NONPARAMETRIC INFERENCE ON THE DIFFERENCE OF LOCATION
PARAMETERS OF CORRELATED VARIABLES FROM
FRAGMENTARY SAMPLES

K. F. CHENG

(Received Oct. 4, 1985; revised Mar. 13, 1986)

Summary

In this paper, two types of robust estimators and approximate
confidence intervals for the difference of location parameters of corre-
lated random variables are proposed and investigated when some ob-
servations are missing. It is shown that the suggested estimators are
consistent and asymptotically normally distributed. In addition, the
proposed approximate confidence intervals are also shown to enjoy some
nice asymptotic properties.

1. Introduction

The problem of estimation of the difference of location parameters
of two correlated random variables with incomplete paired observations
can be described as follows: Let \((X, Y)\) be a random vector with con-
tinuous joint distribution \(H(x, y)\) and marginal distributions \(F_1(x)\) and
\(F_2(y)\) for \(X\) and \(Y\) respectively such that \(F_1(x) = F_2(x - \theta), \theta \in \mathbb{R}\). Assume
that we have a fragmentary random sample \(\{(X_1, Y_1), \ldots, (X_n, Y_n),
(X_{n+1}, \cdot), \ldots, (X_{n+m}, \cdot), (\cdot, Y_{n+1}), \ldots, (\cdot, Y_{n+l})\}\) observed from \((X, Y)\), where
"\(\cdot\)" denotes a missing observation. Then the problem is how to use
these data wisely to make an inference on the shift parameter \(\theta\).

In the situation that \(H(x, y)\) is a bivariate normal distribution, the
problem of estimation of the mean difference \(\theta\) has been extensively
studied; see, in particular, papers by Anderson [1], Lin [10], [11], Lin
and Stivers [12], Mehta and Gurland [13] and Wilks [16]. Assuming
that \(X\) and \(Y\) are linearly related, Gupta and Rohatgi [5] proposed esti-
mators expressed as linear combinations of fragmentary sample means.
Recently Wei [15] proposed a nonparametric approach using the median
of all possible differences \(Y_j - X_i\), \(i=1, \ldots, n+m, j=1, \ldots, n+l\) and

Key words and phrases: Consistent, contaminated bivariate normal distributions, efficiency, linear combinations of sample quantiles, M-estimators.
(Y_i+Y_j-X_i-X_j)/2, 1 \leq i \leq j \leq n$. He compared this median estimator with the estimators suggested by Lin and Stivers [12], Gupta and Rohatgi [5] and a naive estimator $\sum_{j=1}^{n+l} Y_j/(n+l) - \sum_{i=1}^{n+m} X_i/(n+m)$. The results show that the median estimator performs quite satisfactory under bivariate normal distribution and other estimators perform very poorly under Gumbel’s bivariate exponential distribution.

In this article, two large classes of nonparametric estimators of $\theta$ are studied in detail. Each class contains the median estimator as a special case. The first class of estimators is the class of linear combinations of sample quantiles derived from the empirical distributions $\hat{G}(t) = M^{-1}\left[ \alpha \sum_{i=1}^{n+m} \sum_{j=1}^{n+l} I(Y_j-X_i \leq t) + \beta \sum_{1 \leq i \leq j \leq n} I(Y_i+Y_j-X_i-X_j \leq 2t) \right]$, where $\alpha$ and $\beta$ are two nonnegative integers with $\alpha + \beta > 0$, $M = \alpha(n+m)(n+l) + (\beta n(n+1)/2)$, and $I(\cdot)$ is the usual indicator function. More specifically, estimators $\theta^*$ from this class can be expressed as

$$\theta^* = \sum_{i=1}^{k} c_i \hat{\xi}_{p_i},$$

where $c_i$ are nonnegative constants, $0 < p_i < 1$, $i = 1, \ldots, k$ and $\hat{\xi}_{p_i}$ are defined as

$$\hat{\xi}_{p_i} = \inf \{ t : \hat{G}(t) \geq p_i \}.$$

Note that the estimators $\theta^*$ considered here and below depend on $\alpha$ and $\beta$. If we are unable to identify the pairing of $X_i$ with $Y_i$, $i = 1, \ldots, n$, then we simply choose $\beta = 0$. In Section 2, we see that if we choose $c_i$ and $p_i$ properly, then $\theta^*$ will be a consistent estimator of $\theta$. It is believed that suitable members of this class, such as $0.3\hat{\xi}_{1/3} + 0.4\hat{\xi}_{1/2} + 0.3\hat{\xi}_{1/4}$ and $0.25\hat{\xi}_{1/4} + 0.5\hat{\xi}_{1/2} + 0.25\hat{\xi}_{1/4}$ are more distributionally robust and insensitive to spurious observations. The second group of estimators considered here is the class of $M$-estimators. An $M$-estimator for $\theta$ is a solution $\hat{\theta}$ of the equation

$$\alpha \sum_{i=1}^{n+m} \sum_{j=1}^{n+l} F(Y_j-X_i-c) + \beta \sum_{1 \leq i \leq j \leq n} F\left( \frac{W_i + W_j}{2} - c \right) = 0,$$

where $W_i = Y_i - X_i, \ i = 1, \ldots, n$. The function $F$ is usually skew-symmetric about 0. Typical $F$ considered in robust estimation are the Huber [9] family $F(x) = \min (k, \max (-k, x))$ and Hampel’s “redescenders” [6] and [7] etc. In Section 2, sufficient conditions are given to ensure that $\hat{\theta}$ is a consistent estimator of $\theta$ with other good asymptotic properties.

In Section 3, we also study the important problem of constructing approximate confidence intervals for $\theta$. Two approaches are examined. The first approach is based on the sample quantiles. The second method
is to derive approximate confidence intervals from \( M \)-estimates; see Boos [2].

2. Nonparametric estimators of \( \theta \)

Let \( K(t) \) and \( L(t) \) be the distribution functions of \( Y_i - X_i \) and \( (W_1 + W_2)/2 \), respectively. Define \( N = 2n + m + l \) and \( \lambda_{1N} = m/N \), \( \lambda_{2N} = l/N \) and \( \lambda_{3N} = n/N \). We assume that there exists a constant \( \lambda \) such that \( 0 < \lambda \leq \lambda_{3N} \) for all large \( N \). In addition, we let \( \lambda_{iN} \to \lambda_i \), \( i = 1, 2, 3 \). Throughout this paper we assume that both \( K \) and \( L \) are symmetric with symmetry point \( \theta \). This implies that the distribution function \( G(t) = \gamma K(t) + \delta L(t) \), where \( \gamma = 1 - \delta = a(\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3)/(a(\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3) + \beta(\lambda_2/2)) \), is also symmetric with symmetry point \( \theta \). The definitions of the estimators \( \hat{\theta}^* \) and \( \hat{\theta} \) are motivated by the fact that the empirical distribution \( \hat{G}(t) \) is a reliable estimator of the distribution \( G(t) \). Thus if we choose constants \( c_i > 0 \) and \( 0 < p_i < 1 \) \( (i = 1, 2, \ldots, k) \) such that \( \sum_{i=1}^{k} c_i = 1 \) and \( \sum_{i=1}^{k} c_i \hat{\xi}_{p_i} = \theta \), where \( \hat{\xi}_{p_i} = \inf \{ t : \hat{G}(t) \geq p_i \} \), then \( \hat{\theta}^* = \sum_{i=1}^{k} c_i \hat{\xi}_{p_i} \) will be a consistent estimator of \( \theta \). For examples, \( \{c_i = 1, p_i = 1/2\} \) or \( \{c_1 = 0.3, c_2 = 0.4, c_3 = 0.3, p_1 = 1/3, p_2 = 1/2, p_3 = 2/3\} \) or \( \{c_1 = 0.25, c_2 = 0.5, c_3 = 0.25, p_1 = 1/4, p_2 = 1/2, p_3 = 3/4\} \). On the other hand, if, in addition to some mild conditions, we let \( F \) be a skew-symmetric function such that there is only one solution \( \theta \) for \( \int_{-\infty}^{\infty} F(x-t)dG(t) = 0 \), then \( \hat{\theta} \) can also be shown to be a consistent estimator of \( \theta \).

In what follows, we first establish an almost sure representation for sample quantiles \( \hat{\xi}_p \) (Theorem 2.1). Using this representation we are able to obtain some useful asymptotic properties for \( \hat{\theta}^* \) (Corollary 2.1). Asymptotic normality and other properties of the \( M \)-estimators \( \hat{\theta} \) are stated in Theorems 2.2 and 2.3.

**Theorem 2.1.** Let \( 0 < p < 1 \) and \( p_N = p + O(N^{-1/2}(\log N)^{-1/2}), \quad N \to \infty \). Assume that the distributions \( K(t) \) and \( L(t) \) are twice differentiable at \( \xi_p \) with \( K'(\xi_p)L'(\xi_p) > 0 \), and \( \partial \) satisfies a Lipschitz condition of order 1, and \( \lambda_{iN} = \lambda_i + O(N^{-1/2}(\log N)^{1/2}) \), \( i = 1, 2, 3 \), \( N \to \infty \). Then with probability 1,

\[
\hat{\xi}_{p_N} - \xi_p = \frac{p_N - \hat{G}(\xi_p)}{\gamma_NK'(\xi_p) + \alpha N L'(\xi_p)} + O(N^{-1/4}(\log N)^{1/4}), \quad N \to \infty,
\]

where

\[
\gamma_N = 1 - \delta_N = a(n + m)(n + l)/[a(n + m)(n + l) + \beta n(n + 1)/2] \quad \to \gamma, \quad N \to \infty.
\]

**Proof.** See Appendix.
Using this almost sure representation and asymptotic properties of \( \hat{G}(t) \) we can easily establish the following

**Corollary 2.1.** (a) Let \( \lambda_{jN} = \lambda_j + O(N^{-1/2}(\log N)^{1/2}) \), \( N \to \infty \), for \( j = 1, 2, 3 \). Assume that for each \( i = 1, \ldots, k \), \( K(t) \) and \( L(t) \) are both twice differentiable at \( \xi_{pi} \) with \( K'(\xi_{pi})L'(\xi_{pi}) > 0 \) and \( F_1 \) satisfies a Lipschitz condition of order 1. Then with probability 1,

\[
\frac{\sqrt{N}}{\sqrt{\log \log N}} (\theta^* - \theta) = O(1), \quad N \to \infty.
\]

(b) If, in addition to the conditions assumed in (a), we let \( \lambda_{jN} = \lambda_j + O(N^{-1/2}) \), \( N \to \infty \), \( j = 1, 2, 3 \). Then

\[
\sqrt{N}(\theta^* - \theta) \to N(0, \sigma^2), \quad N \to \infty,
\]

where \( 0 < \sigma^2 < \infty \) is the asymptotic variance of \( \sqrt{N} \theta^* \).

**Remark.** If \( c_1 = 1 \) and \( p_1 = 1/2 \), then the value of \( \sigma^2 \) is given in Wei [15]. In general, the asymptotic variances of other linear combinations of sample quantiles are quite complicated. Fortunately, in Section 3, we are able to suggest two types of approximate confidence intervals without requiring the knowledge of the value of \( \sigma^2 \).

We now focus our attention on the class of \( M \)-estimators. For a given function \( \mathcal{F}(x) \), we put \( \lambda_{F}(t) = \int_{-\infty}^{\infty} \mathcal{F}(x-t) d F(x) \) for any distribution \( F \). It is clear that if \( \mathcal{F} \) is skew-symmetric about 0, then \( \theta \) is a solution of \( \lambda_{F}(t) = 0 \) and \( \lambda_{L}(t) = 0 \), and hence also a solution of \( \lambda_{Q}(t) = 0 \). Let the distribution function of \( W_1 = Y_1 - X_1 \) be denoted as \( F_1 \), and we defined:

\[ \mathcal{F}_1(t) = \int_{-\infty}^{\infty} \mathcal{F}(y-t) d F_1(y), \]

\[ \mathcal{F}_2(t) = \int_{-\infty}^{\infty} \mathcal{F}(t-x) d F_1(x), \]

and

\[ \mathcal{F}_3(t) = \int_{-\infty}^{\infty} \mathcal{F}\left( \frac{w}{2} + t \right) d F_3(w). \]

Throughout this paper, we assume \( \mathcal{F}_i(t) \), \( i = 1, 2, 3 \), are measurable functions. Also, we define the following functions involved in the expression of the asymptotic variance of \( \sqrt{N} \hat{\theta} \):

\[ q_1(s) = \int_{-\infty}^{\infty} \frac{\mathcal{F}_1(y-s) d F_1(y)}{\mathcal{F}_1(y)}, \]

\[ q_2(s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}_2(y-x-s) d H(x, y), \]
\[ q_i(s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}_i(x-s) \mathcal{F}_i(y-s) dH(x, y), \]

\[ q_i(s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}_i(x-s) \mathcal{F}_i \left( \frac{y-x}{2} - s \right) dH(x, y), \]

\[ q_i(s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}_i(y-s) \mathcal{F}_i \left( \frac{y-x}{2} - s \right) dH(x, y). \]

In Theorem 2.2, two sets of sufficient conditions are assumed in order to establish asymptotic normality of the $M$-estimators. The method of the proof is based on that given by Huber ([9], Lemmas 4 and 5) in conjunction with the projection technique.

**Theorem 2.2.** (a) Assume the following conditions:

(i) $\mathcal{F}$ is bounded, nondecreasing and skew-symmetric about 0;

(ii) $\lambda_0(t)$ is differentiable at $t=\theta$ with $\lambda_0'(\theta)<0$;

(iii) $\lambda_{iN} = \lambda_i + O \left( \frac{1}{\sqrt{N}} \right)$, $N \to \infty$, $i=1, 2, 3$;

(iv) $q_j(s)$, $j=1, \ldots, 5$, are continuous at $s=\theta$.

Then for any solution sequence $\hat{\theta}$ of the equation $\lambda_0(t)=0$,

\[ \sqrt{N} (\hat{\theta} - \theta) \overset{d}{\to} N(0, \sigma^2), \quad N \to \infty, \]

where

\[ \sigma^2 = \frac{\lambda_0^2}{(\lambda_0'(\theta))^2} \]

and

\[ \sigma_0^2 = \gamma^2 \left( \frac{1}{\lambda_1 + \lambda_3} + \frac{1}{\lambda_2 + \lambda_3} \right) q_1(\theta) + \frac{4\gamma^2}{\lambda_3} q_2(\theta) + \frac{2\gamma^2}{\lambda_1 + \lambda_2} q_3(\theta) + 4\gamma \left( \frac{q_1(\theta)}{\lambda_1 + \lambda_3} + \frac{q_2(\theta)}{\lambda_2 + \lambda_3} \right). \]

(b) Let $\hat{\theta}$ be a solution sequence of $\lambda_0(t)=0$ satisfying $\hat{\theta} \overset{p}{\to} \theta$, $N \to \infty$ and assume the following conditions:

(i) $\theta$ is a root of $\lambda_0(t)=0$ and $\lambda_L(t)=0$;

(ii) $\mathcal{F}$ has a bounded, uniformly continuous derivative $\mathcal{F}'$;

(iii) $\int_{-\infty}^{\infty} \frac{\partial \mathcal{F}(x-t)}{\partial t} dK(x)$ and $\int_{-\infty}^{\infty} \frac{\partial \mathcal{F}(x-t)}{\partial t} dL(x)$ are finite and

\[ \int_{-\infty}^{\infty} \frac{\partial \mathcal{F}(x-t)}{\partial t} dK(x) + \int_{-\infty}^{\infty} \frac{\partial \mathcal{F}(x-t)}{\partial t} dL(x) \neq 0. \]

Then

\[ \sqrt{N} (\hat{\theta} - \theta) \overset{d}{\to} N(0, \sigma^2), \quad N \to \infty, \]
where
\[ \sigma^2 = \sigma_0^2 \left[ \gamma \int_{-\infty}^{\infty} \frac{\partial \psi(x-t)}{\partial t} \bigg|_{t=\theta} dK(x) + \delta \int_{-\infty}^{\infty} \frac{\partial \psi(x-t)}{\partial t} \bigg|_{t=\theta} dL(x) \right]^2. \]

**Proof.** See Appendix.

**Remarks.**

1. In (a), condition (iii) can be relaxed if we assume that \( \lambda_+(t) \) and \( \lambda_-(t) \) are both differentiable at \( t=\theta \) with \( \lambda'_+(\theta) < 0 \).

2. Regarding the condition imposed on \( \hat{\theta} \) in (b), we note that if \( \Psi \) is a continuous and bounded function, \( \lambda_+(\theta) = 0 \) and that at \( \theta \), \( \lambda_+(t) \) changes sign only once in a neighborhood of \( \theta \). Then there is a sequence of solution \( \hat{\theta} \) such that with probability 1,
\[ \hat{\theta} \to \theta, \quad N \to \infty; \]
see Boos and Serfling ([3], Theorem 2.1).

3. In condition (ii) of (b) we assume that \( \Psi' \) is bounded. Actually this condition can be replaced by only assuming that
\[
E \left[ \int_{-\infty}^{\infty} \frac{\partial \Psi(Y_1 - x - t)}{\partial t} \bigg|_{t=\theta} dF_1(x) \right]^2 < \infty,
\]
\[
E \left[ \int_{-\infty}^{\infty} \frac{\partial \Psi(y - X_1 - t)}{\partial t} \bigg|_{t=\theta} dF_2(y) \right]^2 < \infty,
\]
and
\[
E \left[ \int_{-\infty}^{\infty} \frac{\partial \Psi((W_1 + w)/2 - t)}{\partial t} \bigg|_{t=\theta} dF_3(w) \right]^2 < \infty.
\]

In view of the proof for case (b) of Theorem 2.2 (see Appendix), we easily see that if, with probability 1, \( \hat{\theta} \to \theta, \ N \to \infty \), then under regularity conditions similar to (a) of Corollary 2.1, we can obtain a stronger result:
\[ \frac{\sqrt{N}}{\sqrt{\log \log N}}(\hat{\theta} - \theta) = O(1), \quad N \to \infty, \]
with probability 1. The same result is also obtainable for case (a) by using the following almost sure representation, similar to that given in Theorem 2.1, for \( M \)-estimators. This representation is not only useful for the purpose of establishing the LiL type result for \( \hat{\theta} \) but also valuable for constructing confidence intervals for \( \theta \). A weaker version of a similar representation was derived by Boos [2].

Let \( \Psi(t) \) be nondecreasing, left continuous, and strictly positive (negative) for large positive (negative) values of \( t \), and define, for any distribution \( F_1 \),
\[ \gamma(d) = -\int_{-\infty}^{\infty} \mathcal{F}(x - d)dF(x) = -\lambda(d) \]

and

\[ \gamma^{-1}(t) = \inf \{ \delta : \gamma(d) \geq t \}, \quad t \in (\inf_{x \in R} \gamma(x), \sup_{x \in R} \gamma(x)) \]
3. Confidence intervals for the difference of locations

Here we consider two methods of determining approximate confidence intervals for \( \theta \) using Theorems 2.1 and 2.3, respectively.

According to Corollary 2.1, for the median estimator \( \hat{\theta} = \hat{\xi}_{1/2} \) we have \( \sqrt{N}(\theta^* - \theta) \overset{d}{\rightarrow} N(0, \sigma^2) \), \( N \to \infty \). If \( \sigma^2 \) is known and let \( Z_\alpha \) denote the 100(\( 1 - \alpha \)) percentile of the standard normal distribution, then the confidence interval \( I_N = (\theta^* - Z_\alpha(\sigma/\sqrt{N}), \theta^* + Z_\alpha(\sigma/\sqrt{N})) \) has confidence coefficient converging to \( 1 - 2\alpha \) as \( N \to \infty \). Unfortunately, in most cases, \( \sigma^2 \) is not known and therefore the above approach is useless.

To propose a useful confidence interval for \( \theta \), we need to estimate the asymptotic variance of \( \sqrt{N}\theta^* \).

From (2.1), it is not difficult to see that
\[
\sqrt{N}(\theta^* - \theta)/\sigma_N \overset{d}{\rightarrow} N(0, 1), \quad N \to \infty,
\]
where
\[
\sigma_N^2 = \frac{1}{\vartheta_N} (\gamma_N \kappa_1(\theta) + \delta_N L_1(\theta))^2,
\]
where
\[
\vartheta_N^2 = \left[ \frac{1}{12} \left( \frac{\gamma_N^2}{\lambda_1 + \lambda_N} + \frac{\gamma_N^3}{\lambda_2 + \lambda_N} + \frac{4\gamma_N^2}{\lambda_N} \right) + \frac{\gamma_N^2 \lambda_{2N}}{\lambda_N + \lambda_{2N}} \left( \frac{1}{2} - 2\theta_1 \right) + \frac{\gamma_N}{\lambda_N + \lambda_{2N}} \left( \theta_2 - \frac{1}{4} \right) + \frac{\gamma_N}{\lambda_N + \lambda_{2N}} \left( \frac{1}{4} - \theta_3 \right) \right],
\]
\[
\theta_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_1(x)F_1(y-\theta)dH(x, y),
\]
\[
\theta_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_1(x)F_1(2\theta-(y-x))dH(x, y),
\]
and
\[
\theta_3 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_1(y-\theta)F_1(2\theta-(y-x))dH(x, y).
\]
(Note that if \( H(x, y) = H(y, x) \) for all \((x, y) \in \mathbb{R}^2\) then \( \theta_1 = 1/2 - \theta_2 \).) Also, in view of the asymptotic variance \( \sigma^2 \) of \( \sqrt{N}\theta^* \) derived by Wei [15], we have \( \sigma_N^2 \to \sigma^2, \quad N \to \infty \).

Using our method it is enough to estimate simple parameters \( \theta_i \). Let us first define
\[
\hat{F}_i(t) = (n + m)^{-1} \sum_{i=1}^{n+m} I(X_i \leq t) \quad \text{and} \quad \hat{F}_i(t) = -n^{-1} \sum_{i=1}^{n} I(Y_i - X_i \leq t).
\]

Then natural estimators of \( \theta_i \) are \( \hat{\theta}_i \):
\[ \hat{\theta}_i = n^{-1} \sum_{i=1}^{n} \hat{F}'_i(X_i) \cdot \hat{F}'_i(Y_i - \theta^*) , \]
\[ \hat{\theta}_i = n^{-1} \sum_{i=1}^{n} \hat{F}'_i(X_i) \cdot \hat{F}'_i(2\theta^* - Y_i + X_i) \]

and
\[ \hat{\theta}_i = n^{-1} \sum_{i=1}^{n} \hat{F}'_i(Y_i - \theta^*) \cdot \hat{F}'_i(2\theta^* - Y_i + X_i) . \]

Their asymptotic properties are described in

**Corollary 3.1.** If \( F_1 \) and \( F_2 \) satisfy a Lipschitz condition of order 1 and there is a sequence of positive integers \( \alpha_n \to 0 \), \( N \to \infty \), such that with probability 1, \( \alpha_n(\theta^* - \theta) \to 0 \), \( N \to \infty \). Then with probability 1,
\[ \beta_n(\hat{\theta}_i - \theta) \to 0 , \quad N \to \infty , \quad i=1, 2, 3 , \]

where
\[ \beta_n = \min \left( \alpha_n , N^{1/2}(\log N)^{-1/2} \right) . \]

Define \( p_{1N} = -Z \cdot \sigma_n/\sqrt{N} \) and \( p_{2N} = Z \cdot \sigma_n/\sqrt{N} \), where \( \sigma_n \) is the strong consistent estimator of \( \sigma^2 \) obtained by replacing \( \theta \) by \( \hat{\theta}_i \) in the definition of \( \sigma^2 \). Then the suggested distribution-free interval is \( I^*_n = (\hat{\epsilon}_{p_{1N}}, \hat{\epsilon}_{p_{2N}}) \). The advantage of this approach is that we don't need to estimate \( K'(\theta) \) and \( L'(\theta) \). Furthermore, applying Corollary 3.1 and Theorem 2.1, it can be shown that the confidence coefficient of \( I^*_n \to 1 - 2\alpha \) as \( N \to \infty \).

Also,
\[ \sqrt{N} \cdot \text{(length of } I^*_n) \to 2 \sigma Z , \quad N \to \infty . \]

Next we turn our attention to the second approach using the result of Theorem 2.1. From (2.3) we see that under some suitable conditions,
\[ \frac{\sqrt{N}}{\sigma_n}(\gamma_0(p_n) - \theta) \xrightarrow{d} N(0, 1) , \quad N \to \infty , \]

where
\[ (\sigma'_{n})^2 = \frac{\sigma_n^2}{\lambda^2_{\alpha_n}(\theta)} , \]
\[ \sigma_n^2 = \left[ \gamma_n^2 \left( \frac{1}{\lambda_{1N} + \lambda_{2N}} + \frac{1}{\lambda_{2N} + \lambda_{3N}} \right) q_1(\theta) + \frac{4\sigma_n^2}{\lambda_{3N}} q_2(\theta) \right] \]
\[ + \frac{2\gamma_n \lambda_{2N} q_2(\theta)}{\lambda_{1N} + \lambda_{2N}} + 4\gamma_n \lambda_{3N} q_2(\theta) \left( \frac{q_1(\theta)}{\lambda_{1N} + \lambda_{4N}} + \frac{q_3(\theta)}{\lambda_{2N} + \lambda_{3N}} \right) . \]

Hence the interval \( (\gamma_0(p_n) - Z \cdot (\sigma'_n/\sqrt{N}), \gamma_0(p_n) + Z \cdot (\sigma'_n/\sqrt{N})) \) has confi-
dance coefficient converging to $1-2\alpha$ as $N \to \infty$. Again, $\sigma_0$ and $\lambda_{\alpha N}(\theta)$ are in general not known and thus, as above, we have to find a way to estimate $q_i(\theta)$. (It is not necessary to estimate $\lambda_{\alpha N}(\theta)$.)

To estimate $q_i(\theta)$, let us first consider estimation of $\Psi_i(t)$, $i = 1, 2, 3$. The estimator of $\Psi_i(t)$ considered here is $\hat{\Psi}_i(t) = (n+l)^{-1} \sum_{j=1}^{n+l} \Psi(Y_j - t)$ and the estimator of $\Psi_2(t)$ is $\hat{\Psi}_2(t) = (n+m)^{-1} \sum_{i=1}^{n+m} \Psi(t - X_i)$ and the estimator of $\Psi_3$ is $\hat{\Psi}_3(t) = n^{-1} \sum_{i=1}^{n} \Psi\left(\frac{Y_i - X_i}{2} + t\right)$. The asymptotic properties of these estimators are stated in

**Corollary 3.2.** Let $\Psi$ be continuous and of finite variation $\|\Psi\|_\nu$. Then with probability 1,

$$N^{1/2} \left(\log N\right)^{-1/2} \sup_{t \in R} \left|\hat{\Psi}_i(t) - \Psi_i(t)\right| \to 0, \quad N \to \infty, \quad i = 1, 2, 3.$$

The results of Corollary 3.2 lead us to consider the following type of estimators for $q_i(\theta)$:

$$\hat{q}_i(\theta) = (n+l)^{-1} \sum_{j=1}^{n+l} \hat{\Psi}_j^2(Y_j - \hat{\theta}),$$

$$\hat{q}_2(\theta) = n^{-1} \sum_{i=1}^{n} \hat{\Psi}_i^2\left(\frac{Y_i - X_i}{2} - \hat{\theta}\right),$$

$$\hat{q}_3(\theta) = n^{-1} \sum_{i=1}^{n} \hat{\Psi}_i(X_i - \hat{\theta}) \hat{\Psi}_i(Y_i - \hat{\theta}),$$

$$\hat{q}_4(\theta) = n^{-1} \sum_{i=1}^{n} \hat{\Psi}_i(X_i - \hat{\theta}) \hat{\Psi}_i\left(\frac{Y_i - X_i}{2} - \hat{\theta}\right),$$

$$\hat{q}_5(\theta) = n^{-1} \sum_{i=1}^{n} \hat{\Psi}_i(Y_i - \hat{\theta}) \hat{\Psi}_i\left(\frac{Y_i - X_i}{2} - \hat{\theta}\right),$$

where $\hat{\theta} = \gamma_{\hat{\theta}}^{-1}(0)$. Their asymptotic properties are summarized in

**Corollary 3.3.** Let the conditions of Corollary 3.2 be satisfied and $\Psi$ be Lipschitz of order 1. If there is a sequence of positive real numbers $\alpha' \to \infty$, $N \to \infty$ such that with probability 1, $\alpha' \to 0$, $N \to \infty$, then with probability 1,

$$\beta' \left|\hat{q}_i(\theta) - q_i(\theta)\right| \to 0, \quad N \to \infty, \quad i = 1, \ldots, 5,$$

where $\beta' = \min(\alpha', N^{1/4} \left(\log N\right)^{-1/2})$.

In view of the properties of $\hat{q}_i(\theta)$ described in Corollary 3.3, a natural estimator $\hat{\sigma}_i^2$ of $\sigma_i^2$ can be obtained by replacing $q_i(\theta)$ by $\hat{q}_i(\theta)$, $i = 1, \ldots, 5$, in the definition of $\sigma_i^2$. If we do so and if the conditions of
Corollary 3.3 are satisfied then with probability 1,
\[ \beta_N |\hat{\theta}_{n} - \theta_0| \to 0, \quad N \to \infty . \]

Based on the result, we now propose a second type of confidence intervals for \( \theta \). We define two quantities \( q_{1N} = -Z_\alpha (\hat{\sigma}_n / \sqrt{N}) \) and \( q_{2N} = + Z_\alpha (\hat{\sigma}_n / \sqrt{N}) \). Then the distribution-free confidence interval \( \hat{I}_N = (z^{-1}(q_{1N}), z^{-1}(q_{2N})) \) has some desirable properties. That is, the confidence coefficient of \( \hat{I}_N \to 1 - 2\alpha, N \to \infty \) and \( \sqrt{N}(\text{length of } \hat{I}_N) \to 2Z_\alpha \sigma_1, N \to \infty \), where \( \sigma_1 \) is defined in part (a) of Theorem 2.2.

**National Central University, Taiwan**

**REFERENCES**


Appendix

The proof of Theorem 2.1 consists of the following two lemmas. In each of these lemmas, suitable conditions of the theorem are in action.

**Lemma A1.** Let $0 < p < 1$ and $p_N = p + O(N^{-1/2}(\log N)^{1/2})$, $N \to \infty$. If $K'(\xi_p) L'(\xi_p) \geq 0$ and $\lambda_i = \lambda_i + O(N^{-1/2}(\log N)^{1/2})$, $N \to \infty$, $i = 1, 2, 3$, then with probability 1,

\[ \xi_{p_N} = \xi_p + O(N^{-1/3}(\log N)^{1/3}), \quad N \to \infty. \]

**Proof.** Let $\varepsilon_N = c_0 N^{-1/3}(\log N)^{1/3}$, where $c_0$ is some positive constant whose value will be specified later. Define

\[ \hat{K}(t) = [(n + m)(n + l)]^{-1} \sum_{i=1}^{n+m} \sum_{j=1}^{n+l} I(Y_j - X_i \leq t), \]

\[ \hat{L}(t) = [n(n + 1)/2]^{-1} \sum_{1 \leq i < j \leq n} I(W_i + W_j \leq 2t), \]

\[ \hat{T}_1(t) = (n + m)^{-1} \sum_{i=1}^{n+m} [F_i(X_i + t) - K(t)], \]

\[ \hat{T}_3(t) = (n + l)^{-1} \sum_{j=1}^{n+l} [F_i(Y_i - t) - K(t)], \]

where $\bar{F}_1(t) = 1 - F_1(t)$. Write

\[ \hat{S}(t) = [\hat{K}(t) - K(t)] - \hat{T}_1(t) - \hat{T}_3(t) = [(n + m)(n + l)]^{-1} \sum_{i=1}^{n+m} \sum_{j=1}^{n+l} g(X_i, Y_j, t), \]

where

\[ g(X_i, Y_j, t) = [I(Y_j - X_i \leq t) - K(t)] - [F_i(X_i + t) - K(t)] - [\bar{F}_1(Y_j - t) - K(t)], \]

and realize that for any positive integer $d$ and any sequence of constants $t_N \in R$, $\mathbb{E} [\prod_{k=1}^d g(X_{i_k}, Y_{i_k}, t_N)] = 0$, if there is a subscript which appears only once in $\{i_1, \ldots, i_d, j_d\}$. Consequently,

\[ (A1) \quad \mathbb{E} [\hat{S}(t_N)]^d = O(N^{-d}), \quad N \to \infty, \]

for any sequence of real numbers $t_N$.

Now we consider

\[ (A2) \quad P(\xi_{p_N} > \xi_p + \varepsilon_N) \]

\[ \leq P \left( -\gamma_N [\hat{K}(\xi_p + \varepsilon_N) - K(\xi_p + \varepsilon_N)] > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{2} \right) + P \left( -\alpha_N [\hat{L}(\xi_p + \varepsilon_N) - L(\xi_p + \varepsilon_N)] > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{2} \right) \]

\[ = U_{1N} + U_{2N}, \quad \text{say}, \]
where
\[ G_N(t) = \gamma_N K(t) + \delta_N L(t) . \]

It is noted that if \( \alpha = 0 \) or \( \beta = 0 \), then \( U_{1N} = 0 \) or \( U_{2N} = 0 \) for large \( N \). So, in the following we assume \( \alpha \neq 0 \) and \( \beta \neq 0 \). Now regarding the first term on the right-hand side of (A2), we have
\[
U_{1N} \leq P \left( -\gamma_N \hat{S}(\xi_p + \varepsilon_N) > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{6} \right) \\
+ P \left( -\gamma_N \hat{T}_1(\xi_p + \varepsilon_N) > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{6} \right) \\
+ P \left( -\gamma_N \hat{T}_2(\xi_p + \varepsilon_N) > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{6} \right).
\]

Furthermore, according to (A1),
\[
(A3) \quad P \left( -\gamma_N \hat{S}(\xi_p + \varepsilon_N) > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{6} \right) = O(N^{-1}) , \quad N \to \infty ,
\]
since for large \( N \),
\[ G_N(\xi_p + \varepsilon_N) - p_N \geq \frac{1}{2} G'(\xi_p) \varepsilon_N . \]

Also, using Hoeffding's inequality (Hoeffding [8]), the standard argument establishes that
\[
(A4) \quad P \left( -\gamma_N \hat{T}_1(\xi_p + \varepsilon_N) > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{6} \right) \\
+ P \left( -\gamma_N \hat{T}_2(\xi_p + \varepsilon_N) > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{6} \right) = O(N^{-1}) , \quad N \to \infty .
\]

We now treat the second term on the right-hand side of (A2). Define
\[ L^*(t) = \left( \frac{n}{2} \right)^{-1} \sum_{1 \leq i < j \leq n} I(W_i + W_j \leq 2t) , \]
then there exists a constant \( c > 0 \) such that for large \( N \),
\[
(A5) \quad P \left( -\delta_N [\hat{L}(\xi_p + \varepsilon_N) - L(\xi_p + \varepsilon_N)] > \frac{G_N(\xi_p + \varepsilon_N) - p_N}{2} \right) \\
\leq P \left( -L^*(\xi_p + \varepsilon_N) + L(\xi_p + \varepsilon_N) > \frac{G_N(\xi_p + \varepsilon_N) - p_N - cN^{-1}}{2\delta_N} \right) .
\]

Applying Hoeffding's inequality for \( U \)-statistics (Hoeffding [8]) we can show that the probability on the right-hand side of (A5) is also bound-
ed by $O(N^{-1})$ if $c$ is properly chosen.

(A2), (A3), (A4) and (A5) imply that

$$P(\hat{\xi}_{p_N} > \xi_p + \varepsilon_N) = O(N^{-1}), \quad N \to \infty.$$  

A similar argument also shows that

$$P(\hat{\xi}_{p_N} < \xi_p - \varepsilon_N) = O(N^{-1}), \quad N \to \infty.$$  

Thus the assertion of the lemma follows immediately by applying the Borel-Cantelli lemma.

**Lemma A2.** Let $0 < p < 1$. If both $K'(t)$ and $L'(t)$ are bounded in a neighborhood of $\xi_p$ and $F_1$ satisfies a Lipschitz condition of order 1, then with probability 1,

$$\sup_{|t| \leq cN^{-1/2}(\log N)^{1/4}} |[\hat{G}(t + \xi_p) - \hat{G}(\xi_p)] - [G_N(t + \xi_p) - G_N(\xi_p)]| = O(N^{-1/2}(\log N)^{1/2}), \quad N \to \infty,$$

where $c$ is a positive constant.

**Proof.** Write

$$A_N = \sup_{|t| \leq cN^{-1/2}(\log N)^{1/4}} |[\hat{G}(t + \xi_p) - \hat{G}(\xi_p)] - [G_N(t + \xi_p) - G_N(\xi_p)]|$$

$$\leq \gamma_N \sup_{|t| \leq cN^{-1/2}(\log N)^{1/4}} |[\hat{K}(t + \xi_p) - \hat{K}(\xi_p)] - [K(t + \xi_p) - K(\xi_p)]| + \delta_N \sup_{|t| \leq cN^{-1/2}(\log N)^{1/4}} |[\hat{L}(t + \xi_p) - \hat{L}(\xi_p)] - [L(t + \xi_p) - L(\xi_p)]|. $$

Define

$$B_{2N}(t) = [\hat{K}(t + \xi_p) - \hat{K}(\xi_p)] - [K(t + \xi_p) - K(\xi_p)],$$

$$B_{2N}(t) = [\hat{L}(t + \xi_p) - \hat{L}(\xi_p)] - [L(t + \xi_p) - L(\xi_p)].$$

Choose $d_N$ as a sequence of positive integers such that $d_N \sim cN^{1/4}(\log N)^{1/2}$ and put $\eta_{r,N} = r(d_N/d_N)$, where $c_N = cN^{-1/2}(\log N)^{1/2}$ and $-d_N \leq r \leq d_N$. Define $Q_{r,N} = [\eta_{r,N}, \eta_{r-1,N}]$, then for all $t \in Q_{r,N}$,

$$B_{2N}(t) \leq B_{2N}(\eta_{r+1,N}) + \alpha_{r,N}, \quad \text{and} \quad B_{1N}(t) \leq B_{1N}(\eta_{r+1,N}) + \alpha_{r,N},$$

where $\alpha_{r,N} = K(\eta_{r+1,N}) - K(\eta_{r,N})$ and $\alpha'_{r,N} = L(\eta_{r+1,N}) - L(\eta_{r,N})$. Similarly,

$$B_{1N}(t) \geq B_{1N}(\eta_{r,N}) - \alpha_{r,N}, \quad \text{and} \quad B_{2N}(t) \geq B_{2N}(\eta_{r,N}) - \alpha'_{r,N},$$

for all $t \in Q_{r,N}$.

Thus

$$A_N \leq \gamma_N \max \{|B_{1N}(\eta_{r,N})| : -d_N \leq r \leq d_N\}$$
\[ + \delta_N \max \{ |B_N(\tau, N)| : -d_N \leq \tau \leq d_N \} \]
\[ + \gamma_N \max \{ a_{r,N} : -d_N \leq r \leq d_N - 1 \} \]
\[ + \delta_N \max \{ a'_{r,N} : -d_N \leq r \leq d_N - 1 \} . \]

Now, according to the conditions of Lemma A2, it is easy to see that
\[ \gamma_N \max \{ a_{r,N} : -d_N \leq r \leq d_N - 1 \} + \delta_N \max \{ a'_{r,N} : -d_N \leq r \leq d_N - 1 \} \]
\[ = O(N^{-3/4}) , \quad N \to \infty . \]

Define \( a_N = N^{-3/4}(\log N)^{3/4} \), then using the condition that \( F_t \in \text{Lip}(1) \) and Bernstein's inequality and (A1), for all \( r = -d_N, \ldots, d_N \), we have
\[ P \left( |B_N(\tau, N)| > a_N \right) \]
\[ \leq P \left( |\hat{S}(\tau, N)| > a_N / 6 \right) + P \left( |\hat{S}(\xi_p)| > a_N / 6 \right) + P \left( |\hat{T}_1(\tau, N) - \hat{T}_1(\xi_p)| > a_N / 3 \right) \]
\[ + P \left( |\hat{T}_2(\tau, N) - \hat{T}_2(\xi_p)| > a_N / 3 \right) = O(N^{-1}) , \quad N \to \infty . \]

This implies that with probability 1,
\[ \gamma_N \max \{ |B_N(\tau, N)| : -d_N \leq r \leq d_N \} = O(N^{-3/4}(\log N)^{3/4}) , \quad N \to \infty . \]

On the other hand,
\[ \delta_N \max \{ |B_N(\tau, N)| : -d_N \leq \tau \leq d_N \} \]
\[ = \delta_N \max \{ |B'_N(\tau, N)| : -d_N \leq \tau \leq d_N \} + O(N^{-1}) , \]
where
\[ B'_N(t) = [L^*(t + \xi_p) - L^*(\xi_p)] - [L(t + \xi_p) - L(\xi_p)] \]
and \( L^*(t) \) is defined in Lemma A1. Also, for a suitable choice of \( c_i > 0 \) and for all \( r = -d_N, \ldots, d_N \),
\[ P \left( |B'_N(\tau, N)| > c_i N^{-3/4}(\log N)^{3/4} \right) = O(N^{-1}) , \quad N \to \infty , \]
by applying a method similar to that used by Geertsema ([4], Lemma 4.2). As a consequence, with probability 1,
\[ \delta_N \max \{ |B_N(\tau, N)| : -d_N \leq \tau \leq d_N \} = O(N^{-3/4}(\log N)^{3/4}) , \quad N \to \infty . \]

This and the above results establish the proof of the lemma.

**Proof of Theorem 2.1.** First, since both \( K \) and \( L \) are twice differentiable at \( \xi_p \), thus with probability 1,
\[ G_N(\hat{\xi}_p) - G_N(\xi_p) = [\gamma_N K'(\xi_p) + \delta_N L'(\xi_p)] \]
\[ \cdot (\hat{\xi}_p - \xi_p) + O(N^{-1}(\log N)) , \quad N \to \infty , \]
by Lemma A1. Applying Lemma A2, we readily have, with probability 1,
\[ \hat{G}(\hat{\xi}_{2},N) - G(\hat{\xi}_{2}) = \left[ \gamma_{N}K'(\xi_{2}) + \delta_{N}L'(\xi_{2}) \right] \cdot (\hat{\xi}_{2} - \xi_{2}) + O(N^{-\frac{1}{4}}(\log N)^{\frac{1}{4}}) , \quad N \to \infty , \]
and consequently,
\[ \hat{\xi}_{2} - \xi_{2} = \frac{\hat{G}(\hat{\xi}_{2},N) - G(\hat{\xi}_{2})}{\gamma_{N}K'(\xi_{2}) + \delta_{N}L'(\xi_{2})} + O(N^{-\frac{1}{4}}(\log N)^{\frac{1}{4}}) , \quad N \to \infty . \]

This completes the proof.

**Proof of Theorem 2.2.** Since \( \mathcal{F} \) is nondecreasing, therefore
\[ P(\lambda_{\delta}(t) < 0) \leq P(\hat{\theta} \leq t) \leq P(\lambda_{\delta}(t) \leq 0) . \]
Define \( t_{z,N} = \theta + z\sigma_{1}N^{-\frac{1}{4}} \) and let \( \Phi \) denote the standard normal distribution, then it is enough to show that
\[ \lim_{N \to \infty} P(\lambda_{\delta}(t_{z,N}) \leq 0) = \Phi(z) , \quad \text{for each} \quad z \in \mathbb{R} . \]
Recall that \( M = a(n+m)(n+l) + \beta n(n+1)/2 \) and write
\[ P(\lambda_{\delta}(t_{z,N}) \leq 0) = P(M^{-1}[\mathcal{F}(Y_{j} - X_{i} - t_{z,N}) - \lambda_{K}(t_{z,N})] \]
\[ + \beta \sum_{1 \leq i \leq j \leq n} [\mathcal{F}(W_{i} + W_{j} - t_{z,N}) - \lambda_{L}(t_{z,N})] \leq z) \to \Phi(z) , \quad N \to \infty , \quad \text{for all} \quad z \in \mathbb{R} . \]

Define
\[ \hat{V} = \gamma_{N}(n+m)^{-1} \sum_{i=1}^{n+m} [\mathcal{F}(X_{i} + t_{z,N}) - \lambda_{K}(t_{z,N})] \]
\[ + (n+l)^{-1} \sum_{j=1}^{n+l} [\mathcal{F}(Y_{j} - t_{z,N}) - \lambda_{K}(t_{z,N})] \]
\[ - 2\delta_{N} \sum_{i=1}^{n} [\mathcal{F}((W_{i}/2) - t_{z,N}) - \lambda_{L}(t_{z,N})] . \]

Then it can be shown that
\[ N^{\frac{1}{2}}[\lambda_{\delta}(t_{z,N}) - (\gamma_{N}\lambda_{K}(t_{z,N}) + \delta_{N}\lambda_{L}(t_{z,N})) - \hat{V}] \to 0 , \quad N \to \infty , \]
by utilizing the condition that \( \mathcal{F} \) is a bounded function and results similar to (A1) and the properties of the projection of a \( U \)-statistic. Finally applying the Lindeberg-Feller theorem for double arrays of random variables, the first assertion follows easily.

To establish a proof for the second assertion, we first note that
\[ \lambda_{\delta}(\hat{\theta}) - \lambda_{\delta}(\theta) = (\hat{\theta} - \theta)M^{-1} \left[ \alpha \sum_{i=1}^{n+m} \sum_{j=1}^{n+l} \frac{\partial \mathcal{F}(Y_{j} - X_{i} - t)}{\partial t} \right]_{t=\hat{\theta}} \]

\[ \text{(A6)} \quad \lambda_{\delta}(\hat{\theta}) - \lambda_{\delta}(\theta) = (\hat{\theta} - \theta)M^{-1} \left[ \alpha \sum_{i=1}^{n+m} \sum_{j=1}^{n+l} \frac{\partial \mathcal{F}(Y_{j} - X_{i} - t)}{\partial t} \right]_{t=\hat{\theta}} \]
\[ + \beta \sum_{i \leq i \leq j \leq n} \frac{\partial \mathcal{Y}(W_i + W_j/2 - t)}{\partial t} \bigg|_{t=\bar{\theta}} \]

where \(|\bar{\theta} - \theta| \leq |\hat{\theta} - \theta|\). Hence we have

\[ N^{1/2}(\bar{\theta} - \theta) = -N^{1/2} \hat{\lambda}(\theta) \frac{M^{-1} \left[ \alpha \sum_{t=1}^{n+m} \sum_{j=1}^{n+l} \frac{\partial \mathcal{Y}(Y_j - X_i - t)}{\partial t} \bigg|_{t=\hat{\theta}} + \beta \sum_{i \leq i \leq j \leq n} \frac{\partial \mathcal{Y}(W_i + W_j/2 - t)}{\partial t} \bigg|_{t=\hat{\theta}} \right]}{M^{-1} \left[ \alpha \sum_{t=1}^{n+m} \sum_{j=1}^{n+l} \frac{\partial \mathcal{Y}(Y_j - X_i - t)}{\partial t} \bigg|_{t=\hat{\theta}} + \beta \sum_{i \leq i \leq j \leq n} \frac{\partial \mathcal{Y}(W_i + W_j/2 - t)}{\partial t} \bigg|_{t=\hat{\theta}} \right]} . \]

From our previous argument, we have seen that

\[-N^{1/2} \hat{\lambda}(\theta) \xrightarrow{d} N(0, \sigma^2).\]

Also, using the projection technique in connection with the property of \(\hat{\theta}\),

\[ M^{-1} \left[ \alpha \sum_{t=1}^{n+m} \sum_{j=1}^{n+l} \frac{\partial \mathcal{Y}(Y_j - X_i - t)}{\partial t} \bigg|_{t=\hat{\theta}} + \beta \sum_{i \leq i \leq j \leq n} \frac{\partial \mathcal{Y}(W_i + W_j/2 - t)}{\partial t} \bigg|_{t=\hat{\theta}} \right] \xrightarrow{p} \gamma \int_{-\infty}^{\infty} \frac{\partial \mathcal{Y}(x - t)}{\partial t} \bigg|_{t=\hat{\theta}} dK(x) + \gamma \int_{-\infty}^{\infty} \frac{\partial \mathcal{Y}(x - t)}{\partial t} \bigg|_{t=\hat{\theta}} dL(x) . \]

This finishes the proof.

**Proof of Theorem 2.3.** The proof is similar to that given in the proof of Theorem 2.1. The key steps of the proof are to show

(A7) \[ r_{\hat{\theta}}^{-1}(p_N) - \theta = O(N^{-1/2}(\log N)^{1/2}) , \quad N \to \infty , \]

with probability 1 and for constant \(c > 0\),

(A8) \[ \sup_{|t| \leq c N^{-1/4}(\log N)^{1/2}} |(r_{\hat{\theta}}(t + \theta) - r_{\hat{\theta}}(\theta)) - (r_{\hat{\theta}}(t + \theta) - r_{\hat{\theta}}(\theta))| = O(r_N - r) + O(\delta_N - \delta) + O(N^{-3/4}(\log N)^{3/4}) , \quad N \to \infty , \]

with probability 1. Using (A7), (A8), conditions on \(r_{\hat{\theta}}(d)\) and the continuity of \(\mathcal{Y}\), the assertion of the theorem follows.