Ann. Inst. Statist. Math.
Vol. 49, No. 1, 35–55 (1997)

QUANTILE PROCESSES IN THE PRESENCE OF AUXILIARY INFORMATION

BIAO ZHANG

Department of Mathematics, The University of Toledo, Toledo, OH 43606, U.S.A.

(Received March 22, 1995; revised April 24, 1996)

Abstract. We employ the empirical likelihood method to propose a modified quantile process under a nonparametric model in which we have some auxiliary information about the population distribution. Furthermore, we propose a modified bootstrap method for estimating the sampling distribution of the modified quantile process. To explore the asymptotic behavior of the modified quantile process and to justify the bootstrapping of this process, we establish the weak convergence of the modified quantile process to a Gaussian process and the almost-sure weak convergence of the modified bootstrapped quantile process to the same Gaussian process. These results are demonstrated to be applicable, in the presence of auxiliary information, to the construction of asymptotic bootstrap confidence bands for the quantile function. Moreover, we consider estimating the population semi-interquartile range on the basis of the modified quantile process. Results from a simulation study assessing the finite-sample performance of the proposed semi-interquartile range estimator are included.

Key words and phrases: Bootstrap, Brownian bridge, confidence band, empirical likelihood, empirical process, Gaussian process, semi-interquartile range, weak convergence.

1. Introduction

Let $X_1, \dots, X_n$ be independent and identically distributed random variables with unknown distribution function F confined to the interval $[a, b]$, where $-\infty \leq a < b \leq \infty$. The quantile function associated with F is the function defined by $F^{-1}(s) = \inf\{t; F(t) \geq s\}$ for $s \in (0, 1)$. Under the nonparametric setting, the standard estimator of the quantile function $F^{-1}(s)$ is the empirical quantile function $F_n^{-1}(s)$, where F_n is the empirical distribution function of $X_1, \dots, X_n$. The process $Q_n = \sqrt{n}(F_n^{-1} - F^{-1})$ is referred to as the quantile process. Our focus of attention in this paper is to study quantile processes and their corresponding bootstrap versions under the following model. For the unknown population distribution function F underlying the random sample $X_1, \dots, X_n$, we assume that we have some auxiliary information about F in the sense that there exist r ($r \geq 1$)

functionally independent functions $g_1(x), \ldots, g_r(x)$ such that

$$\mathrm{E}_F g(X) = 0, \tag{1.1}$$

where $g(x) = (g_1(x), \ldots, g_r(x))^\tau$. This model is of interest in that in many situations, we may have some partial information about the distribution, although we do not know exactly the distribution function underlying the sample. For example, one may know the population mean, or a population quantile, or that the population distribution is symmetric about a known constant. Several authors have considered nonparametric estimation of a functional of F when auxiliary information about F is available. For instance, Haberman (1984) considered the problem of estimating a Euclidean-valued functional of a probability measure satisfying a finite number of linear constraints. By extending Haberman's work, Sheehy (1988) considered distributions that minimize a distance measure from the empirical distribution subject to linear constraints. Both Haberman and Sheehy's works are based on the minimization of the Kullback-Leibler divergence from the constrained family of probability measures to the empirical measure of the observations. Under simple random sampling, Kuk and Mak (1989) considered estimating the median of a finite population in the presence of auxiliary information. In the context of linking the empirical likelihood method and auxiliary information available in a population, Owen (1991) has shown that when some functionals of the distribution of the data are known, one can obtain sharper inferences on other functionals by constrained empirical likelihood. Chen and Qin (1993) have indicated that the empirical likelihood method can be naturally applied to make more accurate statistical inference in finite population estimation problems by efficiently employing auxiliary information. Qin and Lawless (1994) have shown that empirical likelihood can be used to perform inference on parameters under a semiparametric model. Zhang (1995*a*) has considered the M-estimation problem as well as the quantile estimation problem in the presence of auxiliary information (1.1) in conjunction with the method of empirical likelihood. Zhang (1995*b*) has considered an alternative estimator $\hat{F}_n$ of F satisfying (1.1) and has established the weak convergence of the modified empirical process $\sqrt{n}(\hat{F}_n - F)$ to a Gaussian process. Recently, Zhang (1995*c*) has proposed a modified bootstrap procedure in the case where we have auxiliary information (1.1) and has established the asymptotic validity of the modified bootstrap procedure by proving the almost-sure weak convergence of the modified bootstrapped empirical process.

The empirical likelihood method for constructing confidence regions in nonparametric settings was introduced by Owen (1988, 1990). For a more complete survey of developments in empirical likelihood, see Hall and La Scala (1990) and Owen (1991).

There are two objectives in this paper. Our first objective is to employ the method of empirical likelihood to propose and study, in the presence of auxiliary information (1.1), a modified quantile process defined by $\hat{Q}_n = \sqrt{n}(\hat{F}_n^{-1} - F^{-1})$. To explore the asymptotic behavior of $\hat{Q}_n$, we establish, by representing $\hat{Q}_n$ as the mean of a sequence of independent and identically distributed stochastic processes with a remainder term of order $o_p(n^{-1/2})$, the weak convergence of $\hat{Q}_n$ to a Gaussian process on a proper subinterval of $[0, 1]$ not containing the endpoints.

One advantage of the modified quantile process $\hat{Q}_n$ over the standard quantile process Q_n is that the asymptotic variance function of $\hat{F}_n^{-1}$ is uniformly smaller than that of F_n^{-1}, which enables us to get sharpen inferences on the quantile function F^{-1}. Our second objective in this paper is to bootstrap the modified quantile process $\hat{Q}_n$. To justify the bootstrapping of the process $\hat{Q}_n$, we show that the limiting process of the modified bootstrapped quantile process coincides with that of the original process $\hat{Q}_n$. This result is used to construct asymptotically correct bootstrap confidence bands for F^{-1} with F satisfying (1.1). Other applications include estimating quantiles and the semi-interquartile range in the presence of auxiliary information (1.1).

This paper is organized as follows. In Section 2, we describe the profile empirical likelihood under (1.1) and propose a modified quantile process $\hat{Q}_n = \sqrt{n}(\hat{F}_n^{-1} - F^{-1})$ with F satisfying (1.1). In Section 3, we establish the weak convergence of $\hat{Q}_n$ to a Gaussian process. In Section 4, we bootstrap the modified quantile process $\hat{Q}_n$ and establish the almost-sure weak convergence of the modified bootstrapped quantile process. As an application of this result, we propose in Section 5 the bootstrap confidence bands for F^{-1} with F satisfying (1.1). Section 6 deals with estimating the semi-interquartile range under (1.1). Simulation results are presented in Section 7 to demonstrate the performance of our proposed semi-interquartile range estimator for small samples. Finally, proofs of lemmas and theorems appear in Section 8.

2. Profile empirical likelihood

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with distribution function F confined to the interval $[a, b]$, where $-\infty \leq a < b \leq \infty$. Let $p = (p_1, \ldots, p_n)$ denote a multinomial distribution on the points $X_1, \ldots, X_n$ and put $L(p) = \prod_{i=1}^n p_i$. Under (1.1), since $\mathrm{E}_F g(X) = 0$, the profile empirical likelihood L is defined to be $L = \max_p L(p) = \max_p \prod_{i=1}^n p_i$, subject to the restrictions $\sum_{i=1}^n p_i = 1$, $\sum_{i=1}^n p_i g(X_i) = 0$, and $p_i \geq 0$ for $i = 1, \ldots, n$. If 0 is inside the convex hull of the points $g(X_1), \ldots, g(X_n)$, then L exists uniquely. A little calculus of variations shows that $L = \prod_{i=1}^n \hat{p}_i$, where $\hat{p}_i = \frac{1}{n} \frac{1}{1+\eta^\tau g(X_i)}$ for $1 \leq i \leq n$ with $\eta = (\eta_1, \ldots, \eta_r)^\tau$ being the solution of

$$\sum_{i=1}^n \hat{p}_i g(X_i) = \frac{1}{n} \sum_{i=1}^n \frac{1}{1 + \eta^\tau g(X_i)} g(X_i) = 0. \tag{2.1}$$

Now let

$$\hat{F}_n(x) = \sum_{i=1}^n \hat{p}_i I_{[X_i \leq x]} = \frac{1}{n} \sum_{i=1}^n \frac{I_{[X_i \leq x]}}{1 + \eta^\tau g(X_i)}, \tag{2.2}$$

then $\hat{F}_n$ can be regarded as an alternative estimator for F satisfying (1.1). On the basis of $\hat{F}_n$, we propose, in the presence of auxiliary information (1.1), to estimate the quantile function F^{-1} by $\hat{F}_n^{-1}$, which will be called the modified empirical quantile function. Note that in the absence of (1.1), since $\hat{F}_n$ reduces to the standard empirical distribution function F_n, $\hat{F}_n^{-1}$ reduces to the standard empirical quantile function F_n^{-1}. Throughout this paper, we refer to the process $\hat{Q}_n = \sqrt{n}(\hat{F}_n^{-1} - F^{-1})$ as the modified quantile process.

Zhang (1995*b*) has established the weak convergence of the modified empirical process $\sqrt{n}(\hat{F}_n - F)$ to a Gaussian process in $D[a,b]$ by expressing $\hat{F}_n$ as the mean of a sequence of independent and identically distributed stochastic processes with a remainder term of order $o_p(n^{-1/2})$. These results are summarized in the following theorem.

THEOREM 2.1. *If* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite, then we can write* $\hat{F}_n(x) - F(x) = \frac{1}{n}\sum_{i=1}^n Y_i(x) + R_n(x)$, *where* $Y_i(x) = I_{[X_i \le x]} - F(x) - U^\tau(x)\Sigma^{-1}g(X_i)$, $U(x) = \mathrm{E}[g(X)I_{[X\le x]}]$, *and the remainder term* $R_n(x)$ *satisfies* $\sup_{a\le x\le b}|R_n(x)| = o_p(n^{-1/2})$. *As a result,* $\sqrt{n}(\hat{F}_n - F)\xrightarrow{\mathcal{D}} W$ *on* $D[a,b]$, *where* W *is a Gaussian process with continuous sample paths and satisfies*

$$\begin{aligned} &\mathrm{E}W(x) = 0, \qquad a \le x \le b,\\ &\mathrm{E}W(x)W(y) = F(\min(x,y)) - F(x)F(y)\\ &\qquad\qquad - \mathrm{E}[g^\tau(X)I_{[X\le x]}]\Sigma^{-1}\mathrm{E}[g(X)I_{[X\le y]}], \qquad a \le x, y \le b. \end{aligned}$$

In this paper, the norm of a $n_1 \times n_2$ matrix $A = (a_{ij})_{n_1\times n_2}$ is defined by $\|A\| = (\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} a_{ij}^2)^{1/2}$ for $n_1, n_2 \ge 1$. Moreover, we use o_p^* and O_p^* to stand for o_p and O_p in bootstrap probability under $\hat{F}_n$.

We close this section by establishing the uniform strong consistency of the modified empirical quantile function $\hat{F}_n^{-1}$ over a subinterval of $[0,1]$ not containing the endpoints. This result is a part of Lemma 8.6 in Section 8.

THEOREM 2.2. *Let* $0 < \alpha < \beta < 1$ *be given. Suppose that* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite and* $\mathrm{E}\|g(X)\|^q < \infty$ *with some* $q > 2$. *Suppose further that* F *has continuous positive density* f *on* $[F^{-1}(\alpha), F^{-1}(\beta)]$. *Then, as* $n \to \infty$,

$$\sup_{\alpha\le s\le\beta} |\hat{F}_n^{-1}(s) - F^{-1}(s)| = o(n^{-1/q}) \qquad a.s.$$

3. Weak convergence of $\hat{Q}_n$

We begin with expressing $\hat{F}_n^{-1}$ as the mean of a sequence of independent and identically distributed stochastic processes with a remainder term of order $o_p(n^{-1/2})$ on a subinterval of $[0,1]$.

THEOREM 3.1. *Let* $0 < \alpha < \beta < 1$ *be given. Suppose that* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite and* F *has continuous positive density* f *on* $[F^{-1}(\alpha) - \epsilon, F^{-1}(\beta) + \epsilon]$ *for some* $\epsilon > 0$. *Then, one can write*

$$\hat{F}_n^{-1}(s) - F^{-1}(s) = \frac{1}{n}\sum_{i=1}^n H_i(s) + r_n(s), \tag{3.1}$$

where

$$H_i(s) = -\frac{Y_i(F^{-1}(s))}{f(F^{-1}(s))} = -\frac{I_{[X_i \le F^{-1}(s)]} - s - \mathrm{E}[g^\tau(X_i) I_{[X_i \le F^{-1}(s)]}]\Sigma^{-1} g(X_i)}{f(F^{-1}(s))}$$

and the remainder term $r_n(s)$ *satisfies*

$$\sup_{\alpha \le s \le \beta} |r_n(s)| = o_p(n^{-1/2}). \tag{3.2}$$

Theorem 3.1 enables us to establish the weak convergence of the modified quantile process $\hat{Q}_n$ to a Gaussian process in $D[\alpha, \beta]$.

THEOREM 3.2. *Under the same conditions as in Theorem* 3.1, *as* $n \to \infty$,

$$\sqrt{n}(\hat{F}_n^{-1} - F^{-1}) \xrightarrow{\mathcal{D}} W(F^{-1})/f(F^{-1}) \quad on \quad D[\alpha, \beta],$$

where W *is the Gaussian process defined in Theorem* 2.1.

Remark 3.1. In the absence of auxiliary information (1.1), Theorem 3.2 reduces to $Q_n \xrightarrow{\mathcal{D}} B/f(F^{-1})$ in $D[\alpha, \beta]$ (Bickel (1966)), where $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge.

Remark 3.2. For fixed $p \in (0,1)$, if $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and F has a continuous positive density f in a neighborhood of $F^{-1}(p)$, then Theorem 3.2 implies that as $n \to \infty$, $\sqrt{n}(\hat{F}_n^{-1}(p) - F^{-1}(p)) \xrightarrow{\mathcal{D}} N(0, \sigma_p^2)$, where $\sigma_p^2 = \{p(1-p) - [\int_a^{F^{-1}(p)} g^\tau(x) dF(x)][\mathrm{E}g(X)g^\tau(X)]^{-1}[\int_a^{F^{-1}(p)} g(x) dF(x)]\}/f^2(F^{-1}(p))$. This result was also established by Zhang (1995*a*).

Remark 3.3. For $s \in [\alpha, \beta]$, let

$$\sigma_0^2(s) = \frac{s(1-s)}{[f(F^{-1}(s))]^2},$$

$$\sigma_r^2(s) = \frac{s(1-s) - \mathrm{E}[g^\tau(X_1) I_{[X_1 \le F^{-1}(s)]}]\Sigma^{-1}\mathrm{E}[g(X_1) I_{[X_1 \le F^{-1}(s)]}]}{[f(F^{-1}(s))]^2}, \qquad r \ge 1.$$

Theorem 3.2 indicates that the asymptotic variance function $\sigma_r^2(s)$ of $\hat{F}_n^{-1}(s)$ is uniformly smaller than the asymptotic variance function $\sigma_0^2(s)$ of $F_n^{-1}(s)$ for all $s \in [\alpha, \beta]$. Furthermore, it can be shown that $\sigma_r^2(s) \le \sigma_{r-1}^2(s)$ for $r \ge 1$ and all $s \in [\alpha, \beta]$, which indicates that the more information we have, the smaller asymptotic variance function of $\hat{F}_n^{-1}$ would have.

4. Bootstrapping the modified quantile process $\hat{Q}_n$

On the basis of the modified empirical distribution function $\hat{F}_n$, Zhang (1995c) proposed a modified bootstrap procedure in which we generate bootstrap samples from $\hat{F}_n$, instead of from the standard empirical distribution function F_n. This modified bootstrap procedure is employed in this section to bootstrap the modified quantile process $\hat{Q}_n$. Specifically, let $(X_1^*, \ldots, X_n^*)$ be a random sample of size n from $\hat{F}_n$, then the auxiliary information (1.1) about F carries out to $\hat{F}_n$ in the sense that $\mathrm{E}_{\hat{F}_n} g(X_1^*) = \sum_{i=1}^n \hat{p}_i g(X_i) = 0$ by (2.1). In analogy with the approach in which we propose $\hat{F}_n$ in (2.2), we propose the following modified bootstrap empirical distribution function $\hat{G}_n$ based on $X_1^*, \ldots, X_n^*$:

$$\hat{G}_n(x) = \sum_{j=1}^n \hat{p}_j^* I_{[X_j^* \le x]} = \frac{1}{n} \sum_{j=1}^n \frac{I_{[X_j^* \le x]}}{1 + \lambda^\tau g(X_j^*)}, \tag{4.1}$$

where $\hat{p}_j^* = \frac{1}{n} \frac{1}{1+\lambda^\tau g(X_j^*)}$ for $1 \le j \le n$ with $\lambda = (\lambda_1, \ldots, \lambda_r)^\tau$ being the solution of

$$\frac{1}{n} \sum_{j=1}^n \frac{1}{1 + \lambda^\tau g(X_j^*)} g(X_j^*) = 0. \tag{4.2}$$

Based on $\hat{G}_n$, we define the modified bootstrap quantile process by $\hat{Q}_n^* = \sqrt{n}(\hat{G}_n^{-1} - \hat{F}_n^{-1})$. Note that in the absence of auxiliary information (1.1), $\hat{Q}_n^*$ reduces to the standard bootstrapped quantile process $Q_n^* = \sqrt{n}(G_n^{-1} - F_n^{-1})$, where $G_n(x) = \frac{1}{n} \sum_{j=1}^n I_{[X_j^* \le x]}$ with $X_j^* \sim F_n$ (Bickel and Freedman (1981)).

Zhang (1995c) has established the almost-sure weak convergence of the modified bootstrapped empirical process $\sqrt{n}(\hat{G}_n - \hat{F}_n)$ to a Gaussian process in $D[a,b]$ by expressing $\hat{G}_n$ as the mean of a sequence of independent and identically distributed stochastic processes with a remainder term of order $o_p^*(n^{-1/2})$. These results are summarized in the following theorem.

THEOREM 4.1. *If $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and $\mathrm{E}\|g(X)\|^q < \infty$ with some $q > 2$, then along almost all sample sequences $X_1, X_2, \ldots,$ given $(X_1, \ldots, X_n)$, as $n \to \infty$, we can write $\hat{G}_n(x) - \hat{F}_n(x) = \frac{1}{n} \sum_{j=1}^n Y_j^*(x) + R_n^*(x)$, where $Y_j^*(x) = I_{[X_j^* \le x]} - \hat{F}_n(x) - U_n^\tau(x) \Sigma_n^{-1} g(X_j^*)$, $U_n(x) = \mathrm{E}_{\hat{F}_n}[g(X_1^*) I_{[X_1^* \le x]}]$, $\Sigma_n = \mathrm{E}_{\hat{F}_n}[g(X_1^*) g^\tau(X_1^*)]$, and the remainder term $R_n^*(x)$ satisfies $\sup_{a \le x \le b} |R_n^*(x)| = o_p^*(n^{-1/2})$ almost surely. As a result, along almost all sample sequences $X_1, X_2, \ldots,$ given $(X_1, \ldots, X_n)$, as $n \to \infty$, $\sqrt{n}(\hat{G}_n - \hat{F}_n) \xrightarrow{\mathcal{D}} W$ on $D[a,b]$, where W is the Gaussian process defined in Theorem* 2.1.

On the basis of Theorem 4.1, we can explore the asymptotic behavior of the modified bootstrapped quantile process $\hat{Q}_n^*$ by deriving the bootstrap versions of the representation and the weak convergence as given in Theorems 3.1 and 3.2. We state our results in two theorems.

THEOREM 4.2. *Let $0 < \alpha < \beta < 1$ be given. Suppose that $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and $\mathrm{E}\|g(X)\|^3 < \infty$. Suppose further that F has continuous positive density f on $[F^{-1}(\alpha) - \epsilon, F^{-1}(\beta) + \epsilon]$ for some $\epsilon > 0$. Then, along almost all sample sequences $X_1, X_2, \ldots$, given $(X_1, \ldots, X_n)$, as $n \to \infty$, one can write*

$$\hat{G}_n^{-1}(s) - \hat{F}_n^{-1}(s) = \frac{1}{n}\sum_{i=1}^{n} H_i^*(s) + r_n^*(s), \tag{4.3}$$

where

$$\begin{aligned} H_i^*(s) &= -\frac{Y_i^*(F^{-1}(s))}{f(F^{-1}(s))} \\ &= -\frac{I_{[X_i^* \le F^{-1}(s)]} - \hat{F}_n(F^{-1}(s)) - \mathrm{E}_{\hat{G}_n}[g^\tau(X_i^*)I_{[X_i^* \le F^{-1}(s)]}]\Sigma_n^{-1}g(X_i^*)}{f(F^{-1}(s))}, \end{aligned}$$

and the remainder term $r_n^*(s)$ satisfies

$$\sup_{\alpha \le s \le \beta} |r_n^*(s)| = o_p^*(n^{-1/2}) \qquad a.s. \tag{4.4}$$

Theorem 4.2 enables us to establish the almost-sure weak convergence of the modified bootstrapped quantile process $\hat{Q}_n^*$. This result, together with Theorem 3.2, indicates that the limiting process of $\hat{Q}_n^* = \sqrt{n}(\hat{G}_n^{-1} - \hat{F}_n^{-1})$ agrees with that of $\hat{Q}_n = \sqrt{n}(\hat{F}_n^{-1} - F^{-1})$, thereby we can claim that $\hat{Q}_n^*$ has the same limiting behavior as $\hat{Q}_n$.

THEOREM 4.3. *Under the conditions of Theorem 4.2, along almost all sample sequences $X_1, X_2, \ldots$, given $(X_1, \ldots, X_n)$, as $n \to \infty$, $\sqrt{n}(\hat{G}_n^{-1} - \hat{F}_n^{-1}) \xrightarrow{\mathcal{D}} W(F^{-1})/f(F^{-1})$ on $D[\alpha, \beta]$, where W is the Gaussian process defined in Theorem 2.1.*

5. Confidence bands for F^{-1} in the presence of (1.1)

In this section, we demonstrate how the result of Theorem 4.3 can be employed to set confidence bands on F^{-1} with F satisfying (1.1). According to Theorems 3.2 and 4.3 and the continuous mapping theorem (Theorem 5.1, Billingsley (1968), p. 30), we immediately have the following result.

THEOREM 5.1. *Let $0 < \alpha < \beta < 1$ be given. Suppose that $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and F has continuous positive density f on $[F^{-1}(\alpha) - \epsilon, F^{-1}(\beta) + \epsilon]$ for some $\epsilon > 0$.*

(i) $\sup_{\alpha \le s \le \beta} \sqrt{n}|\hat{F}_n^{-1}(s) - F^{-1}(s)| \xrightarrow{\mathcal{D}} \sup_{\alpha \le s \le \beta} |\frac{W(F^{-1}(s))}{f(F^{-1}(s))}|$.

(ii) *If $\mathrm{E}\|g(X)\|^3 < \infty$, then along almost all sample sequences $X_1, X_2, \ldots$, given $(X_1, \ldots, X_n)$, as $n \to \infty$, $\sup_{\alpha \le s \le \beta} \sqrt{n}|\hat{G}_n^{-1}(s) - \hat{F}_n^{-1}(s)| \xrightarrow{\mathcal{D}} \sup_{\alpha \le s \le \beta} |\frac{W(F^{-1}(s))}{f(F^{-1}(s))}|$.*

Part (i) of Theorem 5.1 indicates that construction of confidence bands for F^{-1} requires of finding the exact distribution of $\sup_{\alpha\le s\le\beta}|\frac{W(F^{-1}(s))}{f(F^{-1}(s))}|$ and its quantiles. Unfortunately, no simple closed-form expressions for these quantiles are known. Nevertheless, an application of the bootstrap procedure described in Section 4 can overcome this difficulty. Indeed, Theorem 5.1 implies that the limiting distribution of $\sup_{\alpha\le s\le\beta}\sqrt{n}|\hat{G}_n^{-1}(s) - \hat{F}_n^{-1}(s)|$ agrees with that of $\sup_{\alpha\le s\le\beta}\sqrt{n}|\hat{F}_n^{-1}(s) - F^{-1}(s)|$ and thus we can approximate the quantiles of $\sup_{\alpha\le s\le\beta}\sqrt{n}|\hat{F}_n^{-1}(s) - F^{-1}(s)|$ by those of $\sup_{\alpha\le s\le\beta}\sqrt{n}|\hat{G}_n^{-1}(s) - \hat{F}_n^{-1}(s)|$. With this heuristic argument in mind, let $v_{1-\alpha}^n = \inf\{t; P^*(\sqrt{n}\hat{V}_n^* \le t) \ge 1-\alpha\}$ with $0<\alpha<1$, $\hat{V}_n^* = \sup_{\alpha\le s\le\beta}|\hat{G}_n^{-1}(s) - \hat{F}_n^{-1}(s)|$, and P^* stands for the bootstrap probability under $\hat{F}_n$. Then under the assumptions in Theorem 5.1 and the additional assumption that the distribution of $\sup_{\alpha\le s\le\beta}|\frac{W(F^{-1}(s))}{f(F^{-1}(s))}|$ is continuous, it can be shown from Theorem 5.1 that as $n\to\infty$, $\lim_{n\to\infty}P(\sup_{\alpha\le s\le\beta}\sqrt{n}|\hat{F}_n^{-1}(s) - F^{-1}(s)| \le v_{1-\alpha}^n) = 1-\alpha$. As a result, a $(1-\alpha)$-level bootstrap confidence band for F^{-1} on $[\alpha,\beta]$ with F satisfying (1.1) is given by $\hat{\mathcal{J}}^* = (\hat{F}_n^{-1}(\cdot) - \frac{v_{1-\alpha}^n}{\sqrt{n}}, \hat{F}_n^{-1}(\cdot) + \frac{v_{1-\alpha}^n}{\sqrt{n}})$. Note that according to Theorem 1 of Tsirel'son (1975), the distribution of $\sup_{\alpha\le s\le\beta}|\frac{W(F^{-1}(s))}{f(F^{-1}(s))}|$ is continuous except perhaps at the lower endpoint of its support. In the absence of (1.1), the bootstrap confidence band $\hat{\mathcal{J}}^*$ reduces to the standard $(1-\alpha)$-level bootstrap confidence band for F^{-1} on $[\alpha,\beta]$ based on F_n^{-1}: $\mathcal{J}^* = (F_n^{-1}(\cdot) - \frac{k_{1-\alpha}^n}{\sqrt{n}}, F_n^{-1}(\cdot) + \frac{k_{1-\alpha}^n}{\sqrt{n}})$, where $k_{1-\alpha}^n = \inf\{t; P^*(\sqrt{n}V_n^* \le t) \ge 1-\alpha\}$ with $V_n^* = \sup_{\alpha\le s\le\beta}|G_n^{-1}(s) - F_n^{-1}(s)|$.

Remark 3.3 predicts, for the same confidence level $1-\alpha$, a certain amount of narrowing in the proposed bootstrap confidence band $\hat{\mathcal{J}}^*$ as compared to the standard bootstrap confidence band $\mathcal{J}^*$.

6. Estimating the semi-interquartile range in the presence of auxiliary information

The population semi-interquartile range $R = \frac{1}{2}(F^{-1}(0.75) - F^{-1}(0.25))$ is an alternative to the population standard deviation σ of F as a measure of dispersion. Under the nonparametric setting, the standard estimator of R is the sample analogue $R_n = \frac{1}{2}(F_n^{-1}(0.75) - F_n^{-1}(0.25))$. However, in the presence of auxiliary information (1.1), we propose to estimate R by $\hat{R}_n = \frac{1}{2}(\hat{F}_n^{-1}(0.75) - \hat{F}_n^{-1}(0.25))$. The bootstrap replication of $\hat{R}_n$ is $\hat{R}_n^* = \frac{1}{2}(\hat{G}_n^{-1}(0.75) - \hat{G}_n^{-1}(0.25))$. To study the asymptotic behavior of $\hat{R}_n$ and $\hat{R}_n^*$, we first present the following theorem which is a straightforward consequence of Theorems 3.2 and 4.3.

THEOREM 6.1. *Let $0 < q_1 < \cdots < q_k < 1$. Suppose that $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and F has continuous positive density f in neighborhoods of $F^{-1}(q_1),\ldots,F^{-1}(q_k)$.*

(i) *As $n\to\infty$, $\sqrt{n}(\hat{F}_n^{-1}(q_1) - F^{-1}(q_1),\ldots,\hat{F}_n^{-1}(q_k) - F^{-1}(q_k)) \xrightarrow{\mathcal{D}} N_k(\mathbf{0},\boldsymbol{B})$, where $\mathbf{0} = (0,\ldots,0)^\tau$ and $\boldsymbol{B} = (b_{ij})_{k\times k}$ with $b_{ij} = \{q_i(1-q_j) - \mathrm{E}[g^\tau(X)I_{[X\le F^{-1}(q_i)]}]\Sigma^{-1}\mathrm{E}[g(X)I_{[X\le F^{-1}(q_j)]}]\}/\{f(F^{-1}(q_i))\cdot f(F^{-1}(q_j))\}$ for $i\le j$ and $b_{ij} = b_{ji}$ for $i > j$.*

(ii) *If* $\mathrm{E}\|g(X)\|^3 < \infty$, *then along almost all sample sequences* $X_1, X_2, \ldots$, *given* $(X_1, \ldots, X_n)$, *as* $n \to \infty$, $\sqrt{n}(\hat{G}_n^{-1}(q_1) - \hat{F}_n^{-1}(q_1), \ldots, \hat{G}_n^{-1}(q_k) - \hat{F}_n^{-1}(q_k)) \xrightarrow{\mathcal{D}} N_k(\mathbf{0}, \boldsymbol{B})$.

Theorem 6.1 and the Cramér-Wold device enables us to establish the asymptotic distribution of $\hat{R}_n$ and its bootstrap counterpart, as given by the following theorem.

THEOREM 6.2. *Suppose that* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite and* F *has continuous positive density* f *in neighborhoods of* $F^{-1}(0.25)$ *and* $F^{-1}(0.75)$.

(i) *As* $n \to \infty$, $\sqrt{n}(\hat{R}_n - R) \xrightarrow{\mathcal{D}} N(0, \sigma_{\hat{R}}^2)$, *where* $\sigma_{\hat{R}}^2 = \sigma_R^2 - (h(0.25) - h(0.75))^\tau \Sigma^{-1}(h(0.25) - h(0.75))$ *with* $\sigma_R^2 = \frac{3}{64}[f(F^{-1}(0.25))]^{-2} + \frac{3}{64}[f(F^{-1}(0.75))]^{-2} - \frac{1}{32}[f(F^{-1}(0.25))f(F^{-1}(0.75))]^{-1}$ *and* $h(y) = \{\mathrm{E}[g(X)I_{[X \leq F^{-1}(y)]}]\}/\{2f(F^{-1}(y))\}$.

(ii) *If* $\mathrm{E}\|g(X)\|^3 < \infty$, *then along almost all sample sequences* $X_1, X_2, \ldots$, *given* $(X_1, \ldots, X_n)$, *as* $n \to \infty$, $\sqrt{n}(\hat{R}_n^* - \hat{R}_n) \xrightarrow{\mathcal{D}} N(0, \sigma_{\hat{R}}^2)$.

Remark 6.1. In the absence of (1.1), $\hat{R}_n = R_n$ and Part (i) of Theorem 6.2 reduces to the well-known result about the asymptotic distribution of R_n (Serfling (1980), p. 86), that is, $\sqrt{n}(R_n - R) \xrightarrow{\mathcal{D}} N(0, \sigma_R^2)$. Likewise, Part (ii) of Theorem 6.2 reduces to the standard bootstrap asymptotic distribution of R_n^*, i.e., $\sqrt{n}(R_n^* - R_n) \xrightarrow{\mathcal{D}} N(0, \sigma_R^2)$ almost surely. Note that the asymptotic variance of $\hat{R}_n$ is less than that of R_n unless $h(0.25) = h(0.75)$. Also, note that the asymptotic relative efficiency of $\hat{R}_n$ relative to R_n is $e(R_n, \hat{R}_n) = \frac{\sigma_{\hat{R}}^2}{\sigma_R^2} = 1 - \{(h(0.25) - h(0.75))^\tau \Sigma^{-1}(h(0.25) - h(0.75))\}/\sigma_R^2$.

As an application of Theorem 6.2, let us consider estimating the semi-interquartile range R when (i) the population mean $\mu_0 = \mathrm{E}_F X$ is known or (ii) the population median $m_0 = F^{-1}(0.5)$ is known. Specifically, suppose $X_1, \ldots, X_n$ is a random sample from a population with unknown distribution function F. We assume that F has continuous positive density f in neighborhoods of $F^{-1}(0.25)$ and $F^{-1}(0.75)$. When μ_0 is known, we have $r = 1$ and $g(x) = g_1(x) = x - \mu_0$ in (1.1). When m_0 is known, we then have $g(x) = I_{[x \leq m_0]} - 0.5$. We now consider three special cases.

Example 6.1. Let $X_1, \ldots, X_n$ be a random sample from a $N(\mu_0, \sigma^2) = N(m_0, \sigma^2)$ population, then since $h(0.25) = h(0.75)$, $\sigma_{\hat{R}}^2 = \sigma_R^2$ for both settings (i) and (ii) described above. As a result, asymptotically, there is no improvement of $\hat{R}_n$ over R_n, whether we know the population mean or the population median.

Example 6.2. Suppose that $X_1, \ldots, X_n$ is a sample from a Gamma distribution $\Gamma(\gamma, \rho)$ with density $f(x) = \frac{\rho^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-\rho x}$ for $x > 0$, $\gamma > 0$, and $\rho > 0$. Let f_γ and F_γ be, respectively, the density and distribution function for a $\Gamma(\gamma, 1)$ distribution, then it is easy to see that $F(x) = F_\gamma(\rho x)$, $F^{-1}(x) = \frac{1}{\rho} F_\gamma^{-1}(x)$, and

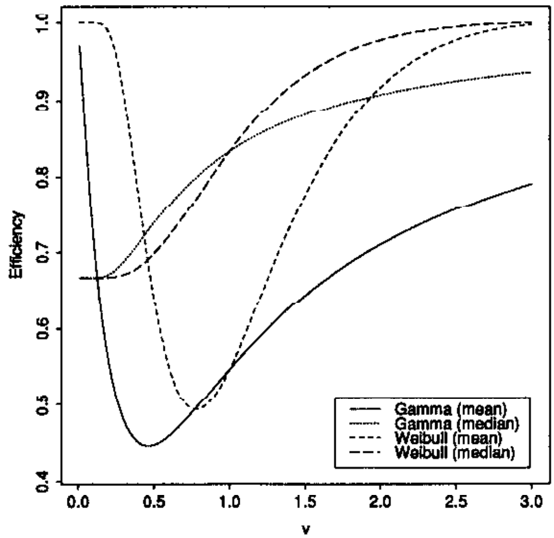

Fig. 1. The curves labeled "Gamma (mean)" and "Gamma (median)" correspond to the asymptotic relative efficiencies $e_1(\gamma)$ and $e_2(\gamma)$, respectively, whereas the curves labeled "Weibull (mean)" and "Weibull (median)" are, respectively, the asymptotic relative efficiencies $e_3(\gamma)$ and $e_4(\gamma)$.

$f(F^{-1}(x)) = \rho f_\gamma(F_\gamma^{-1}(x))$. In addition, we have $\mu_0 = \frac{\gamma}{\rho}$, $m_0 = \frac{1}{\rho}F_\gamma^{-1}(0.5)$, and $R = \frac{F_\gamma^{-1}(0.75)-F_\gamma^{-1}(0.25)}{2\rho}$.

(i) When the population mean μ_0 is known, according to Part (i) of Theorem 6.2 and Remark 6.1, the asymptotic relative efficiency of $\hat{R}_n$ relative to R_n is

$$e_1(\gamma) \equiv e(R_n, \hat{R}_n) = 1 - \frac{8}{\gamma} \cdot \frac{[F_\gamma^{-1}(0.75) - F_\gamma^{-1}(0.25)]^2}{\frac{1.5}{[f_\gamma(F_\gamma^{-1}(0.25))]^2} + \frac{1.5}{[f_\gamma(F_\gamma^{-1}(0.75))]^2} - \frac{1}{f_\gamma(F_\gamma^{-1}(0.25))f_\gamma(F_\gamma^{-1}(0.75))}},$$

which is independent of the scale parameter ρ. When the shape parameter γ is equal to 1, or equivalently, when the population has an exponential distribution with mean $\mu_0 = \frac{1}{\rho}$, it can then be shown that $e_1(1) = 1 - \frac{3}{8}(\log 3)^2 \approx 0.54739$. Figure 1 shows the curve of $e_1(\gamma)$ for $\gamma \in [0.01, 3]$, indicating that $e_1(\gamma)$ attains its minimum around $\gamma = 0.46$ with $e_1(0.46) \approx 0.44597$ and becomes larger for γ away from 0.46.

(ii) When the population median m_0 is known, the asymptotic relative efficiency of $\hat{R}_n$ relative to R_n is

$$e_2(\gamma) \equiv e(R_n, \hat{R}_n) = \frac{1 + a_\gamma^2}{1.5 + 1.5a_\gamma^2 - a_\gamma},$$

where $a_\gamma = f_\gamma(F_\gamma^{-1}(0.25))/f_\gamma(F_\gamma^{-1}(0.75))$. $e_2(\gamma)$ is also independent of ρ. When $\gamma = 1$, $e_2(1) = 5/6 = 0.83333$. Figure 1 shows that $e_2(\gamma)$ is increasing in γ.

Example 6.3. Suppose that $X_1, \ldots, X_n$ is a sample from a Weibull distribution $W(\gamma, \rho)$ with density $f(x) = \frac{\gamma}{\rho} x^{\gamma-1} e^{-x^\gamma/\rho}$ and distribution function $F(x) = 1 - e^{-x^\gamma/\rho}$ for $x > 0$, $\gamma > 0$, and $\rho > 0$. It is seen that $\mu_0 = \rho^{1/\gamma}\Gamma(1 + \frac{1}{\gamma})$, $m_0 = (\rho \log 2)^{1/\gamma}$, $f(F^{-1}(y)) = \gamma \mu_0^{-1}\Gamma(1 + \frac{1}{\gamma})(1 - y)[-\rho \log(1 - y)]^{(\gamma-1)/\gamma}$, and $R = \frac{1}{2}\rho^{1/\gamma}[(\log 4)^{1/\gamma} - (\log \frac{4}{3})^{1/\gamma}]$.

(i) When the population mean μ_0 is known, the asymptotic relative efficiency of $\hat{R}_n$ relative to R_n is, by Part (i) of Theorem 6.2 and Remark 6.1,

$$
\begin{aligned}
e_3(\gamma) &\equiv e(R_n, \hat{R}_n) \\
&= 1 - \frac{48\Gamma^2\left(1 + \frac{1}{\gamma}\right)}{\Gamma\left(1 + \frac{2}{\gamma}\right) - \Gamma^2\left(1 + \frac{1}{\gamma}\right)} \\
&\quad \cdot \frac{\left[\frac{F_{1+1/\gamma}(-\log 0.75) - 0.25}{3(-\log 0.75)^{(\gamma-1)/\gamma}} - \frac{F_{1+1/\gamma}(\log 4) - 0.75}{(\log 4)^{(\gamma-1)/\gamma}}\right]^2}{\left(\log \frac{4}{3}\right)^{-(2(\gamma-1))/\gamma} + 9(\log 4)^{-(2(\gamma-1))/\gamma} - 2\left(\log \frac{4}{3} \log 4\right)^{-(\gamma-1)/\gamma}},
\end{aligned}
$$

where $F_{1+1/\gamma}$ is the distribution function for a $\Gamma(1 + \frac{1}{\gamma}, 1)$ distribution, as defined in the preceding example. Note that $e_3(\gamma)$ does not involve the scale parameter ρ and $e_2(1) \approx 0.54739$ for an exponential distribution with mean $\mu_0 = \rho$. Figure 1 shows the curve of $e_3(\gamma)$, indicating that $e_3(\gamma)$ is minimized around $\gamma = 0.78$ with $e_3(0.78) \approx 0.49348$ and is getting larger and larger when the shape parameter γ is away from 0.78.

(ii) When the population median m_0 is known, the asymptotic relative efficiency of $\hat{R}_n$ relative to R_n is

$$
e_4(\gamma) \equiv e(R_n, \hat{R}_n) = \frac{1 + b_\gamma^2}{1.5 + 1.5 b_\gamma^2 - b_\gamma},
$$

where $b_\gamma = 3(\log \frac{4}{3}/\log 4)^{(\gamma-1)/\gamma}$. $e_4(\gamma)$ is also independent of ρ. When $\gamma = 1$, $e_4(1) = 5/6 = 0.83333$. The curve of $e_4(\gamma)$ is superimposed on those of $e_1(\gamma)$, $e_2(\gamma)$, and $e_3(\gamma)$ in Fig. 1, suggesting that $e_4(\gamma)$ is increasing in γ as $e_2(\gamma)$ is.

Remark 6.2. Examples 6.2 and 6.3 along with Fig. 1 appear to indicate that in the sense of achieving a greater variance reduction, the information about the population mean is, in general, more useful than that about the population median for estimating the semi-interquartile range R. The simulation results in the next section also support this phenomenon.

7. A simulation study

To evaluate the finite sample properties of the proposed semi-interquartile range estimator $\hat{R}_n$ and to compare $\hat{R}_n$ with the existing standard estimator R_n in the presence of auxiliary information (1.1), we perform a small simulation study to verify the asymptotic results of Theorem 6.2 and Remark 6.1 in moderate samples. Specifically, we consider estimating the semi-interquartile range R in situations corresponding to Examples 6.1 to 6.3. Recall that (i) if the population mean μ_0 is known, $g(x) = x - \mu_0$, while (ii) if the population median m_0 is known, $g(x) = I_{[x \leq m_0]} - 0.5$. We now consider the following four population distributions:

(a) The population has the standard normal distribution $N(0,1)$ with (i) $g(x) = x$ and (ii) $g(x) = I_{[x \leq 0]} - 0.5$.

(b) The population has the standard exponential distribution $E(1)$ with (i) $g(x) = x - 1$ and (ii) $g(x) = I_{[x \leq \log 2]} - 0.5$.

(c) The population has the Gamma distribution $\Gamma(\gamma, 1)$ with $\gamma = 0.46$. In this case, we have (i) $g(x) = x - 0.46$ and (ii) $g(x) = I_{[x \leq 0.19362]} - 0.5$. Note that $\gamma = 0.46$ approximately minimizes $e_1(\gamma)$ in Example 6.2.

(d) The population has the Weibull distribution $W(\gamma, 1)$ with $\gamma = 0.78$, $\mu_0 = \Gamma(1 + \frac{1}{\gamma}) = 1.15432$, and $m_0 = (\log 2)^{1/\gamma} = 0.62507$. In this case, we have (i) $g(x) = x - 1.15432$ and (ii) $g(x) = I_{[x \leq 0.62507]} - 0.5$. Notice that $\gamma = 0.78$ approximately minimizes $e_3(\gamma)$ in Example 6.3.

This simulation study aims to compare the performances of $\hat{R}_n$ and R_n by examining their biases, variances, and relative efficiencies. In our simulations, we generate 1000 independent sets of Monte Carlo random samples of size $n = 35$ from each of the four populations described above. For each population, based on 1000 estimates from simulation, we evaluate the biases and variances for R_n and $\hat{R}_n$. To obtain the corresponding bootstrap estimated biases and variances, we generate, for each simulation and population, two sets of 1000 independent bootstrap samples according to F_n and $\hat{F}_n$, respectively. All computations were done in double precision FORTRAN and the simulation results are summarized in Table 1 for setting (i) and in Table 2 for setting (ii).

In Tables 1 and 2, $\text{Bias}(R_n)$ and $\text{Var}(R_n)$ stand for, respectively, the average of 1000 biases of R_n and the sample variance of 1000 estimates R_n, whereas $\text{Bias}(\hat{R}_n)$ and $\text{Var}(\hat{R}_n)$ stand for, respectively, the average of 1000 biases of $\hat{R}_n$ and the sample variance of 1000 estimates $\hat{R}_n$. In addition, we use $\widehat{\text{Bias}}(R_n)$ and $\widehat{\text{Bias}}(\hat{R}_n)$ to represent, respectively, the averages of 1000 bootstrap estimates of the bias of R_n and that of $\hat{R}_n$. Here, each bootstrap estimate of the bias of $\hat{R}_n$ (R_n) is the difference between the average of the 1000 bootstrap replications $\hat{R}_n^*$ (R_n^*) and the estimate $\hat{R}_n$ (R_n). Similarly, we use $\widehat{\text{Var}}(R_n)$ and $\widehat{\text{Var}}(\hat{R}_n)$ to represent, respectively, the averages of 1000 bootstrap estimates of the variance of R_n and that of $\hat{R}_n$. Moreover, $e(R_n, \hat{R}_n)$ denotes the relative efficiency of $\hat{R}_n$ relative to R_n, i.e., $e(R_n, \hat{R}_n) = \text{Var}(\hat{R}_n)/\text{Var}(R_n)$.

For both settings (i) and (ii), the results of the simulation in the normal case are roughly as were expected in Example 6.1, though there is a slight improvement of $\hat{R}_n$ over R_n in terms of their biases and variances. In all other three cases under settings (i) and (ii), $\hat{R}_n$ produces appreciably smaller biases

Table 1. Biases, variances, and relative efficiencies of R_n and $\hat{R}_n$ and their bootstrap estimates with known mean.

Population	$\text{Bias}(R_n)$	$\text{Bias}(\hat{R}_n)$	$\text{Var}(R_n)$	$\text{Var}(\hat{R}_n)$	$\widehat{\text{Bias}}(R_n)$	$\widehat{\text{Bias}}(\hat{R}_n)$	$\widehat{\text{Var}}(R_n)$	$\widehat{\text{Var}}(\hat{R}_n)$	$e(R_n, \hat{R}_n)$
$N(0,1)$	0.01762	0.00439	0.01738	0.01676	0.00280	0.00295	0.02386	0.02167	0.96433
$E(1)$	0.01720	0.00385	0.01869	0.01071	0.00761	0.00403	0.02643	0.01322	0.57303
$\Gamma(0.46,1)$	0.01420	0.00804	0.00949	0.00476	0.01012	0.00427	0.01284	0.00528	0.50158
$W(0.78,1)$	0.03057	0.01244	0.04087	0.02121	0.01163	0.00565	0.05488	0.02478	0.51896

Table 2. Biases, variances, and relative efficiencies of R_n and $\hat{R}_n$ and their bootstrap estimates with known median.

Population	$\text{Bias}(R_n)$	$\text{Bias}(\hat{R}_n)$	$\text{Var}(R_n)$	$\text{Var}(\hat{R}_n)$	$\widehat{\text{Bias}}(R_n)$	$\widehat{\text{Bias}}(\hat{R}_n)$	$\widehat{\text{Var}}(R_n)$	$\widehat{\text{Var}}(\hat{R}_n)$	$e(R_n, \hat{R}_n)$
$N(0,1)$	0.02292	0.01821	0.01692	0.01679	0.01086	0.01063	0.02189	0.02131	0.99230
$E(1)$	0.02534	0.00935	0.02098	0.01769	0.01295	0.01300	0.02654	0.02106	0.84341
$\Gamma(0.46,1)$	0.01755	0.00462	0.00893	0.00692	0.01026	0.00932	0.01310	0.00929	0.77465
$W(0.78,1)$	0.03643	0.00944	0.04130	0.03276	0.02076	0.02005	0.05828	0.04297	0.79325

than R_n, and, as anticipated by the theory in Theorem 6.2, the variances of $\hat{R}_n$ are all considerably smaller than those of R_n. Furthermore, when the population mean is known, the relative efficiencies in the exponential, Gamma, and Weibull cases are, respectively, equal to 0.57303, 0.50158, and 0.51896, which are not far away from the corresponding asymptotic relative efficiencies 0.54739, 0.44597, and 0.49348 as described in Examples 6.2 and 6.3. Moreover, when the population median is known, the relative efficiencies in the exponential, Gamma, and Weibull cases are, respectively, identical to 0.84341, 0.77465, and 0.79325, as compared to the corresponding asymptotic relative efficiencies 0.83333, 0.72877, and 0.77194. These facts indicate that the proposed estimator $\hat{R}_n$ has significantly improved the standard estimator R_n in the presence of auxiliary information (1.1), and that the information about the population mean is more useful than that about the population median for estimating R in the sense of achieving a greater variance reduction, thus supporting the phenomena in Examples 6.2 and 6.3 and in Fig. 1. Finally, with regard to estimating the biases and variances of R_n and $\hat{R}_n$, Tables 1 and 2 reveal that the modified bootstrap estimated biases and variances for $\hat{R}_n$ work better than the standard bootstrap estimated biases and variances for R_n, although, on the whole, the bootstrap estimated variances work better than the bootstrap estimated biases for both R_n and $\hat{R}_n$. In general, biases are harder to estimate than variances and we may need more bootstrap replications to obtain a good estimate of bias.

8. Proofs

We first introduce several lemmas, which are used in the proof of the main results. Here we only present the proofs of Lemma 8.5 and Theorem 3.1. Other proofs are available from the author.

LEMMA 8.1. *If* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite, then* $\eta = \Sigma^{-1}(\frac{1}{n}\sum_{i=1}^n g(X_i)) + o_p(n^{-1/2})$. *As a result,* $\sqrt{n}\eta \xrightarrow{\mathcal{D}} N(0,\Sigma^{-1})$.

LEMMA 8.2. *Suppose* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite and* $\mathrm{E}\|g(X)\|^q < \infty$ *with some* $q > 2$.

(i) $\max_{1\le i\le n}\|g(X_i)\| = o(n^{1/q})$ *a.s.*

(ii) $\eta = o(n^{-1/q})$ *a.s.*

(iii) $\max_{1\le i\le n}|\eta^\tau g(X_i)| = o(1)$ *a.s.*

LEMMA 8.3. *If* $\mathrm{E}\|g(X)\|^q < \infty$ *with some* $q > 2$ *and* h *is a measurable function such that* $\mathrm{E}[|h(X)|\|g(X)\|^k] < \infty$ *for some* $k \in [0,q]$, *then*

(i) $\mathrm{E}_{\hat{F}_n}[|h(X_1^*)|\|g(X_1^*)\|^k] < \infty$ *a.s. for large* n;

(ii) $\mathrm{E}_{\hat{F}_n}[h(X_1^*)\|g(X_1^*)\|^k] = \mathrm{E}[h(X)\|g(X)\|^k] + o(1)$ *a.s. as* $n \to \infty$.

LEMMA 8.4. *Let* $\Sigma_n = \mathrm{E}_{\hat{F}_n}[g(X_1^*)g^\tau(X_1^*)]$. *Suppose* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite and* $\mathrm{E}\|g(X)\|^q < \infty$ *with some* $q > 2$.

(i) *With probability* 1, $\Sigma_n = \Sigma + o(1)$ *as* $n \to \infty$ *and* Σ_n *is positive definite for large* n.

(ii) $\|\Sigma_n^{-1}\|^2 = \|\Sigma^{-1}\|^2 + o(1)$ *with probability* 1 *as* $n \to \infty$.

(iii) *For almost all sample sequences* $X_1, \dots, X_n$ *and given* $(X_1, \dots, X_n)$, *as* $n \to \infty$,

$$\lambda = \Sigma_n^{-1} \left(\frac{1}{n} \sum_{j=1}^{n} g(X_j^*) \right) + o_p^*(n^{-1/2}) \quad as \quad and$$
$$\max_{1 \le j \le n} |\lambda^\tau g(X_j^*)| = o_p^*(1) \quad a.s.$$

As a result, for almost all sample sequences $X_1, \dots, X_n$, *as* $n \to \infty$, $\sqrt{n}\lambda \xrightarrow{\mathcal{D}} N(0, \Sigma^{-1})$.

LEMMA 8.5. *Let* $0 < \alpha < \beta < 1$ *be given. Suppose* $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ *is positive definite.*

(i) $\sup_{a \le x \le b} |\hat{F}_n(x) - F(x)| = O_p(n^{-1/2})$.

(ii) $\sup_{\alpha \le s \le \beta} |\hat{F}_n(\hat{F}_n^{-1}(s)) - s| = O_p(n^{-1})$.

(iii) *If* F *has continuous positive density* f *on* $[F^{-1}(\alpha), F^{-1}(\beta)]$, *then* $\sup_{\alpha \le s \le \beta} |\hat{F}_n^{-1}(s) - F^{-1}(s)| = O_p(n^{-1/2})$.

(iv) *If* F *has continuous positive density* f *on* $[F^{-1}(\alpha) - \epsilon, F^{-1}(\beta) + \epsilon]$ *for some* $\epsilon > 0$, *then for any* $A > 0$, $\sup_{F^{-1}(\alpha) - \epsilon/2 \le x \le F^{-1}(\beta) + \epsilon/2} \sup_{|y - x| \le A n^{-1/2}} |\hat{F}_n(x) - \hat{F}_n(y) - F(x) + F(y)| = o_p(n^{-1/2})$.

PROOF. Part (i) is a straightforward consequence of Theorem 2.1. For Part (ii), applying Lemma 8.1 gives

$$\begin{aligned} \sup_{\alpha \le s \le \beta} |\hat{F}_n(\hat{F}_n^{-1}(s)) - s| &\le \max_{1 \le i \le n} \sup_{\hat{F}_n(X_{(i-1)}) < s \le \hat{F}_n(X_{(i)})} |\hat{F}_n(X_{(i)}) - s| \\ &\le \max_{1 \le i \le n} |\hat{F}_n(X_{(i)}) - \hat{F}_n(X_{(i-1)})| \\ &= \sup_{1 \le i \le n} \hat{p}_i = O_p(n^{-1}), \end{aligned} \tag{8.1}$$

where $X_{(1)} \le \cdots \le X_{(n)}$ are the order statistics of $X_1, \dots, X_n$ and $X_{(0)} = a$. Turning to Part (iii), Part (i) and Lemma 8.1 yields

$$\begin{aligned} \Delta_n &\equiv \max_{1 \le i \le n} \sup_{\hat{F}_n(X_{(i-1)}) < s \le \hat{F}_n(X_{(i)})} |s - F(X_{(i)})| \\ &\le 2 \sup_{a \le x \le b} |\hat{F}_n(x) - F(x)| + \max_{1 \le i \le n} \hat{p}_i = O_p(n^{-1/2}). \end{aligned} \tag{8.2}$$

Since F has continuous positive density f on $[F^{-1}(\alpha), F^{-1}(\beta)]$, $f(F^{-1}(s))$ is bounded away from zero on $[\alpha, \beta]$, which, together with (8.2), implies that

$$\begin{aligned} &\sup_{\alpha \le s \le \beta} |\hat{F}_n^{-1}(s) - F^{-1}(s)| \\ &\quad \le \max_{1 \le i \le n} \sup_{\hat{F}_n(X_{(i-1)}) < s \le \hat{F}_n(X_{(i)})} |X_{(i)} - F^{-1}(s)| \end{aligned} \tag{8.3}$$

$$= \max_{1\le i\le n} \sup_{\hat F_n(X_{(i-1)})<s\le \hat F_n(X_{(i)})} |F^{-1}(F(X_{(i)})) - F^{-1}(s)|$$
$$\le \left[\sup_{\alpha\le s\le\beta} \frac{1}{f(F^{-1}(s))}\right] \Delta_n + o_p(\Delta_n) = O_p(n^{-1/2}).$$

To prove Part (iv), let $\delta_n = An^{-1/2}$ and $\{m_n\}$ be a sequence of positive integers such that $m_n = [n^\theta]$ with $\theta > \frac{1}{2}$, where $[n^\theta]$ denotes the largest integer less or equal to n^θ. We first divide the interval $[[m_n\{F^{-1}(\alpha) - \epsilon/2\}]/m_n, [m_n\{F^{-1}(\beta) + \epsilon/2\} + 1]/m_n]$ into subintervals $I_i = [t_i, t_{i+1}]$ for $i = 0, 1, \ldots, d_n - 1$, where $t_k = ([m_n\{F^{-1}(\alpha) - \epsilon/2\}] + k)/m_n$ for $k = 0, 1, \ldots, d_n$ with $d_n = [m_n\{F^{-1}(\beta) + \epsilon/2\} + 1] - [m_n\{F^{-1}(\alpha) - \epsilon/2\}]$. We then subdivide each interval $[t_i - [m_n\delta_n + 1]/m_n, t_i + [m_n\delta_n + 1]/m_n]$ for $i = 0, 1, \ldots, d_n$ into subintervals $I_{ij} = [t_{ij}, t_{i(j+1)}]$ for $j = -b_n, \ldots, b_n - 1$, where $t_{ij} = t_i + \frac{j}{m_n}$ for $j = -b_n, \ldots, b_n$ with $b_n = [m_n\delta_n + 1]$. Now let

$$\begin{aligned}
&H_n(x, y) = \hat F_n(x) - \hat F_n(y) - F(x) + F(y),\\
&K_n = \max_{0\le i\le d_n} \max_{-b_n\le j\le b_n} |H_n(t_i, t_{ij})|,\\
&a_i = F(t_{i+1}) - F(t_i), \qquad a_{ij} = F(t_{i(j+1)}) - F(t_{ij}),\\
&\qquad\qquad i = 0, \ldots, d_n,\ \ j = -b_n, \ldots, b_n,\\
&Z_{ijk} = Y_k(t_i) - Y_k(t_{ij}),\\
&\qquad\qquad i = 0, \ldots, d_n,\ \ j = -b_n, \ldots, b_n,\ \ k = 1, 2, \ldots, n,
\end{aligned} \tag{8.4}$$

where Y_k is defined in Theorem 2.1. Since F has continuous density f on $[F^{-1}(\alpha) - \epsilon, F^{-1}(\beta) + \epsilon]$, f is bounded on $[F^{-1}(\alpha) - \epsilon, F^{-1}(\beta) + \epsilon]$, i.e., $f(x) \le M_f$ for all x in $[F^{-1}(\alpha) - \epsilon, F^{-1}(\beta) + \epsilon]$ with M_f being some constant. Furthermore, applying the Mean Value Theorem gives $\max_{0\le i\le d_n} a_i \le \frac{M_f}{m_n}$ and $\max_{0\le i\le d_n} \max_{-b_n\le j\le b_n} \cdot a_{ij} \le \frac{M_f}{m_n}$ for large n. These facts, along with the monotonicity of $\hat F_n$ and F, implies that, for large n,

$$\max_{t_i\le x\le t_{i+1}} \max_{t_{ij}\le y\le t_{i(j+1)}} |H_n(x, y)| \le K_n + \frac{2M_f}{m_n}. \tag{8.5}$$

Inequality (8.5) further implies that, for large n,

$$\begin{aligned}
&\sup_{F^{-1}(\alpha)-\epsilon/2\le x\le F^{-1}(\beta)+\epsilon/2} \sup_{|y-x|\le\delta_n} |\hat F_n(x) - \hat F_n(y) - F(x) + F(y)|\\
&\quad\le K_n + \frac{2M_f}{m_n} = K_n + o(n^{-1/2}).
\end{aligned} \tag{8.6}$$

Thus, in order to prove Part (iv), it is enough to show that $K_n = o_p(n^{-1/2})$. For this purpose, by (8.4) and Theorem 2.1, we have

$$\begin{aligned}
K_n &= \max_{0<i<d_n} \max_{-b_n<j<b_n} \left|\frac{1}{n}\sum_{k=1}^n Y_k(t_i) + R_n(t_i) - \frac{1}{n}\sum_{k=1}^n Y_k(t_{ij}) - R_n(t_{ij})\right|\\
&\le \max_{0\le i\le d_n} \max_{-b_n<j<b_n} \left|\frac{1}{n}\sum_{k=1}^n Z_{ijk}\right| + o_p(n^{-1/2}).
\end{aligned} \tag{8.7}$$

It now suffices to show that $\max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}|\frac{1}{n}\sum_{k=1}^n Z_{ijk}| = o_p(n^{-1/2})$. By (8.4) and the definition of Y_k, it is easy to see that

$$(8.8)\quad \max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\left|\frac{1}{n}\sum_{k=1}^n Z_{ijk}\right|$$
$$\le \max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\left|\frac{1}{n}\sum_{k=1}^n\{[I_{[X_k\le t_i]}-I_{[X_k\le t_{ij}]}]-[F(t_i)-F(t_{ij})]\}\right|$$
$$+\max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\left|\mathrm{E}[g^{\tau}(X_1)\{I_{[X_1\le t_i]}-I_{[X_1\le t_{ij}]}\}]\Sigma^{-1}\right.$$
$$\left.\cdot\left(\frac{1}{n}\sum_{k=1}^n g(X_k)\right)\right|.$$

Moreover, it can be shown by the Mean Value Theorem that

$$(8.9)\qquad \max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\mathrm{Var}(I_{[X_1\le t_i]}-I_{[X_1\le t_{ij}]})$$
$$\le \max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}|F(t_i)-F(t_{ij})|\le M_f\left(\delta_n+\frac{1}{m_n}\right).$$

For any $\varepsilon>0$, using (8.4), (8.9), and Bernstein's inequality (Serfling (1980), p. 95) gives

$$\sum_{n=1}^{\infty}P\left(\sqrt{n}\max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\left|\frac{1}{n}\sum_{k=1}^n[[I_{[X_k\le t_i]}-I_{[X_k\le t_{ij}]}]-[F(t_i)-F(t_{ij})]]\right|>\varepsilon\right)$$
$$\le\sum_{n=1}^{\infty}\sum_{i=0}^{d_n}\sum_{j=-b_n}^{b_n}P\left(\left|\sum_{k=1}^n\{[I_{[X_k\le t_i]}-I_{[X_k\le t_{ij}]}]-[F(t_i)-F(t_{ij})]\}\right|>\delta_n n\varepsilon\right)$$
$$\le 2\sum_{n=1}^{\infty}(d_n+1)(2b_n+1)\exp\left(-\frac{n\delta_n^2\varepsilon^2}{2M_f(\delta_n+m_n^{-1})+\frac{4}{3}\delta_n\varepsilon}\right)<\infty,$$

which, along with the first Borel-Cantelli Lemma, implies that

$$(8.10)\qquad \max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\left|\frac{1}{n}\sum_{k=1}^n[I_{[X_k\le t_i]}-I_{[X_k\le t_{ij}]}]-[F(t_i)-F(t_{ij})]\right|$$
$$=o(n^{-1/2})\qquad \text{a.s.}$$

To deal with the second term in (8.8), applying (8.9) and Cauchy-Schwarz's inequality gives, for $0\le s\le r$,

$$(8.11)\qquad \max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}|\mathrm{E}[g_s(X_1)\{I_{[X_1\le t_i]}-I_{[X_1\le t_{ij}]}\}]|$$
$$\le[\mathrm{E}|g_s(X_1)|^2]^{1/2}\left[\max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}|F(t_i)-F(t_{ij})|^{1/2}\right]$$
$$=O(n^{-1/4}).$$

Let $\Sigma^{-1} = (\sigma^{st})_{r\times r}$, combining (8.11) and the central limit theorem gives

$$(8.12)\quad \max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\left|\mathrm{E}[g^\tau(X_1)\{I_{[X_1\le t_i]}-I_{[X_1\le t_{ij}]}\}]\Sigma^{-1}\left(\frac{1}{n}\sum_{k=1}^n g(X_k)\right)\right|$$
$$\le \sum_{s=1}^r\sum_{t=1}^r|\sigma^{st}|\left\{\max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}|\mathrm{E}[g_s(X_1)\{I_{[X_1\le t_i]}-I_{[X_1\le t_{ij}]}\}]|\right\}$$
$$\cdot\left|\frac{1}{n}\sum_{k=1}^n g_t(X_k)\right|$$
$$= O_p(n^{-3/4}).$$

Combining (8.8), (8.10) and (8.12) yields

$$\max_{0\le i\le d_n}\max_{-b_n\le j\le b_n}\left|\frac{1}{n}\sum_{k=1}^n Z_{ijk}\right| = o_p(n^{-1/2})+O_p(n^{-3/4}) = o_p(n^{-1/2}),$$

which, together with (8.7), implies $K_n = o_p(n^{-1/2})$, and this completes the proof of Lemma 8.5.

LEMMA 8.6. *Let $0<\alpha<\beta<1$ be given. Suppose that $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and $\mathrm{E}\|g(X)\|^q<\infty$ with some $q>2$.*

(i) $\sup_{a\le x\le b}|\hat F_n(x)-F(x)| = o(n^{-1/q})$ *a.s.*

(ii) $\sup_{\alpha\le s\le\beta}|\hat F_n(\hat F_n^{-1}(s))-s| = O(n^{-1})$ *a.s.*

(iii) *If F has continuous positive density f on $[F^{-1}(\alpha),F^{-1}(\beta)]$, then* $\sup_{\alpha\le s\le\beta}|\hat F_n^{-1}(s)-F^{-1}(s)| = o(n^{-1/q})$ *a.s.*

(iv) *If $\mathrm{E}\|g(X)\|^3<\infty$ and F has continuous positive density f on $[F^{-1}(\alpha)-\epsilon, F^{-1}(\beta)+\epsilon]$ for some $\epsilon>0$, then for any $A>0$,* $\sup_{F^{-1}(\alpha)-\epsilon/2\le x\le F^{-1}(\beta)+\epsilon/2}\cdot\sup_{|y-x|\le An^{-5/12}}|\hat F_n(x)-\hat F_n(y)-F(x)+F(y)| = o(n^{-1/2})$ *a.s.*

LEMMA 8.7. *Let $0<\alpha<\beta<1$ be given. Suppose $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and $\mathrm{E}\|g(X)\|^q<\infty$ with some $q>2$.*

(i) $\sup_{a\le x\le b}|\hat G_n(x)-\hat F_n(x)| = O_p^*(n^{-1/2})$ *a.s.*

(ii) $\sup_{\alpha\le s\le\beta}|\hat G_n(\hat G_n^{-1}(s))-s| = O_p^*(n^{-1})$ *a.s.*

(iii) *If F has continuous positive density f on $[F^{-1}(\alpha)-\epsilon,F^{-1}(\beta)+\epsilon]$ for some $\epsilon>0$, then we have* $\sup_{\alpha\le s\le\beta}|\hat G_n^{-1}(s)-\hat F_n^{-1}(s)| = o_p^*(n^{-1/q})$ *a.s.*

(iv) *If $\mathrm{E}\|g(X)\|^3<\infty$ and F has continuous positive density f on $[F^{-1}(\alpha)-\epsilon,F^{-1}(\beta)+\epsilon]$ for some $\epsilon>0$, then for any $A>0$,* $\sup_{F^{-1}(\alpha)\le x\le F^{-1}(\beta)}\cdot\sup_{|y-x|\le An^{-5/12}}|\hat G_n(x)-\hat G_n(y)-\hat F_n(x)+\hat F_n(y)| = o_p^*(n^{-1/2})$ *a.s.*

LEMMA 8.8. *Let $0<\alpha<\beta<1$ be given. Suppose that $\Sigma = \mathrm{E}[g(X)g^\tau(X)]$ is positive definite and $\mathrm{E}\|g(X)\|^3<\infty$. Suppose further that F has continuous positive density f on $[F^{-1}(\alpha)-\epsilon,F^{-1}(\beta)+\epsilon]$ for some $\epsilon>0$. Then,*

$$\sup_{\alpha\le s\le\beta}|F(\hat G_n^{-1}(s))-F(\hat F_n^{-1}(s))+\hat G_n(F^{-1}(s))-\hat F_n(F^{-1}(s))| = o_p^*(n^{-1/2})\qquad a.s.$$

PROOF OF THEOREM 3.1. For convenience, we use $F^{-1}\hat{F}_n F^{-1}(\cdot)$ etc. to denote composite functions. Let

$$r_{1n}(s) = -\left[F^{-1}\hat{F}_n F^{-1}(s) - F^{-1}FF^{-1}(s) + \frac{1}{n}\sum_{i=1}^{n} H_i(s)\right],$$
$$r_{2n}(s) = F^{-1}F\hat{F}_n^{-1}(s) - F^{-1}\hat{F}_n\hat{F}_n^{-1}(s) + F^{-1}\hat{F}_n F^{-1}(s) - F^{-1}FF^{-1}(s),$$
$$r_{3n}(s) = F^{-1}\hat{F}_n\hat{F}_n^{-1}(s) - F^{-1}(s),$$

then it is easy to see that

$$\hat{F}_n^{-1}(s) - F^{-1}(s) = \frac{1}{n}\sum_{i=1}^{n} H_i(s) + r_{1n}(s) + r_{2n}(s) + r_{3n}(s). \tag{8.13}$$

Let $x_n(s) = \hat{F}_n(F^{-1}(s))$, then $\sup_{\alpha\le s\le\beta}|x_n(s) - s| = O_p(n^{-1/2})$ by Part (i) of Lemma 8.5. Since F has continuous positive density f on $[F^{-1}(\alpha)-\epsilon, F^{-1}(\beta)+\epsilon]$, $f(F^{-1}(s))$ is bounded away from zero on $[\alpha,\beta]$. As a result, applying Taylor expansion and Theorem 2.1 gives

$$\begin{aligned}
&\sup_{\alpha\le s\le\beta}|r_{1n}(s)| \\
&\quad= \sup_{\alpha\le s\le\beta}\left|F^{-1}(x_n(s)) - F^{-1}(s) + \frac{1}{n}\sum_{i=1}^{n} H_i(s)\right| \\
&\quad\le \sup_{\alpha\le s\le\beta}\left|\frac{1}{f(F^{-1}(s))}[\hat{F}_n(F^{-1}(s)) - F(F^{-1}(s))] + \frac{1}{n}\sum_{i=1}^{n} H_i(s)\right| \\
&\qquad + o_p(n^{-1/2}) \\
&\quad= \sup_{\alpha\le s\le\beta}\left|\frac{R_n(F^{-1}(s))}{f(F^{-1}(s))}\right| + o_p(n^{-1/2}) = o_p(n^{-1/2}).
\end{aligned}\tag{8.14}$$

Let $\phi_n = \sup_{\alpha\le s\le\beta}|\hat{F}_n^{-1}(s) - F^{-1}(s)|$, then $\phi_n = O_p(n^{-1/2})$ by Part (iii) of Lemma 8.5. Applying Parts (i) and (iv) of Lemma 8.5 yields

$$\begin{aligned}
&\sup_{\alpha\le s\le\beta}|r_{2n}(s)| \\
&\quad\le \sup_{\alpha\le s\le\beta}\left|\frac{\hat{F}_n(F^{-1}(s)) - \hat{F}_n(\hat{F}_n^{-1}(s)) - F(F^{-1}(s)) - F(\hat{F}_n^{-1}(s))}{f(F^{-1}(s))}\right| \\
&\qquad + o_p(n^{-1/2}) \\
&\quad\le \left[\sup_{\alpha\le s\le\beta}\frac{1}{f(F^{-1}(s))}\right] \\
&\qquad \cdot \sup_{F^{-1}(\alpha)\le x\le F^{-1}(\beta)}\ \sup_{|y-x|\le\phi_n}|\hat{F}_n(x) - \hat{F}_n(y) - F(x) + F(y)| \\
&\qquad + o_p(n^{-1/2}) \\
&\quad= o_p(n^{-1/2}).
\end{aligned}\tag{8.15}$$

Finally, by Part (ii) of Lemma 8.5, we have

$$\sup_{\alpha\le s\le\beta}|r_{3n}(s)| \le \sup_{\alpha\le s\le\beta}\left|\frac{\hat{F}_n(\hat{F}_n^{-1}(s)) - s}{f(F^{-1}(s))}\right| + o_p(n^{-1}) - O_p(n^{-1}). \tag{8.16}$$

Let $r_n(s) = r_{1n}(s) + r_{2n}(s) + r_{3n}(s)$, then both (3.1) and (3.2) hold by combining (8.13)–(8.16), and this completes the proof of Theorem 3.1.

PROOF OF THEOREM 3.2. Theorem 3.2 is a straightforward consequence of (3.1), (3.2), and Theorem 2.1.

PROOF OF THEOREM 4.2. Theorem 4.2 can be proved by employing a similar approach as in the proof of Theorem 3.1.

PROOF OF THEOREM 4.3. Theorem 4.3 is a straightforward consequence of (4.3), (4.4), and Theorem 4.1.

Acknowledgements

The author wishes to thank two referees, an Associate Editor, and the Editor for making many useful comments and constructive suggestions on an earlier version of the paper, which led to a substantially improved presentation. This research was supported by an URAFP grant from the University of Toledo.

REFERENCES

Bickel, P. J. (1966). Some contributions to the theory of order statistics, *Proceedings Fifth Berkeley Symp. Math. Statist. and Prob.*, **1**, 575–592.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap, *Ann. Statist.*, **9**, 1196–1217.

Billingsley, P. (1968). *Weak Convergence of Probability Measures*, Wiley, New York.

Chen, J. and Qin, J. (1993). Empirical likelihood estimation for finite populations and the effective usage of auxiliary information, *Biometrika*, **80**, 107–116.

Haberman, S. J. (1984). Adjustment by minimum discriminant information, *Ann. Statist.*, **12**, 971–988.

Hall, P. and La Scala, B. (1990). Methodology and algorithms of empirical likelihood, *Internat. Statist. Rev.*, **58**, 109–127.

Kuk, A. Y. C. and Mak, T. K. (1989). Median estimation in the presence of auxiliary information, *J. Roy. Statist. Soc. Ser. B*, **51**, 261–269.

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional, *Biometrika*, **75**, 237–249.

Owen, A. B. (1990). Empirical likelihood confidence regions, *Ann. Statist.*, **18**, 90–120.

Owen, A. B. (1991). Empirical likelihood for linear models, *Ann. Statist.*, **19**, 1725–1747.

Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations, *Ann. Statist.*, **22**, 300–325.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

Sheehy, A. (1988). Kullback-Leibler constrained estimation of probability measures, Tech. Report, No. 137, Department of Statistics, University of Washington.

Tsirel'son, V. S. (1975). The density of the distribution of the maximum of a Gaussian process, *Theory Probab. Appl.*, **20**, 847–856.

Zhang, B. (1995*a*). M-estimation and quantile estimation in the presence of auxiliary information, *J. Statist. Plann. Inference*, **44**, 77–94.
Zhang, B. (1995*b*). Estimating a distribution function in the presence of auxiliary information, unpublished.
Zhang, B. (1995*c*). Bootstrapping with auxiliary information, unpublished.