process. We show that indeed this is the case.

For the random vector (X, Y), let $f(y)$ denote the marginal density function (pdf) of Y, and $g(\cdot|y)$ the conditional pdf of X given $Y = y$ with corresponding conditional distribution function (cdf) $G(\cdot|y)$. For any fixed point $y_0 \in \mathcal{C}$ (compact set) it is assumed that:

[A1:] (a)

$$f(y_0) > 0. \tag{2.1}$$

(b) $f''(y)$ exists in a neighborhood of y_0, and there exist $\epsilon > 0$ and a constant $A^* < \infty$ such that $|y - y_0| < \epsilon$ implies

$$|f''(y) - f''(y_0)| \le A^*|y - y_0|. \tag{2.2}$$

[A2:] (a) $g(\cdot|y_0) > 0$.

(b) The partial derivatives $g_{yy}(x|y)$ of $g(x|y)$ and $G_{yy}(x|y)$ of $G(x|y)$ exist in a neighborhood of y_0, and there exist $\epsilon > 0$ and Lebesgue measurable functions $u_1^*(x)$, $u_2^*(x)$ and $u_3^*(x)$ such that for $|y - y_0| < \epsilon$,

$$\begin{aligned} |g_y(x|y)| \le u_1^*(x), \qquad & |g_{yy}(x|y)| \le u_2^*(x), \\ |g_{yy}(x|y) - g_{yy}(x|y_0)| \quad \le & \quad u_3^*(x)|y - y_0|. \end{aligned} \tag{2.3}$$

Remarks: (1) Recall that $Z_i = |Y_i - y_0|$, let $F_Z(z),\ z \geq 0$, be the marginal distribution of Z_i. Then $F_Z(z)$ has a density function

$$f_Z(z) = f(y_0 + z) + f(y_0 - z),\ z \geq 0. \tag{2.4}$$

Thus, from (2.1) and (2.2), we conclude that [A1] holds for f_Z too.
(2) Let $G^*(x|z)$ and $g^*(x|z)$ be the conditional distribution function and conditional density function of X given $Z = z$, respectively, then

$$G^*(x|z) = \{f(y_0+z)G(x|y_0+z) + f(y_0-z)G(x|y_0-z)\}/f_Z(z), \tag{2.5}$$

$$g^*(x|z) = \{f(y_0+z)g(x|y_0+z) + f(y_0-z)g(x|y_0-z)\}/f_Z(z). \tag{2.6}$$

(3) Note that $G^*(x|0) = G(x)$ and $g^*(x|0) = g(x|y_0)$. For notational simplicity, from now on $G(x|y_0)$ and $g(x|y_0)$ will be denoted by $G(x)$ and $g(x)$, respectively.

Theorem 1. *Under the above conditions, at any fixed point y_0,*

$$\{W_n(t,x,y_0),\ t \times x \in [a,b]\times[0,1],\ 0<a<b<\infty\}$$
$$\xrightarrow{\mathcal{D}} \{W(t,x,y_0),\ t \times x \in [a,b]\times[0,1],\ 0<a<b<\infty\},$$

where $W(t,x,y_0)$ is a Gaussian process with mean $\beta(x,y_0)t^{5/2}$ and covariance $EW(t_1,x_1,y_0)W(t_2,x_2,y_0)$
$= (t_2/t_1)^{1/2}\{G(x_1|y_0) - G(x_1|y_0)G(x_2|y_0)\} + \beta(x_1,y_0)\beta(x_2,y_0)(t_1t_2)^{5/2}$, for $t_1 \leq t_2$, and $x_1 \leq x_2$, where $\beta(x,y_0) = \frac{1}{24}\{G_{yy}(x|y_0)f(y_0) + 2f'(y_0)G_y(x|y_0)\}/f^3(y_0)$.

Remark: This theorem remains true if (2.2) and (2.3) are replaced, respectively, by existing an $\alpha > 0$ such that

$$|f''(y) - f''(y_0)| \leq A^*|y-y_0|^\alpha, \tag{2.7}$$

$$|g_{yy}(x|y) - g_{yy}(x|y_0)| \leq u_3^*(x)|y-y_0|^\alpha. \tag{2.8}$$

The proof of the theorem is given in section 4.

2.2 Local Asymptotic Behavior of the Conditional Empirical Process. It is known that for distinct two points y_1 and y_2 the processes $W_n(t,x,y_1)$ and $W_n(t,x,y_2)$ are asymptotically independent and each converges to its corresponding Gaussian processes. However, when y_1 and y_2 are "close" enough (in a shrinking neighborhood of a fixed point), these two processes ($W_n(t,x,y_1)$ and $W_n(t,x,y_2)$) may have some dependence structure. In fact, let $I_n(y_0) = \{y:\ |y-y_0| \leq n^{-1}k_nC\}$ be such a shrinking neighborhood of y_0, where C satisfies $Cf(y_0) \leq 1$, and $f(y)$ is the marginal density of random variable Y. For any two points y_n and

y_n' belonging to $I_n(y_0)$, let $k(y_n, y_n')$ be the number of elements in common between this two neighbor sets defined as before. Consider the intersection set $I_n(y_n, y_n')$:

$$[y_n - (2nf(y_0))^{-1}k_n, y_n + (2nf(y_0))^{-1}k_n] \cap [y_n' - (2nf(y_0))^{-1}k_n, y_n' + (2nf(y_0))^{-1}k_n]. \tag{2.9}$$

Let $l_n(y_n, y_n')$ be the length of this intersection set, and $k(y_n, y_n')$ be the number of observations in common, then by the Bahadur (1966) representation, we may claim that

$$n^{-1}k(y_n, y_n') - f(y_0)l_n(y_n, y_n) = o(n^{-1}k_n) \text{ } a.s., \text{ } as \text{ } n \to \infty.$$

As such, we may reach the conclusion that the number of common observations used in estimating $G_{nk_n}(x|y_n)$ and $G_{nk_n}(x|y_n')$ asymptotically is

$$nf(y_0)l_n(y_n, y_n')/k_n. \tag{2.10}$$

Proposition 2. *Let $y_n = y_0 + n^{-1}k_n s_1$, $y_n' = y_0 + n^{-1}k_n s_2$, $s_1, s_2 \in [-C/2, C/2]$ and $s_1 < s_2$. Then*

$$nf(y_0)l_n(y_n, y_n')/k_n \to [s_1 f(y_0) - 1/2, s_1 f(y_0) + 1/2] \cap [s_2 f(y_0) - 1/2, s_2 f(y_0) + 1/2], \text{ } s_1 < s_2.$$

Proof: note that here $k_n = [n^{4/5}t]$, plug $y_n = y_0 + n^{-1}k_n s_1$, $y_n' = y_0 + n^{-1}k_n s_2$ into (2.9) which gives $l_n(y_n, y_n')$. Simple calculation shows that (2.10) is nothing but

$$[s_1 f(y_0) - 1/2, s_1 f(y_0) + 1/2] \cap [s_2 f(y_0) - 1/2, s_2 f(y_0) + 1/2].$$

Since when y_n and y_n' belong to a shrinking neighborhood of a fixed point y_0, it has such a interesting dependence structure between $W_n(t, x, y_n)$ and $W_n(t, x, y_n')$, it leads us to further investigate the properties of the stochastic process $W_n(t, x, y)$ when y varying in some shrinking neighborhood of a fixed point y_0.

Let

$$W_n^0(s) = W_n(t, x(s), y_0 + n^{-1}k_n s), \text{ } s \in [-C/2, C/2],$$

where $x(s)$ is defined in a way that for any given $0 \le p \le 1$, $G(x(s)|y_0 + n^{-1}k_n s) = p$. Namely, for any fixed p $x(s) = x_p(s) \equiv [G(\cdot|y_0 + n^{-1}k_n s)]^{-1}(p)$. Then, we have

Theorem 2. *Suppose [A1], [A2] hold, then*

$$\{W_n^0(s), s \in [-C/2, C/2]\} \xrightarrow{\mathcal{D}} \{W^0(s), s \in [-C/2, C/2]\},$$

where $W^0(s)$ is a Gaussian process with mean $\beta(x, y_0)t^{5/2}$ and covariance ($s_1 < s_2$) $EW^0(s_1)W^0(s_2) = l(G(x) - G^2(x|y_0)) + \beta^2(x, y_0)t^5$, where $l = 1 - f(y_0)(s_2 - s_1)$.

Remarks: This theorem also remains true if the conditions (2.2) and (2.3) are replaced by (2.7) and (2.8).

3. Limiting Behavior of the Conditional Sample Functions

In this section, we will study the properties of

$$V_n(y) = k_n^{1/2}[\sup_{1\le i\le k_n} \Psi(X_{ni}, G_{nk_n}(X_{ni}|y)) - \sup_x \Psi(x, G(x|y))].$$

Notice that for distinct two points y_1 and y_2 (fixed) $V_n(y_1)$ and $V_n(y_2)$ are asymptotically independent since the data set involved in estimating $G(x|y_1)$ and $G(x|y_2)$ are asymptotically disjoint. Therefore, it is expected that at each fixed point, say y_0, under some conditions $V_n(y_0)$ converges to a normal random variable. However, when y varies in a shrinking neighborhood of the fixed point $y_0 \in [-C/2, C/2]$, as we saw in the previous section that for such y's, $G_{nk_n}(x|y)$ has some dependence structure, it is not surprise that for those y's, $V_n(y)$ also has that kind of structure. More precisely, we reach the conclusion that under some conditions $V_n(y)$ does converge to a Gaussian process as y varying in the shrinking neighborhood of a fixed point y_0. Explicitly, for $y = y_0 + n^{-1}k_n s$, where $s \in [-C/2, C/2]$ and C satisfies $Cf(y_0) \le 1$, we show that under suitable c onditions, $V_n(y)$ is uniformly in probability proportional to $k_n^{1/2}(G_{nk}(x^0(y)|y) - G(x^0(y)|y))$, where $x^0(y)$ is the point at which $\Psi(x, G(x|y))$ reaching its maximum.

Let

$$G'(x|s) = G(x|y_0 + n^{-1}k_n s), \qquad h(x,y) = \Psi(x, G(x|y)),$$
$$h^0(y) = \sup_x \Psi(x, G(x|y)), \qquad h_n(x,y) = \Psi(x, G_{nk_n}(x|y)).$$

The following conditions are assumed:

[B1:] For each y, $h(x,y)$ reaches its maximum $h^0(y)$ at a unique point $x^0(y)$, and $0 < \inf_y h^0(y) \le \sup_y h^0(y) < \infty$, for $y \in I_n(y_0)$.

[B2:] The conditional density function $g^*(x|y)$ is greater or equal some constant, i.e. $g^*(x|y) \ge c > 0$ for y in the neighborhood of y_0.

[B3:] The partial derivatives of $\Psi(\cdot,\cdot)$ exist and continuous. Further, it is assumed that there exist positive finite constant $c_i, m_i (> 1), i = 1, 2$, independent of y such that for all $x : |x - x^0(y)| \le \rho(> 0)$ and uniformly in y

$$h(x^0(y), y) - c_1|x - x^0(y)|^{m_1} \le h(x,y) \le h(x^0(y), y) - c_2|x - x^0(y)|^{m_2} \tag{3.1}$$

holds, where ρ may be taken small. In order to avoid confusion, let $\Psi_{10}(\cdot,\cdot)$ and $\Psi_{01}(\cdot,\cdot)$ denote the partial derivatives of $\Psi(\cdot,\cdot)$ w.r.t. its first and second variable, respectively. Further, it is assumed that

(a) if $\xi(y) = \Psi_{01}(x^0(y), G(x^0(y)|y))$, then $0 < \inf_y \xi(y) \le \sup_y \xi(y) < \infty$

(b) for all y, $|\Psi_{01}(x,y)| \le \pi(x)$, where $\pi(x)$ is continuous and $\pi^2(x)$ is uniformly integrable w.r.t. $G^*(x|y)$, namely

$$\sup_y \int_{\pi(x)>t} \pi^2(x) dG^*(x|y) \to 0 \text{ as } t \to \infty \text{ uniformly in } y. \tag{3.2}$$

[B4:] To avoid the possibility of having another local maximum of $h(x,y)$ to be arbitrary close to $h^0(y)$, we impose the following separability condition, i.e. if $h^1(y)$ be the second largest local maximum of $h(x,y)$, then

$$\inf_y [h^0(y) - h^1(y)] > h^* > 0. \tag{3.3}$$

Remark: if $h(x,y)$ has a continuous second derivative (w.r.t. x) which is negative and uniformly bounded away from 0, and bounded below in the interval $[x^0(y)-\rho, x^0(y)+\rho]$, then (3.1) holds with $m_1 = m_2 = 2$.

Then, the main theorem is the following:

Theorem 3. *under the conditions [A1], [A2], [B1] up to [B4], $V_n(y)$ is uniformly (in $y \in I_n(y_0)$) in probability proportional to $k_n^{1/2}(G_{nk_n}(x^0(y)|y) - G(x^0(y)|y))$, namely*

$$\begin{aligned} V_n(y) &= k_n^{1/2}[\sup_{1\le i\le k_n} \Psi(X_{ni}, G_{nk_n}(X_{ni}|y)) - \sup_x \Psi(x, G(x|y))] \\ &= \xi(y_0) k_n^{1/2}(G_{nk_n}(x^0(y)|y) - G(x^0(y)|y)) + o_p(1) \end{aligned} \tag{3.4}$$

where x^0 is the point at which $\Psi(x, G(x|y))$ reaching its unique maximum, and $\xi(y_0) = \Psi_{01}(x^0(y_0), G(x^0(y_0)|y_0))$.

The above theorem has some obvious but important consequences. One is at any fixed point , say, y_0, $V_n(y_0)$ tends to a normal random variable as n tending to infinity; another fact is that when y varying in a shrinking neighborhood of y_0, $V_n(y)$ converges to a Gaussian process. We state these facts as corollaries:

Corollary 1. *Under the same conditions as those in theorem 3, at any given point y_0, $V_n(y_0)$ is asymptotically normal.*

Corollary 2. *under the conditions of theorem 3, when y varying in a shrinking neighborhood of y_0, i.e. $y = y_0 + n^{-1}k_n s$, $s \in [-C/2, C/2]$ and C satisfies $Cf(y_0) \le 1$, $V_n(y)$ converges to a Gaussian process.*

4. Proofs

4.1 Preliminary Results (all equalities and inequalities below hold a.s. without specification):

Lemma 1. *Let $k = [n^{\lambda} t]$ for some λ $(0 < \lambda < 1)$ and t $(0 < a < t < b < \infty)$. Assumption [A1] holds. Then (a) for every A such that $Af(y_0) > b > a > 0$, there exists an $n_0 < \infty$ such that*

$$P[Z_{n[bn^{\lambda}]} > An^{-1+\lambda}] \leq \exp[-2n^{-1+2\lambda}(Af(y_0) - b)^2],\ \forall n \geq n_0.$$

(b) If $U_{n1} < U_{n2} < \cdots < U_{nn}$ denote the order statistics in a random sample $U_1, U_2, \cdots, U_n$ of size n from uniform distribution $[0,1]$, then for $A > b$ there exists an $n_0 (< \infty)$ such that

$$P[U_{n[bn^{\lambda}]} > An^{-1+\lambda}] \leq \exp[-2n^{-1+2\lambda}(A - b)^2],\ \forall n \geq n_0.$$

Hence, $U_{n[bn^{\lambda}]}$ and $Z_{n[bn^{\lambda}]}$ are stochastically $\mathcal{O}(n^{-1+\lambda})$.

For the proof of this lemma, we may refer to Bhattacharya and Mack (1987).

Lemma 2. *Let $k = [n^{4/5} t]$ $(0 < a < b < \infty)$. Suppose that the assumption [A1] holds, then*

$$k^{-1} \sum_{i=1}^{k} Z_{ni}^2 = \{12 f^2(y_0)\}^{-1} (k/n)^2 + R_{nk},\ \textit{where} \tag{4.1}$$

$$\max_{k \leq [n^{4/5} b]} |R_{nk}| = \mathcal{O}(n^{-3/5}) \tag{4.2}$$

A similar result was proved by Bhattachy and Gangopadhyay (1990) under a little stronger conditions. However, the result still is true under [A1] by the similar proof.

Note that if $k_1 = [n^{4/5} t_1]$ and $k_2 = [n^{4/5} t_2]$ $(0 < a < t_1 < t_2 < b < \infty)$, then by (4.1) and (4.2) we have

$$\begin{aligned} k_2^{-1/2} \sum_{i=1}^{k_2} Z_{ni}^2 - k_1^{-1/2} \sum_{i=1}^{k_1} Z_{ni}^2 &= \{12 f^2(y_0)\}^{-1} [(k_2/n)^2 k_2^{1/2} - (k_1/n)^2 k_1^{1/2}] + R*_{nk} \\ &\longrightarrow \{12 f^2(y_0)\}^{-1} [t_2^{5/2} - t_1^{5/2}], \end{aligned} \tag{4.3}$$

where $\max_{k_i \leq [n^{4/5} b]} |R^*_{nk}| = \mathcal{O}(n^{-1/5})$.

Lemma 3. *Under conditions[A1] and [A2], the following expansions hold for the conditional pdf $g^*(x|z)$ and the conditional cdf $G^*(x|z)$:*

$$g^*(x|z) = g(x) + (1/2) q(x) z^2 + r(x,z) z^3, \tag{4.4}$$

$$G^*(x|z) = G(x) + (1/2) Q(x) z^2 + R(x,z) z^3, \tag{4.5}$$

where

$$g(x) = g(x|y_0), \qquad G(x) = G(x|y_0),$$

$$q(x) = g_{yy}(x|y_0) + 2f'(y_0) g_y(x|y_0)/f(y_0), \qquad Q(x) = G_{yy}(x|y_0) + 2f'(y_0) G_y(x|y_0)/f(y_0),$$

and there exist $\epsilon > 0$ and $M < \infty$ such that $|q(x)|, |Q(x)|, |r(x,z)|$ and $|R(x,z)|$ are all bounded by M for $0 < z < \epsilon$.

For the proof, we refer to Bhattacharya and Gangopadhyay (1990). 4.2 PROOF OF THEOREM 1: First, the normality of $W_n(t,x,y_0)$: Let $k_{nj}=[n^{4/5}t_j],\ j=1,2,\cdots,m$, from Wald theorem, it is equivalent to show that $\forall\lambda\in\mathcal{R}^m$ arbitrary but fixed, $\sum_{j=1}^m\lambda_jW_n(t_j,x_j,y_0)$ tends to a normal random variable as $n(k_n)$ tending to ∞. Without loss of generality, we assume that $t_1\le t_2\le\cdots\le t_m$. Rewrite $\sum_{j=1}^m\lambda_jW_n(t_j,x_j,y_0)$ into

$$\begin{aligned}&\sum_{i=1}^{k_{n1}}\sum_{j=1}^{m}\lambda_jk_{nj}^{-1/2}(I_{\{X_{ni}\le x_j\}}-G(x_j))+\sum_{i=k_{n1+1}}^{k_{n2}}\sum_{j=2}^{m}\lambda_jk_{nj}^{-1/2}(I_{\{X_{ni}\le x_j\}}-G(x_j))+\cdots\\&+\sum_{i=k(m-2)+1}^{k_{n(m-1)}}\sum_{j=m-2}^{m-1}\lambda_jk_{nj}^{-1/2}(I_{\{X_{ni}\le x_j\}}-G(x_j))+\sum_{i=k(m-1)+1}^{k_{nm}}\lambda_mk_{nm}^{-1/2}(I_{\{X_{ni}\le x_m\}}-G(x_m))\\&\equiv\sum_{i=1}^{k_{nm}}V_{ni},\ \text{where}\end{aligned}$$

$$V_{ni}=\begin{cases}\sum_{j=1}^m\lambda_jk_{nj}^{-1/2}(I_{\{X_{ni}\le x_j\}}-G(x_j)), & \text{for } 1\le i\le k_{n1},\\ \sum_{j=2}^m\lambda_jk_{nj}^{-1/2}(I_{\{X_{ni}\le x_j\}}-G(x_j)), & \text{for } k_{n1}+1\le i\le k_{n2},\\ \cdots\cdots & \\ \lambda_mk_{nm}^{-1/2}(I_{\{X_{ni}\le x_m\}}-G(x_m)), & \text{for } k_{n(m-1)}+1\le i\le k_{nm}.\end{cases}$$

Observe that V_{ni} are (conditionally) independent random variables. Using lemma 1 ($k_{nj}^{-1}Z_{ni}^2=\mathcal{O}(n^{-6/5})$) and lemma 3, it can be showed that for $k_{nh}+1\le i\le k_{n(h+1)},\ 1\le h\le m-2$ $(k_{n0}\equiv 0)$,

$$\begin{aligned}\sigma^2(V_{ni}) &= \sum_{j=h}^{m}(\lambda_jk_{nj}^{-1/2})^2[G(x_j)-G^2(x_j)]\\&\quad+\sum_{j\ne l=h}^{m}\lambda_j\lambda_l(k_{nj}k_{nl})^{-1/2}[G(\min(x_j,x_l))-G(x_j)G(x_l)]+\mathcal{O}(n^{-6/5}),\end{aligned}$$

and for $k_{n(m-1)}+1\le i\le k_{nm}$,

$$\sigma^2(V_{ni})=\lambda^2k_{nm}^{-1}(G(x_m)-G^2(x_m)). \tag{4.6}$$

Therefore,

$$\begin{aligned}
s_{nm}^2 &\equiv \sum_{i=1}^{k_{nm}} \sigma^2(V_{ni}) \\
\rightarrow\ & t_1 \sum_{j=1}^{m} \lambda_j^2 t_j^{-1}[G(x_j) - G^2(x_j)] + t_1 \sum_{j\neq l=1}^{m} \lambda_j \lambda_l (t_j t_l)^{-1/2}[G(\min(x_j, x_l)) - G(x_j)G(x_l)] \\
& + (t_2 - t_1)\{\sum_{j=2}^{m} \lambda_j^2 t_j^{-1})[G(x_j) - G^2(x_j)] \\
& + \sum_{j\neq l=2}^{m} \lambda_j \lambda_l (t_j t_l)^{-1/2}[G(\min(x_j, x_l)) - G(x_j)G(x_l)]\} \\
& + \cdots + (t_m - t_{(m-1)}0\lambda_m^2 t_m^{-1}(G(x_m) - G^2(x_m)) > 0.
\end{aligned}$$

Note that for all $1 \leq i \leq k_{nm}$, $EV_{ni} = \mathcal{O}(n^{-4/5})$. $\forall \epsilon > 0$

$$\begin{aligned}
& \frac{1}{s_{nm}^2} \sum_{i=1}^{k_{nm}} E[(V_{ni} - EV_{ni})^2 I_{\{|V_{ni} - EV_{ni}| > \epsilon s_{nm}\}}] \\
\leq\ & \frac{1}{s_{nm}^2} \sum_{i=1}^{k_{nm}} E[(V_{ni} - EV_{ni})^2 \times I_{\{\sum_{j=1}^{m} |\lambda_j||I_{\{X_{ni} \leq x_j\}} - G^*(x_j|Z_{ni})| \geq k_{n1}^{1/2} \epsilon s_{nm}\}} \\
\longrightarrow\ & 0, \quad \text{as } n \to \infty,
\end{aligned}$$

since for all i, $\sum_{j=1}^{m} |\lambda_j||I_{\{X_{ni} \leq x_j\}} - G^*(x_j|Z_{ni})| \leq 2\sum_{j=1}^{m} |\lambda_j|$, but $k_{n1}^{1/2} s_{nm} \to \infty$. Thus, the Lindeberg condition is satisfied, and it means that $\sum_{j=1}^{m} \lambda_j W_n(t_j, x_j, y_0) \longrightarrow Normal$.

Next, we prove that $W_n(t, x, y_0)$ $(t \times x \in [a, b] \times (0, 1))$ is tight. Let $T^n = \{\frac{i}{k_n}, [k_n a] \leq i \leq [k_n b]\} \times [0, 1]\}$, then, T_n is dense in $[a, b] \times [0, 1]$. Define

$$\begin{aligned}
W_n^*(t, x) &\equiv k_n^{1/2}[G_{nk}(x|y_0) - G(x)] - k_{0n}^{1/2}[G_{0n}(x|y_0) - G(x)] \\
&= W_n(t, x, y_0) - W_n(t_0, x, y_0),
\end{aligned}$$

where $t_0 = a, k_{0n} = [n^{4/5}a]$. Thus, $W_n^*(t, x)$ vanishes along its lower boundary $a \times [0, 1] \cup [a, b] \times \{0\}$ and its increment over block $B = (t_1, t_2] \times (a_1, a_2]$ is

$$\begin{aligned}
W_n^*(B) &= [W_n^*(t_2, a_2) - W_n^*(t_2, a_1)] - [W_n^*(t_1, a_2) - W_n^*(t_1, a_1)] \\
&= [W_n(t_2, a_2, y_0) - W_n(t_2, a_1, y_0)] - [W_n(t_1, a_2, y_0) - W_n(t_1, a_1, y_0)],
\end{aligned}$$

which is the same as that of $W_n(t, x, y_0)$ over the block B. Likewise, for the neighboring blocks $C = (t_1, t_2] \times (a_2, a_3]$ and $D = (t_2, t_3] \times (a_1, a_2]$, the corresponding increments can be similary obtained. Let

$$U_{ni}(x) \equiv I_{\{X_{ni} \leq x\}} - G(x),$$

for $B_2 = (a_1, a_2]$, define

$$U_{ni}(B_2) \equiv U_{ni}(a_2) - U_{ni}(a_1) = I_{\{a_1 < X_{ni} \le a_2\}} - G(B_2),$$

where $G(B_2) = G(a_2) - G(a_1)$. Similarly we define $U_{ni}(C_2)$ for block $C_2 = (a_2, a_3]$. It's easy to show that

$$\begin{aligned} EU_{ni}^2(B_2) &\le G(B_2) + \mathcal{O}(n^{-2/5}), & (4.7)\\ EU_{ni}^2(B_2)U_{ni}(C_2) &\le G(B_2)G(C_2) + \mathcal{O}(n^{-2/5}), & (4.8)\\ EU_{ni}^2(B_2)U_{ni}^2(C_2) &\le G(B_2)G(C_2) + \mathcal{O}(n^{-2/5}). & (4.9) \end{aligned}$$

Define

$$W_n^{\star\star}(t,x) \equiv n^{-2/5} \sum_{i=1}^{k_n} (I_{\{X_{ni} \le x\}} - G(x)),\ t \times x \in [a,b] \times [0,1].$$

If $W_n^{\star\star}(t,x)$ is tight on $[a,b] \times [0,1]$, so is $W_n^{\star}(t,x)$ (the reason we define $W_n^{\star\star}(t,x)$ is to simplify the calculation below, such defined $W_n^{\star\star}(t,x)$ has a nice representation by $U_{ni}(\cdot)$, but $W_n^*(t,x)$ does not). Note that

$$\begin{aligned} W_n^{\star\star}(B) &= [W_n^{\star\star}(t_2,a_2) - W_n^{\star\star}(t_2,a_1)] - [W_n^{\star\star}(t_1,a_2) - W_n^{\star\star}(t_1,a_1)]\\ &= n^{-2/5} \sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}(B_2), \end{aligned}$$

here $k_{nj} = [n^{4/5} t_j],\ j = 1,2,3$, and B, B_2 as before. Similarly

$$W_n^{\star\star}(C) = n^{-2/5} \sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}(C_2), \quad W_n^{\star\star}(D) = n^{-2/5} \sum_{i=k_{n2}+1}^{k_{n3}} U_{ni}(B_2).$$

Thus (by noting (4.7))

$$\begin{aligned} &E[W_n^{\star\star}(B) W_n^{\star\star}(D)]^2\\ &= n^{-8/5} E\{[\sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}(B_2)]^2 [\sum_{i=k_{n2}+1}^{k_{n3}} U_{ni}(B_2)]^2\} & (4.10)\\ &\le n^{-8/5} E\{[(k_{n2} - k_{n1}) \sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}^2(B_2)][(k_{n3} - k_{n2}) \sum_{i=k_{n2}+1}^{k_{n3}} U_{ni}^2(B_2)]\} & (4.11)\\ &\le 16[(t_2 - t_1)G(B_2)][(t_3 - t_2)G(B_2)] + \mathcal{O}(n^{-2/5}) & (4.12) \end{aligned}$$

Note that from (4.10) to (4.11), we used the following inequality: for any real numbers $x_1, x_2, \cdots, x_k$, $(\sum_{i=1}^k x_i)^2 \le k \sum_{i=1}^k x_i^2$. From (4.11) to (4.12) we used the property that for

$k_{n1}+1 \leq i \leq k_{n2}$, U_{ni} (X_{ni}) are (conditionally) independent of U_{nj} (X_{nj}) for $k_{n2}+1 \leq j \leq k_{n3}$.

$$\begin{aligned}
&E\{[W_n^{\star\star}(B)]^2[W_n^{\star\star}(C)]^2\} \qquad (4.13)\\
&= n^{-8/5}E\{\sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}^2(B_2) \sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}^2(C_2) + \sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}^2(B_2)\\
&\quad \times \sum_{k\neq l=k_{n1}+1}^{k_{n2}} U_{nk}(C_2)U_{nl}(C_2) + \sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}^2(C_2) \sum_{i\neq j=k_{n1}+1}^{k_{n2}} U_{ni}(B_2)U_{nj}(B_2)\\
&\quad + \sum_{i\neq j=k_{n1}+1}^{k_{n2}} U_{ni}(B_2)U_{nj}(B_2) \sum_{k\neq l=k_{n1}+1}^{k_{n2}} U_{nk}(C_2)U_{nl}(C_2)\}\\
&\equiv I_1+I_2+I_3+I_4.
\end{aligned}$$

By virtue of (4.7), (4.8) and (4.9),

$$\begin{aligned}
I_1 &\leq n^{-8/5}[(k_{n2}-k_{n1})\max_i EU_{ni}^2(B_2)U_{ni}^2(C_2)\\
&\quad +(k_{n2}-k_{n1})(k_{n2}-k_{n1}-1)\max_i EU_{ni}^2(B_2)\max_j EU_{nj}^2(C_2)\\
&\leq 4[(t_2-t_1)G(B_2)][(t_2-t_1)G(C_2)]+\mathcal{O}(n^{-2/5})\\
I_2 &= n^{-8/5}E[\sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}^2(B_2) \sum_{j=i\neq l=k_{n1}+1}^{k_{n2}} U_{nj}(C_2)U_{nl}(C_2)\\
&\quad + \sum_{i=k_{n1}+1}^{k_{n2}} U_{ni}^2(B_2) \sum_{j\neq i\neq l=k_{n1}+1}^{k_{n2}} U_{nj}(C_2)U_{nl}(C_2)]\\
&\leq n^{-8/5}(k_{n2}-k_{n1})(k_{n2}-k_{n1}-1)G(B_2)G(C_2)\\
&\quad +(1/4)n^{-8/5}(k_{n2}-k_{n1})(k_{n2}-k_{n1}-1)(k_{n2}-k_{n1}-2)G(B_2)(\max_i Z_{ni}^2)^2+\mathcal{O}(n^{-2/5})\\
&= [(t_2-t_1)G(B_2)][(t_2-t_1)G(C_2)]+A_1(t_2-t_1)(t_2-t_1)G(B_2)Q^2(C_2)+\mathcal{O}(n^{-1/5}),
\end{aligned}$$

where A_1 is a constant associated with lemma 1 and $Q^2(C_2)=(Q(a_3)-Q(a_2))^2$. Likewise

$$I_3 \leq [(t_2-t_1)G(B_2)][(t_2-t_1)G(C_2)]+A_2(t_2-t_1)(t_2-t_1)G(C_2)Q^2(B_2)+\mathcal{O}(n^{-1/5}),$$

and

$$\begin{aligned}
|I_4| &= |n^{-8/5} E \sum_{i\neq j=k_{n1}+1}^{k_{n2}} U_{ni}(B_2)U_{nj}(B_2) \sum_{k=i\neq l=k_{n1}+1}^{k_{n2}} U_{ni}(C_2)U_{nl}(C_2)\} \\
&\quad + \sum_{i\neq j\neq k\neq l=k_{n1}+1}^{k_{n2}} U_{ni}(B_2)U_{nj}(B_2)U_{nk}(C_2)U_{nl}(C_2)| \\
&\leq (1/4)n^{-8/5}(k_{n2}-k_{n1})(k_{n2}-k_{n1}-1)(k_{n2}-k_{n1}-2) \\
&\quad \times G(B_2)G(C_2)|Q(B_2)Q(C_2)|(A_2^* n^{-4/5}+\mathcal{O}(n^{-1})) \\
&\quad + n^{-8/5}(k_{n2}-k_{n1})(k_{n2}-k_{n1}-1)(k_{n2}-k_{n1}-2)(k_{n2}-k_{n1}-3) \\
&\quad \times Q^2(B_2)Q^2(C_2)|(A_3^* n^{-8/5}+\mathcal{O}(n^{-9/5})) \\
&\leq A_4^*[(t_2-t_1)G(B_2)][(t_2-t_1)G(C_2)] \\
&\quad + A_5^*[(t_2-t_1)G(B_2)][(t_2-t_1)G(C_2)]Q^2(B_2)Q^2(C_2)+\mathcal{O}(n^{-1/5}),
\end{aligned}$$

where $A_2, A_2^*, A_3^*, A_4^*, A_5^*$ are constants. From condition [A2], we know both $|Q(B_2)| = |Q(a_2) - Q(a_1)|$ and $|Q(C_2)| = |Q(a_3) - Q(a_2)|$ are bounded. Now for $B^* = (x_1, x_2] \in [0,1]$, define

$$\mu(B^*) = G(B^*|y_0) + \int_{x_1}^{x_2} |g_{yy}(x|y_0) + \frac{2f'(y_0)}{f(y_0)} g_y(x|y_0)|dx.$$

It is easy to prove that $\mu(B^*)$ (its extension) is a finite measure on $[0,1]$. By noting the boundness of $|Q(B_2)|(|Q(C_2)|)$ we conclude that there exists a constant A such that

$$|I_j| \leq A[(t_2-t_1)\mu(B_2)][(t_2-t_1)\mu(C_2)],\ j=1,2,3,4.$$

According to Bickel and Wichura (1971) $W_n^{\star\star}(t,x)$ is tight, so is $W_n^\star(t,x)$.
Note that the process $W_n(a,x,y_0)$ is also tight. It can be seen easily from the calculation of Billingsley type inequality: for $0 \leq x_1 < x < x_2$

$$\begin{aligned}
& E[W_n(a,x,y_0) - W_n(a,x_1,y_0)]^2[W_n(a,x_2,y_0) - W_n(a,x,y_0)]^2 \\
&= k_{na}^{-2} E(\sum_{i=1}^{k_{na}} U_{na}((x_1,x]))^2 (\sum_{i=1}^{k_{na}} U_{na}((x,x_2]))^2
\end{aligned}$$

which is the special case of (4.13). Therefore $W_n(t,x,y_0)$ is tight. The calculation for the covariance structure is straight forward. The proof of theorem 1 is complete.

If (2.2) and (2.3) are replaced by (2.7) and (2.8), after carefully checking the proofs, one can see that they go through without any difficulty.

PROOF OF THEOREM 2: In this subsection, for the notation simplicity, let

$$\begin{aligned}\tilde{G}(x(s)|s) &\equiv G(x(s)|y_0+n^{-1}k_n s),\\ Q(x(s)) &\equiv \tilde{G}_{yy}(x(s)|s)+2\tilde{G}_y(x(s)|s)f'(y_0+n^{-1}k_n s)/f(y_0+n^{-1}k_n s).\end{aligned}$$

Expand $\tilde{G}(x(s)|s), Q(x(s))$, it is easy to see that

$$\tilde{G}(x(s)|s) = G(x(0))+\mathcal{O}(n^{-1/5}),\ Q(x(s))=Q(x(0))+\mathcal{O}(n^{-1/5}),$$

where $x(s)\equiv x_p(y_0+n^{-1}k_n s)=[G(\cdot|y_0+n^{-1}k_n s)]^{-1}(p)$, for any given $p\in[0,1]$ (p may be taken as $G(x(0)|y_0)$). Then, $x(s)-x(0)=O(n^{-\frac{1}{5}})$.
For the normality of $W_n^0(s)$, we need to show that for any $\lambda=(\lambda_1,\lambda_2,\cdots,\lambda_m)\in\mathcal{R}^m$ arbitrary but fixed and $s_1,s_2,\cdots,s_m$, $\sum_{j=1}^m\lambda_j W_n^0(s_j)\to Normal$. For notation simplicity, we give the proof of the two dimensional case. The higher dimensional situation can be proved by the same routine. $\forall\lambda\in\mathcal{R}^2$ arbitrary but fixed and $s_1<s_2$, rewrite $\lambda_1W_n^0(s_1)+\lambda_2W_n^0(s_2)$ as

$$\begin{aligned}&\lambda_1 k_n^{-1/2}\sum_{\{i=1,\cdots,k_n\}\backslash\{common\}}(I_{\{X_{ni}^{s_1}\le x(s_1)\}}-\tilde{G}(x(s_1)|s_1))\\ &+[\lambda_1 k_n^{-1/2}\sum_{common}(I_{\{X_{ni}^{s_1}\le x(s_1)\}}-\tilde{G}(x(s_1)|s_1))\\ &+\lambda_2 k_n^{-1/2}\sum_{common}(I_{\{X_{nj}^{s_2}\le x(s_2)\}}-\tilde{G}(x(s_2)|s_2))]\\ &+\lambda_2 k_n^{-1/2}\sum_{\{j=1,\cdots,k_n\}\backslash\{common\}}(I_{\{X_{nj}^{s_2}\le x(s_2)\}}-\tilde{G}(x(s_2)|s_2))\\ &\equiv I_{n,5}+I_{n,6}+I_{n,7},\end{aligned}$$

where $X_{ni}^{s_j}$ $(j=1,2)$ are the induced order statistics with respect to $y_0+n^{-1}k_n s_j$.

First, observe that $I_{n,5},I_{n,6},I_{n,7}$ are (conditionally) independent. Therefore, if we can prove that each of the above terms tends to a normal random variable respectively, then $\lambda_1W_n^0(s_1)+\lambda_2W_n^0(s_2)$ also tends to a normal random variable.

Suppose there are k_n' terms in common, then according to proposition 2,

$$k_n'/k_n\to l=1-f(y_0)(s_2-s_1).$$

We assume that $0 < l < 1/2$ (the proofs for $1/2 \leq l < 1$ are exactly the same), then

$$\begin{aligned}
EI_{n,5} &= (1/2)\lambda_1 k_n^{-1/2} Q^*(x(s_1), s_1) \sum_{\{i=1,\cdots,k_n\}\setminus\{common\}} (Z_{ni}^{s_1})^2 + \mathcal{O}(n^{-3/5}) \\
&= (1/2)\lambda_1 k_n^{-1/2} (Q(x(0)) + \mathcal{O}(n^{-1/5}))[\sum_{i=1}^{k_n-2k_n'} (Z_{ni}^{s_1})^2 + (1/2) \sum_{i=k_n-2k_n'}^{k_n-2k_n'} (Z_{ni}^{s_1})^2] \\
&\quad + \mathcal{O}(n^{-3/5}) \\
&= (1/4)\lambda_1 k_n^{-1/2} Q(x(0))/(12f^2(y_0))\{(k_n - 2k_n')[(k_n - 2k_n')/n]^2 \\
&\quad + (1/2)(k_n(k_n/n)^2 - (k_n - 2k_n')[(k_n - 2k_n')/n]^2)\} + \mathcal{O}(n^{-1/5}) \\
&\to \lambda_1 Q(x(0))/(48f^2(y_0))t^{5/2}[(1-2l)^3 + 1] \qquad (4.14)
\end{aligned}$$

$$E[\lambda_1 k_n^{-1/2}(I_{\{X_{ni}^{s_1} \leq x(s_1)\}} - \tilde{G}(x(s_1)|s_1))]^2 = \lambda_1^2 k_n^{-1}(G(x(0)) - G^2(x(0)) + \mathcal{O}(n^{-1/5})).$$

Note that $E[\lambda_1 k_n^{-1/2}(I_{\{X_{ni}^{s_1} \leq x(s_1)\}} - \tilde{G}(x(s_1)|s_1))] = \mathcal{O}(n^{-4/5})$, we have $\sigma^2(\lambda_1 k_n^{-1/2}(I_{\{X_{ni}^{s_1} \leq x(s_1)\}} - \tilde{G}(x(s_1)|s_1)) = \lambda_1^2 k_n^{-1}(G(x(0)) - G^2(x(0))) + \mathcal{O}(n^{-1})$. Therefore, the sum of variances for $I_{n,5}$ is

$$\begin{aligned}
s_{n1}^2 &= \sum_{\{i=1,\cdots,k_n\}\setminus\{common\}} \sigma^2(I_{\{X_{ni}^{s_1} \leq x(s_1)\}} - \tilde{G}(x(s_1)|s_1)) \\
&= \lambda_1^2 k_n^{-1}(k_n - 2k_n')(G(x(0)) - G^2(x(0))) + \mathcal{O}(n^{-1/5}) \\
&\to \lambda_1^2(1-2l)(G(x(0)) - G^2(x(0))) > 0.
\end{aligned}$$

Hence, $\forall \epsilon > 0$, by noting that

$$I_{\{\lambda_1 k_n^{-1/2}|(I_{\{X_{ni}^{s_1} \leq x(s_1)\}} - \tilde{G}(x(s_1)|s_1) - E(I_{\{X_{ni}^{s_1} \leq x(s_1)\}} - \tilde{G}(x(s_1)|s_1))| > \epsilon s_{n1}\}} = 0$$

when n is large enough, it's easy to see the Lindeberg condition is satisfied. Therefore, $I_{n,5}$ tends to a normal random variable. Likewise, $I_{n,7}$ also tends to a normal random variable. As for $I_{n,6}$,

$$\begin{aligned}
EI_{n,6} &= (1/2)\lambda_1 k_n^{-1/2}(Q(x(0) + \mathcal{O}(n^{-1/5})) \sum_{common} (Z_{ni}^{(s_1)})^2 \\
&\quad + (1/2)\lambda_2 k_n^{-1/2}(Q(x(0)) + \mathcal{O}(n^{-1/5})) \sum_{common} (Z_{nj}^{(s_2)})^2 + \mathcal{O}(n^{-1/5})) \\
&= \lambda_1 k_n^{-1/2} Q(x(0))(48f^2(y_0))^{-1}\{k_n(k_n/n)^2 - (k_n - 2k_n')[(k_n - 2k_n')/n]^2\} + \\
&\quad \lambda_2 k_n^{-1/2} Q(x(0))(48f^2(y_0))^{-1} 1\{k_n(k_n/n)^2 - (k_n - 2k_n')[(k_n - 2k_n')/n]^2\} + \mathcal{O}(n^{-1/5}) \\
&\to Q(x(0))(48f^2(y_0))^{-1} t^{5/2}[1 - (1-2l)^3](\lambda_1 + \lambda_2).
\end{aligned}$$

For the variance of each term in $I_{n,6}$, since they are in common, for any j, X_{nj} (w.r.t. s_2) must be equal to certain X_{ni} (w.r.t. s_1). Pairing those X_{ni}'s and X_{nj}'s, then we have

$$\begin{aligned}
&E[\lambda_1 k_n^{-1/2}(I_{\{X_{ni}^{s_1}\le x(s_1)\}} - \tilde{G}(x(s_1)|s_1)) + \lambda_2 k_n^{-1/2}(I_{\{X_{nj}^{s_2}\le x(s_2)\}} - \tilde{G}(x(s_2)|s_2))]^2 \\
&= k_n^{-1}(\lambda_1+\lambda_2)^2 E(I_{\{X_{ni}\le x(0)\}} - G(x(0)) + \mathcal{O}(n^{-1/5}))^2 \\
&= k_n^{-1}(\lambda_1+\lambda_2)^2 (G(x(0)) - G^2(x(0)) + \mathcal{O}(n^{-1/5})).
\end{aligned}$$

So the sum of the variance of $I_{n,6}$ is

$$\begin{aligned}
s_{n2}^2 &= k_n' k_n^{-1}(\lambda_1+\lambda_2)^2(G(x(0)) - G^2(x(0)) + \mathcal{O}(n^{-1/5})) \\
&\to l(\lambda_1+\lambda_2)^2(G(x(0)) - G^2(x(0))) > 0.
\end{aligned}$$

Again, the Lindeberg condition is satisfied for $I_{n,6}$ since the indicator function $I_{A_{i,j}}$ in Lindeberg condition is 0 when n is large enough (for all $i, j = 1, 2, \cdots, k_n$), where

$$\begin{aligned}
&A_{i,j} = \{|\lambda_1 k_n^{-1/2}(I_{\{X_{ni}^{s_1}\le x(s_1)\}} - \tilde{G}(x(s_1)|s_1)) + \lambda_2 k_n^{-1/2}(I_{\{X_{nj}^{s_2}\le x(s_2)\}} - \tilde{G}(x(s_2)|s_2)) \\
&-E[\lambda_1 k_n^{-1/2}(I_{\{X_{ni}^{s_1}\le x(s_1)\}} - \tilde{G}(x(s_1)|s_1)) + \lambda_2 k_n^{-1/2}(I_{\{X_{nj}^{s_2}\le x(s_2)\}} - \tilde{G}(x(s_2)|s_2))]| > \epsilon\} = \emptyset
\end{aligned}$$

for n large enough. Thus, $I_{n,5}, I_{n,6}, I_{n,7}$ all tend to a normal random variable, respectively. Since they are independent, so $\lambda_1 W_n^0(s_1) + \lambda_2 W_n^0(s_2)$ tends to a normal random variable too.

The verification of the tightness for $W_n^0(s)$ $s \in [-C/2, C/2]$, where C satisfies $Cf(y_0) < 1$: first note that for any two points $-C/2 < s_1 < s_2 < C/2$,

$$\begin{aligned}
&W_n^0(s_2) - W_n^0(s_1) = k_n^{-1/2} \sum_{\{1,\cdots,k_n\}\backslash\{common\}} [I_{\{X_{ni}^{s_2}\le x(s_2)\}} - \tilde{G}(x(s_2)|s_2)] \\
&- k_n^{-1/2} \sum_{\{1,\cdots,k_n\}\backslash\{common\}} [I_{\{X_{nj}^{s_1}\le x(s_1)\}} - \tilde{G}(x(s_1)|s_1)] \\
&+ k_n^{-1/2}[\sum_{common} (I_{\{X_{ni}^{s_2}\le x(s_2)\}} - \tilde{G}(x(s_2)|s_2)) - \sum_{common} (I_{\{X_{nj}^{s_1}\le x(s_1)\}} - \tilde{G}(x(s_1)|s_1))] \quad (4.15)
\end{aligned}$$

For the third term in the above expression since they are in common (corresponding to same set of $Y_k's$ in the original data set), so for any X_{ni} (w.r.t. s_1) which is corresponding to Y_{k_0}, say, in the original data set (i.e. $X_{ni} = X_{k_0}$), then X_{ni} must be equal to certain $X_{nj}, j \in$ common (w.r.t. s_2). Therefor,

$$
\begin{aligned}
&k_n^{-1}E|\sum_{common}(I_{\{X_{ni}^{s_2}\le x(s_2)\}}-\tilde{G}(x(s_2)|s_2))-\sum_{common}(I_{\{X_{nj}^{s_1}\le x(s_1)\}}-\tilde{G}(x(s_1)|s_1))|^2\\
&\quad=k_n^{-1}E[\sum_{common}I_{\{\min(x(s_1),x(s_2))\le X_{ni}\le\max(x(s_1),x(s_2))\}}]^2\\
&\quad=k_n^{-1}E[\sum_{common}I_{\{\min(x(s_1),x(s_2))\le X_{ni}\le\max(x(s_1),x(s_2))\}}\\
&\quad+\sum_{i\neq j\in common}I_{\{\min(x(s_1),x(s_2))\le X_{ni}\le\max(x(s_1),x(s_2))\}}\times I_{\{\min(x(s_1),x(s_2))\le X_{nj}\le\max(x(s_1),x(s_2))\}}]\\
&\quad=k_n^{-1}\{Q(x(0))/2\sum_{common}[(Z_{ni}^{s_2})^2-(Z_{ni}^{s_1})^2]\\
&\quad+Q^2(x(0))/4\sum_{i\neq j\in common}[(Z_{ni}^{s_2})^2-(Z_{ni}^{s_1})^2][(Z_{nj}^{s_2})^2-(Z_{nj}^{s_1})^2]\}+O(n^{-\frac{1}{5}})\\
&\quad=k_n^{-1}k'^2_n\max_{i,j\in common}[(Z_{ni}^{s_2})^2-(Z_{ni}^{s_1})^2][(Z_{nj}^{s_2})^2-(Z_{nj}^{s_1})^2]+O(n^{-\frac{1}{5}})\\
&\quad\le A_3(s_2-s_1)^2, \qquad (4.16)
\end{aligned}
$$

where A_3 is a constant. This is due to the fact that

$$
\begin{aligned}
(Z_{ni}^{s_2})^2-(Z_{ni}^{s_1})^2 &= (Z_{ni}^{s_2}+Z_{ni}^{s_1})(Z_{ni}^{s_2}-Z_{ni}^{s_1}),\\
Z_{ni}^{s_2}+Z_{ni}^{s_1}\le c_1n^{-1/5}, &\qquad |Z_{ni}^{s_2}-Z_{ni}^{s_1}|\le c_2n^{-1/5}(s_2-s_1),
\end{aligned}
$$

where c_1 and c_2 are some constants.

So for $-C/2\le s_1<s<s_2\le C/2$,

$$
\begin{aligned}
&W_n^0(s)-W_n^0(s_1)=k_n^{-1/2}\sum_{I_{11}}[I_{\{X_{ni}^{s}\le x(s)\}}-\tilde{G}(x(s)|s)]\\
&-k_n^{-1/2}\sum_{I_8}[I_{\{X_{ni}^{s_1}\le x(s_1)\}}-\tilde{G}(x(s_1)|s_1)]+k_n^{-1/2}\sum_{I_9\cup I_{10}}[I_{\{X_{ni}^{s}\le x(s)\}}-I_{\{X_{ni}^{s_1}\le x(s_1)\}}]\\
&\equiv I_{n,11}(s)-I_{n,8}(s_1)+I_{n,9}(s_1,s)+I_{n,10}(s_1,s)
\end{aligned}
$$

Likewise

$$
\begin{aligned}
&W_n^0(s_2)-W_n^0(s)=k_n^{-1/2}\sum_{I_{12}}[I_{\{X_{ni}^{s_2}\le x(s_2)\}}-\tilde{G}(x(s_2)|s_2)]\\
&-k_n^{-1/2}\sum_{I_9}[I_{\{X_{ni}^{s}\le x(s)\}}-\tilde{G}(x(s)|s)]+k_n^{-1/2}\sum_{I_{10}\cup I_{11}}[I_{\{X_{ni}^{s_2}\le x(s_2)\}}-I_{\{X_{ni}^{s}\le x(s)\}}]\\
&\equiv I_{n,12}(s_2)-I_{n,9}(s)+I_{n,10}(s_2,s)+I_{n,11}(s_2,s)
\end{aligned}
$$

where I_8 is the set of those terms $(X'_{ni}s)$ which only involve in estimating $G(\cdot|y_0+n^{-1}k_ns_1)$, I_9 is the set of those terms which involve in estimating $G(\cdot|y_0+n^{-1}k_ns_1)$ and $G(\cdot|y_0+n^{-1}k_ns)$ but not $G(\cdot|y_0+n^{-1}k_ns_2)$, while I_{10} is the set of those terms which involve in estimating $G(\cdot|y_0+n^{-1}k_ns_1)$, $G(\cdot|y_0+n^{-1}k_ns)$ and $G(\cdot|y_0+n^{-1}k_ns_2)$. Likewise for I_{11} and I_{12}.

Note that by (4.15) and (4.16), we have

$$EI_{n,9}^2(s,s_1) \le A_4(s_2-s_1)^2, \qquad EI_{n,10}^2(s,s_1) \le A_5(s_2-s_1)^2.$$

Similar results hold for $I_{n,10}(s_2,s)$ and $I_{n,11}(s_2,s)$. Keep in mind that $EI_{n,j}^2$ $(j=8,9,10,11,12)$ are bounded (so are $E|I_{n,j}|$) and, $I_{n,j}$ are conditionally independent, so if we can reach that

$$\begin{aligned} E(I_{n,11}(s)-I_{n,8}(s_1))^2 &\le constant \times (s_2-s_1), \\ E(I_{n,12}(s_2)-I_{n,9}(s))^2 &\le constant \times (s_2-s_1), \\ E((I_{n,11}(s)-I_{n,8}(s_1))I_{n,11}(s_2,s))^2 &\le constant \times (s_2-s_1)^2, \end{aligned}$$

then

$$E[W_n^0(s)-W_n^0(s_1)]^2[W_n^0(s_2)-W_n^0(s)]^2 \le constant \times (s_2-s_1)^2, \tag{4.17}$$

which means the tightness for $W_n^0(s)$. But

$$\begin{aligned} &E(I_{n,11}(s)-I_{n,8}(s_1))^2 \\ &= k_n^{-1}E[\sum_{I_{11}(s)}\{I_{\{X_{ni}^s\le x(s)\}}-\tilde G(x(s)|s)\}]^2 - 2k_n^{-1}E[\sum_{I_{11}(s)}\{I_{\{X_{ni}^s\le x(s)\}}-\tilde G(x(s)|s)\}] \\ &\quad \times E[\sum_{I_8(s_1)}\{I_{\{X_{nj}^{s_1}\le x(s_1)\}}-\tilde G(x(s_1)|s_1)\}] + k_n^{-1}E[\sum_{I_8(s_1)}\{I_{\{X_{nj}^{s_1}\le x(s_1)\}}-\tilde G(x(s_1)|s_1)\}]^2 \\ &\equiv I_n + II_n + III_n. \end{aligned}$$

Note that (suppose that $W_n^0(s)$ and $W_n^0(s_1)$ have k_n' observations in common)

$$\begin{aligned} I_n &= k_n^{-1}E[\sum_{I_{11}(s)}\{I_{\{X_{ni}^s\le x(s)\}}-\tilde G(x(s)|s)\}]^2 \\ &= \frac{k_n-k_n'}{k_n}\frac{1}{k_n-k_n'}\sum_{I_{11}(s)}E\{I_{\{X_{ni}^s\le x(s)\}}-\tilde G(x(s)|s)\}^2 \\ &\quad + \frac{k_n-k_n'}{k_n}\frac{1}{k_n-k_n'}\sum_{i\ne j\in I_{11}(s)}(G^*(x(s)|Z_{ni}^s)-\tilde G(x(s)|s)) \\ &\quad \times (G^*(x(s)|Z_{nj}^s)-\tilde G(x(s)|s_1)) \\ &\le \frac{k_n-k_n'}{k_n}(G(x(0))-G^2(x(0)) \\ &\quad + \frac14 Q^2(x(0))\frac{(k_n-k_n')^2}{k_n}\max_{i,j}(Z_{ni}^s)^2(Z_{nj}^{s_1})^2 + \mathcal{O}(n^{-1/5}) \qquad (4.18) \\ &\le A_7(s-s_1) \le A_7(s_2-s_1), \qquad (4.19) \end{aligned}$$

where A_7 is a constant, and from (4.18) to (4.19) by noting the facts that

$$\begin{aligned} (k_n-k_n')/k_n \to 1-l &= 1-[1-(s-s_1)f(y_0)] \le f(y_0)(s_2-s_1), \\ \max_{i,j}(Z_{ni}^s)^2(Z_{nj}^{s_1})^2 &\le constant \times n^{-4/5}. \end{aligned}$$

Likewise

$$III_n \leq A_8(s_2 - s_1). \tag{4.20}$$

As for II_n, since

$$\begin{aligned} |II_n| &= (2/k_n)|\sum_{I_{11}(s)}(G^*(x(s)|Z^s_{ni}) - \tilde{G}(x(s)|s)) \times \sum_{I_8(s_1)}(G^*(x(s_1)|Z^{s_1}_{nj}) - \tilde{G}(x(s_1)|s_1)| \\ &= 2\frac{k_n - k'_n}{k_n}\frac{1}{k_n - k'_n}[\sum_{I_{11}(s)}(Z^s_{ni})^2\sum_{I_8(s_1)}(Z^{s_1}_{nj})^2 + \mathcal{O}(n^{-1})] \\ &\leq A_9(s_2 - s) \leq A_9(s_2 - s_1). \end{aligned} \tag{4.21}$$

Hence, by (4.19), (4.20) and (4.21), there is a constant A_{10} such that

$$E[I_{n,11}(s) - I_{n,8}(s_1)]^2 \leq A_{10}(s_2 - s_1). \tag{4.22}$$

Similarly

$$E[I_{n,12}(s_2) - I_{n,9}(s)]^2 \leq A_{11}(s_2 - s_1). \tag{4.23}$$

Note that

$$\begin{aligned} &E((I_{n,11}(s) - I_{n,8}(s_1))I_{n,11}(s_2,s))^2 \leq 2[EI^2_{n,11}(s_2,s)I^2_8(s_1) \\ &+EI^2_{n,11}(s_2,s)(k_n^{-1/2}\sum_{j\neq i\in I_{11}}(I_{\{X_{ni}\leq x(s)\}} - \tilde{G}(x(s)|s)))^2 + O(n^{-1/5})] \\ &\leq A_{12}(s_2 - s_1)^2 \end{aligned} \tag{4.24}$$

Combining (4.22), (4.23) and (4.24), we conclude that (4.17) holds, which finishes the proof of the tightness for $W^0_n(s)$. The covariance structure can be obtained by straight calculation. Theorem 2 is proved.

4.4 PROOF OF THEOREM 3: Let $\mu_{n,1} \leq \mu_{n,2} \leq \cdots \leq \mu_{n,k_n}$ be the order statistics corresponding to $\mu_{ni} \equiv \Psi(X_{ni}, G(X_{ni}|y)), i = 1, 2, \cdots, k_n$. Then μ_{ni} are (conditionally) independent with (conditional) distribution $F^n_1, F^n_2, \cdots, F^n_{k_n}$. Since $\sup_x \Psi(x, G(x|y)) = h^0(y)$, we have

$$F^n_i(h^0(y)|\mathcal{F}) = 1 \text{ for all } 1 \leq i \leq k_n \text{ and } y,$$

where $\mathcal{F} = \sigma(Z_{ni}, 1 \leq i \leq n)$.

Lemma 4. *Under the regularity conditions [B1] up to [B4], for every $\epsilon > 0$*

$$\lim_{n\to\infty} P\{\mu_{n,k_n} \geq h^0(y) - \epsilon/k_n\} = 1 \textit{ uniformly for all } y. \tag{4.25}$$

Proof: by virtue of (3.3) and (3.1), it follows that for every $\epsilon > 0$, there exists an k_n, say $k_{n_0}(\epsilon)$, such that $\epsilon/k_n < h^*$ for all $k_n \geq k_{n_0}(\epsilon)$ and uniformly in y, also there exists $\{\epsilon(k_n)\}$ such that for all $k_n \geq k_{n_0}(\epsilon)$ and uniformly in y,

$$\{x : h(x,y) \geq h^0(y) - \epsilon/k_n\} \supset \{x : |x - x^0(y)| \leq (1/2)\epsilon(k_n)\},$$

where (by (3.1))

$$\lim_{n\to\infty} \epsilon(k_n) = 0,\ but\ \lim_{n\to\infty} k_n\epsilon(k_n) = \infty.$$

Hence, for every $\epsilon > 0$, $k_n \geq k_{n_0}$

$$\begin{aligned}
1 - F_i^n(h^0(y) - \epsilon/k_n|\mathcal{F}) &\geq P\{|X_{ni} - x^0(y)| \leq (1/2)\epsilon(k_n)|\mathcal{F}\} \\
= &\quad P\{x^0(y) - (1/2)\epsilon(k_n) \leq X_{ni} \leq x^0(y) - (1/2)\epsilon(k_n)|\mathcal{F}\} \\
= &\quad P\{X_{ni} \in [x^0(y) - (1/2)\epsilon(k_n), x^0(y) + (1/2)\epsilon(k_n)]|\mathcal{F}\} \\
= &\quad [G^*(x^0(y) + (1/2)\epsilon(k_n)|Z_{ni}) - G^*(x^0(y) - (1/2)\epsilon(k_n)|Z_{ni}) \cong \epsilon(k_n)g^*(x^0(y)|Z_{ni}) \geq \epsilon(k_n)c \\
\Rightarrow &\quad P\{\mu_{n,k_n} \leq h^0(y) - \epsilon/k_n\} = P\{\mu_{ni} \leq h^0(y) - \epsilon/k_n, i = 1,2,\cdots,k_n\} \\
= &\quad E\{P[\mu_{ni} \leq h^0(y) - \epsilon/k_n, i = 1,2,\cdots,k_n]|\mathcal{F}\} = E\{\prod_{i=1}^{k_n} F_i^n(h^0(y) - \epsilon/k_n)|\mathcal{F}\} \\
\leq &\quad E\{[k_n^{-1}\sum_1^{k_n} F_i^n(h^0(y) - \epsilon/k_n|\mathcal{F})]^{k_n}\} \leq (1 - \epsilon(k_n)c)^{k_n} \to 0
\end{aligned}$$

uniformly in y since $k_n\epsilon(k_n) \to \infty$.

By the definition just defined, μ_{n,k_n} must be equal to $\max_{1\leq i\leq k_n} \Psi(X_{ni}, G(X_{ni}|y)) = \Psi(X_{n(r)}, G(X_{n(r)}|y)$, where r is a random variable taking integral values between 1 and k_n, and $X_{n(1)}, \cdots, X_{n(k_n)}$ are the order statistics corresponding to $X_{n1}, \cdots, X_{nk_n}$. Note that r also depends on y.

Lemma 5.

$$X_{n(r)} - X^0(y) \xrightarrow{p} 0 \ \textit{uniformly in } y. \tag{4.26}$$

Proof: from lemma 4 and (3.1), and by the uniqueness of point at which the $\sup_x \Psi(x, G(x|y))$ reaching its maximum, it follows $X_{n(r)} - X^0(y) \xrightarrow{p} 0$ uniformly in y.

Consider a set of intervals

$$I(\delta, y) = \{x : x^0(y) - \delta \leq x^0(y) \leq x^0(y) + \delta\},$$

and let

$$\begin{aligned} J_n(\delta,y) &= \sup_{x\in I(\delta,y)} \{|k_n^{1/2}[G_{nk_n}(x|y_0+n^{-1}k_n s)-G(x|y_0+n^{-1}k_n s)] \\ &\quad - k_n^{1/2}[G_{nk_n}(x^0(y)|y_0+n^{-1}k_n s)-G(x^0(y)|y_0+n^{-1}k_n s)|]. \end{aligned}$$

Lemma 6.

$$J_n(\delta,y)\xrightarrow{p} 0 \text{ as } n\to\infty$$

uniformly in $y\in I_n(y_0)$ *as* $\delta\to 0$.

Proof: Using the same techniques and routines as those in the proofs of theorem 1 and theorem 2, we can prove that

$$\tilde{W}_n(x,s)=k_n^{-1/2}[G_{nk_n}(x|y_0+n^{-1}k_n s)-G(x|y_0+n^{-1}k_n s)] \tag{4.27}$$

is tight for $x\times s\in[0,1]\times[-C/2,C/2]$. Hence , the lemma follows.

The following result is from Suh, Bhattacharyya and Grandage (1970), we state the result without proof.

Lemma 7. *Suppose random variable* $X_i's$ *are independent with distribution functions* F_i*, and*

$$\limsup\int_{-\infty}^{\infty}|x|^p<\infty,\ \textit{then}$$

$$\lim P[max_{1\le i\le n}|X_i|/n^{1/p}>\epsilon]=0 \textit{ for every } \epsilon>0. \tag{4.28}$$

Lemma 8. *Under 3.2,*

$$max_{1\le i\le k_n}k_n^{-1/2}\pi(X_{ni})=o_p(1)$$

Proof: Since X_{ni} are (conditionally) independent with distribution function $G^*(\cdot|Z_{ni})$, from (3.2), the result follows from lemma 7.

Lemma 9. *Under condition [B3], as* $n\to\infty$

$$\max_{1\le i\le k_n}|\Psi(X_{ni},G(X_{ni}|y))-\Psi(X_{ni},G_{nk_n}(X_{ni}|y))|=o_p(1). \tag{4.29}$$

uniformly for y.

Proof: Note that

$$\begin{aligned} &\max_{1\le i\le k_n}|\Psi(X_{ni},G(X_{ni}|y))-\Psi(X_{ni},G_{nk_n}(X_{ni}|y))| \\ &\le\ [\max_{1\le i\le k_n}k_n^{-1/2}\pi(X_{ni})][\max_{1\le i\le k_n}k_n^{1/2}(G(X_{ni}|y)-G_{nk_n}(X_{ni}|y)). \end{aligned}$$

Since for $y \in I_n(y_0)(s \in [-C/2, C/2])$, $k_n^{1/2}(G(X_{ni}|y) - G_{nk_n}(X_{ni}|y)$ is tight, it means $\sup_x k_n^{1/2}(G(X_{ni}|y) - G_{nk_n}(X_{ni}|y)$ is bounded in probability uniformly in y. Thus, the lemma follows from lemma 8.

Now, we are ready to prove the theorem: let

$$\mu_{ni}(y) = \Psi(X_{ni}, G(X_{ni}|y)) \ (1 \le i \le k_n), \qquad \nu_{ni}(y) = \Psi(X_{ni}, G_{nk_n}(X_{ni}|y)) \ (1 \le i \le k_n),$$
$$\nu_n^*(y) = \max_{1 \le i \le k_n} \nu_{ni}(y), \qquad \mu_n^*(y) = \max_{1 \le i \le k_n} \mu_{ni} = \Psi(X_{n(r)}, G(X_{n(r)}|y)),$$

where r is an integral random variable $(1 \le r \le k_n)$. Let $a_n^1(y), a_n^2(y)$ be so defined that

$$h(x^0(y) - a_n^1(y), y) = h(x^0(y) + a_n^2(y), y) = h^0(y) - ck_n^{-1/2} \log k_n; c > 0.$$

By the uniqueness of the maximum point, both a_n^1 and a_n^2 tend to zero for any y.
Noting that $X_{n(r)}(y) \to x^0(y)$ uniformly in y, we have

$$P\{X_{n(r)}(y) \in [x^0(y) - a_n^1(y), x^0(y) + a_n^2(y)]\} \to 1 \text{ uniformly in } y.$$

Consider three random sets

$$\begin{aligned} S_n^1 &= \{X_{ni}(y) : x^0(y) - a_n^1 < X_{ni}(y) < x^0(y) + a_n^2\}, \\ S_n^2 &= \{X_{ni}(y) : X_{ni}(y) \ni [x^0(y) - a_n^1, x^0(y) + a_n^2], but\, |X_{ni}(y) - x^0(y)| < \rho\}, \\ S_n^3 &= \{X_{ni}(y) : |X_{ni}(y) - x^0(y)| > \rho\}. \end{aligned}$$

Let

$$W_n^*(y) = k_n^{-1/2}(G(x^0(y)|y) - G_{nk_n}(x^0(y)|y)), \text{ then}$$

$$\sup_i \{|k_n^{-1/2}(\mu_{ni}(y) - \nu_{ni}(y)) - \Psi_{01}(x^0(y), G(x^0(y)|y))W_n^*(y)|, i \in S_n^1\} = o_p(1).$$

It means $\max_i\{\mu_{ni}(y), i \in S_n^1\} = \mu_n^*(y)$, in probability uniformly in y as $n \to \infty$. Thus,

$$\begin{aligned} \nu_{n1}^*(y) &\equiv \max_i\{\nu_{ni}(y), i \in S_n^1\} \\ &= \mu_n^*(y) + k_n^{-1/2}\Psi_{01}(x^0(y), G(x^0(y)|y))W_n^*(y) + o_p(k_n^{-1/2}) \\ &= h^0(y) + k_n^{-1/2}\Psi_{01}(x^0(y), G(x^0(y)|y))W_n^*(y) + o_p(k_n^{-1/2}). \end{aligned}$$

Note that for every $c > 0$, uniformly in y

$$P[\nu_{n1}^* > h^0(y) - 1/2ck_n^{-1/2} \log k_n] \to 1.$$

To prove the theorem, it is sufficient to show that as $n(k_n) \to \infty$, uniformly in y

$$\sum_{j=2}^{3} P\{\max_i[\nu_{ni}(y), i \in S_n^j] \ge h^0(y) - 1/2k_n^{-1/2} \log k_n\} \to 0$$

(which means $k_n^{1/2}(\nu_n^*(y) - \nu_{n1}^*(y)) \to 0$ in probability as $n \to \infty$ uniformly in y). Note that

$$\begin{aligned} & \sup_i\{|k_n^{1/2}(\mu_{ni}(y) - \nu_{ni}(y))| : i \in S_n^2\} \\ \leq \; & \sup_x\{\Psi_{01}(x, G(x|y)) : x \in [x^0(y) - \rho, x^0(y) + \rho]\} \\ & \times \sup_x\{k_n^{1/2}|G_{nk_n}(x|y) - G(x|y)| : x \in [x^0(y) - \rho, x^0(y) + \rho]\} \\ = \; & \mathcal{O}_p(1), \end{aligned}$$

and $\max_i\{\mu_{ni}(y), i \in S_n^2\} \leq h^0(y) - ck_n^{-1/2}\log k_n$, it gives

$$\max_i\{\nu_{ni}(y) : i \in S_n^2\} \leq h^0(y) - ck_n^{-1/2}\log k_n \text{ in probability.}$$

By the distinguishability of the maximum and lemma 9,

$$\max_i\{\nu_{ni}(y) : i \in S_n^3\} \leq h^0(y) - \rho^*$$

for some $\rho^* > 0$ uniformly in y, which finishes the proof of the theorem.

The corollaries are the direct consequences of the theorem.

References

[1] Bhattacharya, P. K. (1974), Convergence of sample paths of normalized sums of induced order statistics. *Ann. Stat.*, **2**,1034-1039.

[2] Bhattacharya, P. K. and Gangopadhyay, A. K. (1990), Kernel and nearest neighbor estimation of a conditional quantile. *Ann. Stat.*, **18**, 1400-1415.

[3] Bhattacharya, P. K. and Mack, Y. P. (1987), Weak convergence of k-NN density and regression estimators with varying k and applications. *Ann. Stat.*, **15**,976-994.

[4] Bickel, P. J. and Wichura, M. J. (1971), Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Stat.*, **42**,1656-1670.

[5] Billingsley, P.(1968), Convergence of probability measures. *Wiley, New York*

[6] Breiman, L. (1968), Probability. *Addison-Wesley, Reading Mass.*.

[7] Crowder, M.J., Kimber, A.C., Smith, R.L. and Sweeting, T.J. (1991). Statistical analysis of reliability data. *Chapman & Hall*.

[8] Csőrgő, M. and Révész, P. (1981) Strong approximations in probability and statistics, *Academic Press*.

[9] Daniels, H.E. (1945) The statistical theory of the strength of bundles of threads. *Proc. Roy. Soc. London, Ser. A*, bf 183,405-435.

[10] Daniels, H.E. (1974) The maximum size of a closed epidemic. *Adv. Prob.*, **6**, 607-621.

[11] Daniels, H.E. and Skyrme, T.H.R. (1985) The maximum of a random walk whose mean path has a maximum. *Adv. Appl. Prob.*, **17** 85-99.

[12] Daniels, H.E. (1989) The maximum of a gaussian process whose mean path has a maximum, with an application to the strength of bundles of fibres. *Adv. Appl. Prob.*, **21**, 315-333.

[13] Mack, Y.P. and Rosenblatt (1979) Multivariate k-nearest neighbor density estimates. *J. Multivariate Analysis*, **9**, 1-15.

[14] Neuhaus, Georg (1971) On weak convergence of stochastic processes with multidimensional time parameters. *Ann. Math. Stat.*, **42**,1285-1295.

[15] Phoenix, S.L. and Taylor, H.M. (1973) The asymptotic strength distribution of a general fibre bundle. *Adv. Appl. Prob.*, **5**, 200-216.

[16] Pollard, David (1991) Empirical processes: theory and applications. *NSF-CBMS Regional Conference Series in Probability and Statistics, Volume 2.*

[17] Sen, P.K. (1981) Sequential nonparametrics. *Wiley*

[18] Sen, P.K. (1993) Regression quantiles in nonparametric regression. *J. Nonparamet. Statist.* **3**, 237-253.

[19] Sen, P.K., Bhattacharyya, B.B. and Suh, M.W. (1973) Limiting behavior of the extremum of certain sample functions. *Ann. Prob.*, **1**, 297-311.

[20] Smith, R.L. (1982) The Asymptotic distribution of the strength of a series-parallel system with equal load-sharing. *Ann. Prob.*, **10**, 137-171.

[21] Suh, M.W. Bhattacharyya, B.B. and Grandage, A. (1970) On the distributions of the strength of a bundle of filaments. *J. Ann. Prob.*, **7**, 712-720.

[22] Zhou, Zhenwei (1996) Limiting behavior of extrema of certain conditional sample functions and some applications. *Ph.D dissertation*, Department of Statistics, UNC-CH.

Dept. of Statistics, University of California at Los Angeles, Los Angeles, CA 90095-1554

Dept. of Statistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260