Statistica Sinica **14**(2004), 1179-1198

REDUCED BOOTSTRAP FOR THE MEDIAN

M. D. Jiménez-Gamero, J. Muñoz-García and R. Pino-Mejías

Universidad de Sevilla

Abstract: In this paper we study a modified bootstrap that consists of only considering those bootstrap samples satisfying $k_1 \leq \nu_n \leq k_2$, for some $1 \leq k_1 \leq k_2 \leq n$, where ν_n is the number of distinct original observations in the bootstrap sample. We call it reduced bootstrap, since it only uses a portion of the set of all possible bootstrap samples. We show that, under some conditions on k_1 and k_2, the reduced bootstrap consistently estimates the distribution and the variance of the sample median. Unlike the ordinary bootstrap, the reduced bootstrap variance estimator does not require conditions on the population generating the data to be a consistent estimator, but does rely an adequate choice of k_1 and k_2. Since several choices of k_1 and k_2 yield consistent estimators, we compare the finite sample performance of the corresponding estimators through a simulation study. The simulation study also considers consistent variance estimators proposed by other authors.

Key words and phrases: Bootstrap, consistency, distribution estimation, sample median, variance estimation.

1. Introduction

Let $X_1, \ldots, X_n$ be a random sample of size n from a univariate population with distribution function F, and let $\theta = \inf\{t/F(t) \geq 1/2\}$ be the population median. If F has a positive derivative f at θ, $f(\theta) > 0$, then $Z_n = \sqrt{n}(\theta_n - \theta)$ converges weakly to $N(0, \sigma^2)$, where θ_n is the sample median, $\theta_n = \inf\{t/F_n(t) \geq 1/2\}$, F_n is the empirical distribution function of the sample and $\sigma^2 = 1/\{4f^2(\theta)\}$. If $f(\theta)$ were known, one could approximate the distribution of Z_n by its weak limit. However, $f(\theta)$ is rarely known. Another way to approximate the distribution of Z_n is by its bootstrap distribution. Bickel and Freedman (1981) have shown that if F has a unique median θ and a positive derivative f at θ that is continuous in a neighborhood of θ, then the bootstrap consistently estimates the distribution of Z_n.

The bootstrap can be also used to estimate standard errors. Hence, one can estimate the variance of Z_n, σ_n^2, through its bootstrap variance, σ_n^{*2}. Nevertheless, Ghosh, Parr, Singh and Babu (1984) have shown that σ_n^{*2} may not be a consistent estimator of σ_n^2. This is caused by the fact that Z_n^* may take some exceptionally large values. To solve this inconsistency one can put additional

conditions on F (Ghosh, Parr, Singh and Babu (1984) and Babu (1986)), modify the original sample by winsorizing or trimming it (Ghosh, Parr, Singh and Babu (1984)), modify the usual bootstrap variance estimator (Shao (1990, 1992)), or modify the resampling scheme generating the bootstrap samples (Jiménez-Gamero, Muñoz-García and Muñoz-Reyes (1998)).

In this paper we generalize the method in Jiménez-Gamero, Muñoz-García and Muñoz-Reyes (1998). The method considered by these authors, termed OBS, consists of only considering those bootstrap samples having a number of distinct original observations, ν_n, greater or equal than some constant k. The main advantage of OBS over the usual bootstrap is that the breakdown point of the OBS version of Z_n is greater than that of the usual bootstrap, which is $1/n$ regardless of the breakdown point of the estimator (Stromberg (1997)). This way, the OBS bootstrap variance estimator is not affected by exceptionally large values that Z_n^* may take.

The generalization considered here is motivated by the following observation made by Rao, Pathak and Koltchinskii (1997): as bootstrap samples are simple random samples of size n selected with replacement from the original sample, not all bootstrap samples are equally informative, due to the randomness in ν_n that occurs in different bootstrap samples. As these authors assert, the variability of ν_n is neither necessary nor desirable. To reduce this variability, which causes the inconsistency of σ_n^{*2}, we propose trimming ν_n, that is, only considering those bootstrap samples satisfying $k_1 \leq \nu_n \leq k_2$, for some $1 \leq k_1 \leq k_2 \leq n$. We call it the reduced bootstrap (RB), since it only uses a portion of all possible bootstrap samples. OBS is a particular case of RB with $k_2 = n$.

In this paper we show that, for suitable choices of k_1 and k_2, the RB estimator of the distribution of Z_n has the same asymptotic accuracy as the usual bootstrap estimator (ordinary bootstrap variability of ν_n is not necessary) and that, in contrast to usual bootstrap, the RB estimator of the variance of Z_n is consistent (ordinary bootstrap variability of ν_n is not desirable).

The paper is organized as follows. In Section 2 we introduce some notation, describe RB and discuss the choice of k_1 and k_2. In Section 3 we give a preliminary result that will be used in the proofs of the results in the following sections. (To shorten the proofs of our main results, we include an appendix containing most of the required technical results we need to demonstrate them). The result of Section 3 deserves to be displayed since it has interest per se − it gives the order of the error of the RB distribution estimator of the sample mean, for certain choices of k_1 and k_2. In Section 4 we show that, under the conditions on k_1 and k_2 previously stated, the RB consistently estimates the distribution of the

sample median. Section 5 gives a similar result for the RB variance estimator of the sample median. Simulation results are displayed in Section 6. Finally, the Appendix contains some auxiliary lemmas used in the proofs of the results in Sections 3, 4 and 5.

2. The Reduced Bootstrap

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample from an unknown distribution F, and let $T_n = T_n(\mathbf{X}; F)$ be a statistic of interest. Let F_n be the empirical distribution function of $\mathbf{X}$ and let $\mathbf{X}^* = (X_1^*, \ldots, X_n^*)$ be a random sample drawn from F_n. $\mathbf{X}^*$ is called a bootstrap sample. The bootstrap method estimates the distribution of T_n through the conditional distribution of $T_n^* = T_n(\mathbf{X}^*; F_n)$, given $X_1, \ldots, X_n$. This conditional distribution is called the bootstrap distribution of T_n. In particular, the bootstrap estimates the variance of T_n by the conditional variance of T_n^*. Throughout this paper P_*, E_* and var_* denote the bootstrap conditional probability law, the bootstrap conditional expectation and the bootstrap conditional variance, given $X_1, \ldots, X_n$, respectively.

For each bootstrap sample $\mathbf{X}^*$, let $N_i = card\{X_j^* = X_i\}$, $1 \leq i \leq n$. The vector $N = (N_1, \ldots, N_n)$ is the resampling vector and, under the ordinary bootstrap, has a multinomial distribution, $\mathcal{M}(n; 1/n, \ldots, 1/n)$. Let $\nu_n = \nu_n(\mathbf{X}^*)$ be the number of different elements contained in $\mathbf{X}^*$, that is, $\nu_n = \sum_{i=1}^n I(N_i > 0)$, where $I(A) = 1$ if A holds and $I(A) = 0$ otherwise. Given k_1 and k_2, with $1 \leq k_1 \leq k_2 \leq n$, the RB estimates the distribution of a statistic T_n, $P(T_n \leq x)$, through $P_*(T_n^* \leq x/k_1 \leq \nu_n \leq k_2)$, the RB distribution of T_n. This way, only αn^n bootstrap samples are used to estimate the distribution of T, where $\alpha = P(k_1 \leq \nu_n \leq k_2)$.

To get a bootstrap sample $\mathbf{X}^*$ with $k_1 \leq \nu_n(\mathbf{X}^*) \leq k_2$, for some fixed $1 \leq k_1 \leq k_2 \leq n$, we can proceed in several fashions. One way to do this is to imitate the algorithm in Muñoz-García, Pino-Mejías, Muñoz-Pichardo and Cubiles-de-la-Vega (1997): Step 1, draw a bootstrap sample $\mathbf{X}^*$; Step 2, calculate $\nu_n = \nu_n(\mathbf{X}^*)$; Step 3, if $\nu_n < k_1$ or $k_2 < \nu_n$ then throw away the generated bootstrap sample and go to Step 1, otherwise the generated bootstrap sample is considered to be valid. Another way is as follows: first, select a simple random sample of size k_2 without replacement from $\{1, \ldots, n\}$, say $\mathcal{I}_1$; second, select a simple random sample of size k_1 without replacement from $\mathcal{I}_1$, say $\mathcal{I}_2 = \{i_1, i_2, \ldots, i_{k_1}\}$; third, select a simple random sample of size $n - k_1$ with replacement from $\mathcal{I}_1$, say $J = (j_1, \ldots, j_{n-k_1})$; fourth, let $(l_1, \ldots, l_n)$ be an n-vector whose components are obtained by randomly permuting the string $(i_1, \ldots, i_{k_1}, j_1, \ldots, j_{n-k_1})$; finally, the n-vector $(X_{l_1}, \ldots, X_{l_n})$ is a bootstrap sample satisfying $k_1 \leq \nu_n \leq k_2$.

The choice of an algorithm to get $\mathbf{X}^*$ with $k_1 \leq \nu_n(\mathbf{X}^*) \leq k_2$ depends on $\alpha = P(k_1 \leq \nu_n \leq k_2)$. If α is large then the first algorithm requires less computational effort than the second one, and the reverse is true when α is small. A minimum requirement is to choose k_1 and k_2 so that RB consistently estimates the distribution of Z_n. To discuss this choice we first introduce some notation.

Let Φ denote the standard normal distribution function. Let $k_1 = k_1(n)$ and $k_2 = k_2(n)$ be two integers with $1 \leq k_1 \leq k_2 \leq n$, $w_1 = (k_1 - 1 - np)n^{-1/2}{\sigma_0}^{-1}$, $w_2 = (k_2 - np)n^{-1/2}{\sigma_0}^{-1}$, ${\sigma_0}^2 = pq - q^2$, $p = 1 - \mathrm{e}^{-1}$, $q = 1 - p$ and, for any fixed $\varepsilon \geq 0$, let

$$v_\varepsilon = \begin{cases} w_1 + \varepsilon & \text{if } w_1 \to \infty, \\ |w_2 - \varepsilon| & \text{if } w_2 \to -\infty. \end{cases}$$

Since $P(\nu_n \leq k) = \Phi(w) + o(1)$ with $w = (k - np)/\sqrt{n}\sigma_0$ (Johnson and Kotz (1977, p.318)), if k_1 and k_2 are such that condition C.1 below holds, then imitating the proof of Proposition 3.1 in Jiménez-Gamero, Muñoz-García and Muñoz-Reyes (1998) one can show that the RB consistently estimates the distribution of Z_n.

Condition C.1. $\Phi(w_2) - \Phi(w_1) \geq \alpha_0$, $\forall n \geq n_0$, for some $n_0 \in \mathbb{N}$ and some fixed constant $\alpha_0 > 0$.

Condition C.1 means that, at least for large n, the percentage of bootstrap samples used by the RB is greater or equal than $100\alpha_0\%$. The question that arises is that if RB can also estimate consistently the distribution of Z_n when k_1 and k_2 are chosen such that the percentage of used bootstrap samples tends to 0. The method of proof of Proposition 3.1 in Jiménez-Gamero, Muñoz-García and Muñoz-Reyes (1998) is not useful to handle this case. In this paper we will see that, employing the results in Section 3, it is possible to show the consistency of the RB distribution of Z_n when k_1 and k_2 are chosen such that $P(k_1 \leq \nu_n \leq k_2) \to 0$ in the way specified by the following conditions.

Condition C.2. $|w_1 - w_2| \to 0$ and $w_1 \to l$, for some $l \in \mathbb{R}$.

Condition C.3. $w_1 \to \infty$ or $w_2 \to -\infty$, $|w_2 - w_1| \geq \varepsilon$ and $2v_\varepsilon^2 \leq a + \log n + 2\log\log\log n$, $\forall n \geq n_0$, for some $n_0 \in \mathbb{N}$ and some fixed constants $\varepsilon \geq 0$ and $a \in \mathbb{R}$.

Condition C.4. $w_1 \to \infty$ or $w_2 \to -\infty$ and $v_0 \leq a\log^{1/2} n$, $\forall n \geq n_0$, for some $n_0 \in \mathbb{N}$ and some fixed constant $a < 1$.

If C.2 holds then, in the limit, all the considered bootstrap samples have the same number of different elements; if C.3 or C.4 hold, then they all have

a very large number of different elements ($w_1 \to \infty$) or, by contrast, they all have a very small number of different elements ($w_2 \to -\infty$). Condition C.3 is more restrictive than C.4. Although under both conditions the RB consistently estimates the distribution of Z_n, if C.3 holds we can get a better approximation to the distribution of Z_n.

3. A Preliminary Result

To demonstrate that the RB consistently estimates the distribution of Z_n, we use the result in Theorem 3.1 below. It has interest on its own, because it gives the order of the error of the RB distribution estimator of the sample mean. Let

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \;\; s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \text{ and } \bar{X}_n^* = \frac{1}{n}\sum_{i=1}^{n} X_i^*.$$

Theorem 3.1. *If $E(|X_1|^3) < \infty$ then, as $n \to \infty$, $P_*\{n^{1/2}(\bar{X}_n^* - \bar{X}_n) \le x s_n \mid k_1 \le \nu_n \le k_2\} = \Phi(x) + r_n$ uniformly in x, where*

$$r_n = \begin{cases} O(n^{-1/2}) & \text{if C.1 or C.2 holds,} \\ O\left(n^{-1/4}(\log\log n)^{1/2}\right) & \text{if C.3 holds,} \\ o(1) & \text{if C.4 holds.} \end{cases} \tag{1}$$

Proof. Let $Y_1, Y_2, \ldots$ be a sequence of independent and identically distributed Poisson variables with mean 1, $S = \sum_{j=1}^{n} Y_j$, $T = \sum_{j=1}^{n} I(Y_j > 0)$ and $D_i = D_{i,n} = (X_i - \bar{X}_n)/s_n$, $1 \le i \le n$. With this notation,

$$\begin{aligned} &P_*\left\{n^{1/2}(\bar{X}_n^* - \bar{X}_n) \le x s_n \mid k_1 \le \nu_n \le k_2\right\} \\ &= \frac{P_*\left\{\sum_{i=1}^{n} D_i Y_i \le n^{1/2} x,\, S = n,\, k_1 \le T \le k_2\right\}}{P(S = n,\, k_1 \le T \le k_2)}. \end{aligned} \tag{2}$$

From (2) and Lemmas 7.1 and 7.4 in the Appendix, $P_*\{n^{1/2}(\bar{X}_n^* - \bar{X}_n) \le x \mid k_1 \le \nu_n \le k_2\} = \Phi(x) + O(t_n^{-1})$, where

$$t_n = \begin{cases} n^{1/2}\{\Phi(w_2) - \Phi(w_1)\} & \text{if} \quad |w_2 - w_1| \ge \varepsilon, \text{ for some } \varepsilon > 0, \\ n^{1/2} \displaystyle\sum_{k=k_1}^{k_2} \phi_0(v_k)/(k_2 - k_1 + 1) & \text{if} \quad |w_2 - w_1| \to 0. \end{cases} \tag{3}$$

The assumed conditions on k_1 and k_2 give the orders in (1).

Corollary 3.1. *Let* $\mu = E(X_1)$ *and* $\varrho^2 = var(X_1)$. *Under the conditions of Theorem* 3.1, *as* $n \to \infty$, *if* C.1 *or* C.2 *hold then*

$$\sup_x n^{1/2} \left| P_* \left\{ n^{1/2}(\bar{X}_n^* - \bar{X}_n) \leq x s_n \mid k_1 \leq \nu_n \leq k_2 \right\} - P\left\{ n^{1/2}(\bar{X}_n - \mu) \leq x\varrho \right\} \right| = O(1);$$

if C.3 *holds then*

$$\sup_x n^{1/4}(\log\log n)^{-1/2} \left| P_* \left\{ n^{1/2}(\bar{X}_n^* - \bar{X}_n) \leq x s_n \mid k_1 \leq \nu_n \leq k_2 \right\} - P\left\{ n^{1/2}(\bar{X}_n - \mu) \leq x\varrho \right\} \right| = O(1);$$

if C.4 *holds then*

$$\sup_x \left| P_* \left\{ n^{1/2}(\bar{X}_n^* - \bar{X}_n) \leq x s_n \mid k_1 \leq \nu_n \leq k_2 \right\} - P\left\{ n^{1/2}(\bar{X}_n - \mu) \leq x\varrho \right\} \right| = o(1).$$

Note that if $E(|X_1|^3) < \infty$ and k_1 and k_2 satisfy C.1 or C.2, then the RB distribution estimator of the sample mean has the same asymptotic accuracy as the ordinary bootstrap estimator (see Theorem 1.C of Singh (1981)), while under C.3 and C.4, the RB distribution estimator of the sample mean behaves worse than the ordinary bootstrap one. Nevertheless, in all cases, the RB distribution estimator of the sample mean is consistent, which is the only condition we need to prove the consistency of the RB distribution estimator of the sample median. We will use this fact in the proof of Theorem 4.1 in next Section, which shows that, for all the considered conditions on k_1 and k_2, the RB consistently estimates the distribution of the sample median.

4. Consistency of the RB Distribution of the Sample Median

Bickel and Freedman (1981) have shown that if

$$\begin{aligned} &F \text{ has a unique median } \theta \text{ and a positive derivative } F' = f \text{ at } \theta \\ &\text{that is continuous in a neighborhood of } \theta, \end{aligned} \tag{4}$$

then the bootstrap estimates consistently the distribution of Z_n. Next we show that, under some conditions on k_1 and k_2, the RB distribution of Z_n is also a consistent estimator. For each bootstrap sample $\mathbf{X}^*$, let F_n^* be its empirical distribution function and $\theta_n^* = \inf\{t/F_n^*(t) \geq 1/2\}$.

Theorem 4.1. *If* F *satisfies* (4) *and* k_1, k_2 *satisfy one of* C.1, C.2, C.3, C.4, *then as* $n \to \infty$, $\sup_x \left| P_*\{n^{1/2}(\theta_n^* - \theta_n) \leq x/k_1 \leq \nu_n \leq k_2\} - P\{n^{1/2}(\theta_n - \theta) \leq x\} \right| = o(1)$.

Proof. Since

$$\begin{aligned} &P_*\Big\{n^{1/2}(\theta_n^*-\theta_n)\le x \,/\, k_1\le\nu_n\le k_2\Big\}\\ &=P_*\left[\frac{W_n^*-n\mu_n^*}{\{n\mu_n^*(1-\mu_n^*)\}^{1/2}}\ge-\frac{n\mu_n^*-n/2}{\{n\mu_n^*(1-\mu_n^*)\}^{1/2}}\,/\,k_1\le\nu_n\le k_2\right], \end{aligned} \tag{5}$$

where $W_n^*=nF_n^*(\theta_n+xn^{-1/2})$ and $\mu_n^*=F_n(\theta_n+xn^{-1/2})$, the result follows from (5), Theorem 3.1 and the fact that Z_n converges to $N(0,\sigma^2)$.

To prove Theorem 4.1 we use two facts: under the assumed conditions on k_1 and k_2, the RB consistently estimates the distribution of the sample mean; if F satisfies (4), then $\{F_n(\theta_n+xn^{-1/2})-1/2\}/xn^{-1/2}=f(\theta)+u_n$, with $u_n=o(1)$. By assuming stronger conditions on F we can reduce the order of u_n. More precisely, if

$$\begin{aligned} &F \text{ has a bounded second derivative in a neighborhood of } \theta \text{ and}\\ &f(\theta)>0, \end{aligned} \tag{6}$$

then Theorem 2 in Singh (1981) shows that the difference between the distribution of Z_n and its bootstrap estimator is $O(n^{-1/4}(\log\log n)^{1/2})$. The next Theorem shows that for appropriate choices of k_1 and k_2 the RB also satisfies this property.

Theorem 4.2. *If F satisfies* (6) *and k_1, k_2 satisfy one of* C.1, C.2, C.3, *then as $n\to\infty$,*

$$\begin{aligned} &\sup_x\Big|P_*\Big\{n^{1/2}(\theta_n^*-\theta_n)\le x\,/\,k_1\le\nu_n\le k_2\Big\}-P\Big\{n^{1/2}(\theta_n-\theta)\le x\Big\}\Big|\\ &=O(n^{-1/4}(\log\log n)^{1/2}). \end{aligned}$$

Proof. We have

$$P_*\Big\{n^{1/2}|\theta_n^*-\theta_n|>\log n\,/\,k_1\le\nu_n\le k_2\Big\}\le\frac{P_*\Big\{n^{1/2}|\theta_n^*-\theta_n|>\log n\Big\}}{P(k_1\le\nu_n\le k_2)},$$

$$\begin{aligned} &P_*\Big\{n^{1/2}|\theta_n^*-\theta_n|>\log n\Big\}\\ &\le P_*\Big\{\sum_{i=1}^n V_i^*-\sum_{i=1}^n E_*(V_i^*)>n\delta_1\Big\}+P_*\Big\{\sum_{i=1}^n W_i^*-\sum_{i=1}^n E_*(W_i^*)\ge n\delta_2\Big\}, \end{aligned}$$

where $V_i^*=I(X_i^*>\theta_n+n^{-1/2}\log n)$, $\delta_1=F_n(\theta_n+n^{-1/2}\log n)-1/2$, $W_i^*=I(X_i^*\le\theta_n-n^{-1/2}\log n)$ and $\delta_2=1/2-F_n(\theta_n-n^{-1/2}\log n)$. Using these bounds and Lemma 2.3.2 in Serfling (1980), we get

$$P_*\Big\{n^{1/2}|\theta_n^*-\theta_n|>\log n\,/\,k_1\le\nu_n\le k_2\Big\}\le\frac{\mathrm{e}^{-2n\delta_1^2}+\mathrm{e}^{-2n\delta_2^2}}{P(k_1\le\nu_n\le k_2)}.$$

According to Lemma 3.2 in Singh (1981), $n\delta_i^2 = f^2(\theta)\log^2 n + o(1)$, $i = 1, 2$. Using this and Lemma 7.2, we find that the right side of the above inequality is $O(n^{-1})$, that is,

$$P_*\left\{n^{1/2}|\theta_n^* - \theta_n| > \log n \,/\, k_1 \leq \nu_n \leq k_2\right\} = O(n^{-1}). \tag{7}$$

Now, from (5), Theorem 3.1 and Lemma 3.2 in Singh (1981), we obtain

$$P_*\left\{n^{1/2}(\theta_n^* - \theta_n) \leq x \,/\, k_1 \leq \nu_n \leq k_2\right\} = \Phi(x\sigma^{-1}) + O(n^{-1/4}(\log\log n)^{1/2}), \tag{8}$$

uniformly in $|x| \leq \log n$. Finally, the result follows from (7), (8) and the Berry-Esseén bound for $\theta_n - \theta$ (see, for example, Theorem 2.3.3.C in Serfling (1980)).

Needless to say the results in Theorems 4.1 and 4.2 extend appropriately for any general quantile, $\theta_\xi = \inf\{t/F(t) \geq \xi\}$, $0 < \xi < 1$.

5. Consistency of the RB Variance of the Sample Median

It is well known that convergence in distribution of a random sequence does not imply convergence of moments. An example in the sample median: although its bootstrap distribution is strongly consistent, its bootstrap variance estimator may diverge to infinity, while the asymptotic variance of θ_n is finite. This inconsistency is caused by the fact that $|\theta_n^* - \theta_n|$ may take some exceptionally large values (see the example in Ghosh, Parr, Singh and Babu (1984)). To ensure the consistency of the bootstrap variance estimator we need some additional tail condition on F. Examples are the moment conditions in Ghosh, Parr, Singh and Babu (1984) and Babu (1986). Another way to get a consistent variance estimator is to truncate $|\theta_n^* - \theta_n|$, as proposed in Shao (1992). As noted earlier, the RB estimator is also based on truncating (it truncates ν_n) but, in contrast to Shao's method, it truncates before evaluating θ_n^*. Both methods restrict θ_n^* from attaining large deviations from θ_n. One might say that Shao's method is a post-sampling correction to the ordinary bootstrap and that the RB is a pre-sampling or while-sampling correction to the ordinary bootstrap.

The following Lemma shows that, for certain choices of k_1 and k_2, under RB, the absolute difference $|\theta_n^* - \theta_n|$ is bounded with probability one. Hence we do not have to impose additional tail conditions on F to get a consistent variance estimator, as is shown in Theorem 5.1 below.

Lemma 5.1. *If $n\beta_0 \leq k_1 \leq k_2 \leq n$, for some $\beta_0 > 1/2$, then $P_*(X_{[n\beta]:n} \leq \theta_n^* \leq X_{[n(1-\beta)]:n} \,/\, k_1 \leq \nu_n \leq k_2) = 1$, where $X_{r:n}$ denotes the rth order statistic, $1 \leq r \leq n$, $\beta = \beta_0 - 1/2$ and $[x]$ denotes the greatest integer less or equal than x.*

Theorem 5.1. *If the assumptions of Theorem* 4.1 *hold and* $n\beta_0 \le k_1$ *for some* $1/2 < \beta_0 \le p$ *if* $w_1 \not\to -\infty$ *and* $1/2 < \beta_0 < p$ *if* $w_1 \to -\infty$, *then* $var_*\{\sqrt{n}(\theta_n^* - \theta_n) / k_1 \le \nu_n \le k_2\} = \sigma_n^2 + o(1)$.

Proof. By the assumed conditions on k_1 and k_2 and Theorem 4.1, it suffices to show that

$$\begin{aligned} &E_*\{|\sqrt{n}(\theta_n^* - \theta_n)|^{2+\delta} / k_1 \le \nu_n \le k_2\} \\ &= (1+\delta)\int_0^\infty t^{1+\delta} P_*(\sqrt{n}|\theta_n^* - \theta_n| > t / k_1 \le \nu_n \le k_2)dt < \infty \end{aligned} \tag{9}$$

for some $\delta > 0$. To bound the integral in (9), we consider three zones: (I) $t \in [1, c\log^{1/2} n]$, (II) $t \in (c\log^{1/2} n, d\sqrt{n})$ and (III) $t \in [d\sqrt{n}, \infty)$, where the constants c and d will be specified later. From Lemma 7.6, Markov's inequality and Lemma 7.5, in zone (I) of t, $P_*(\sqrt{n}|\theta_n^* - \theta_n| > t / k_1 \le \nu_n \le k_2) = O(t^{-4})$ and hence for any $0 < \delta < 2$,

$$\int_{\text{zone (I)}} t^{1+\delta} P_*(\sqrt{n}|\theta_n^* - \theta_n| > t / k_1 \le \nu_n \le k_2)dt < \infty.$$

In zone (II) of t, by our Lemma 7.6 and Lemma 2.3.2 in Serfling (1980), we obtain

$$\begin{aligned} P_*(\sqrt{n}|\theta_n^* - \theta_n| > t/k_1 \le \nu_n \le k_2) &\le P_*(\sqrt{n}|\theta_n^* - \theta_n| > c\log^{1/2} n/k_1 \le \nu_n \le k_2) \\ &\le \frac{2}{n^{2c^2} P(k_1 \le \nu_n \le k_2)}, \end{aligned}$$

and therefore

$$\int_{\text{zone (II)}} t^{1+\delta} P_*(\sqrt{n}|\theta_n^* - \theta_n| > t / k_1 \le \nu_n \le k_2)dt = O\Big(\frac{n^{1+\delta/2-2c^2}}{P(k_1 \le \nu_n \le k_2)}\Big).$$

From Lemma 7.2, to ensure that the integral in zone (II) is finite, it suffices to take the constant c such that $c \ge (2+\delta/2)^{1/2}$. Finally, since $m = \max\{\theta_n - X_{[\beta n]:n}, X_{[(1-\beta)n]:n} - \theta_n\} = O(1)$, with $\beta = \beta_0 - 1/2$, if we take the constant d such that $m \le d$, then by Lemma 5.1 we have that

$$\int_{\text{zone (III)}} t^{1+\delta} P_*(\sqrt{n}|\theta_n^* - \theta_n| > t / k_1 \le \nu_n \le k_2)dt = 0.$$

This completes the proof.

6. A Simulation Study

In Sections 4 and 5 we have seen that, for adequate choices of k_1 and k_2, the RB estimators of the distribution and the variance of Z_n are consistent. Since

several choices of k_1 and k_2 are reasonable, we have carried out a simulation experiment to compare the finite sample performance of the corresponding estimators, including the ordinary bootstrap. We consider six choices for k_1 and k_2, displayed in Table 1. Note that Method 1 in this Table is the ordinary bootstrap.

Table 1. Choices of k_1 and k_2.

Method	k_1	k_2
1	1	n
2	$[np - \sqrt{npq}] + 1$	n
3	$[np - \sqrt{npq}] + 1$	$[np + \sqrt{npq}]$
4	$[np] + 1$	$[np] + 1$
5	$[np + \sqrt{npq}] + 1$	$[np + \sqrt{npq}] + 1$
6	$[np + \sqrt{npq}] + 1$	n

For the estimation of the variance of Z_n we have also considered the following consistent estimators:

- The estimator developed by Bloch and Gastwirth (1968), with $m = 0.5n^{4/5}$, denoted estimator 7.
- The estimator based on the interquartile range, $(IQR^*/\{\Phi^{-1}(3/4) - \Phi^{-1}(1/4)\})^2$, where IQR^* is the interquartile range of the ordinary bootstrap distribution of Z_n, denoted estimator 8.
- The estimator proposed by Shao (1992); for estimating the variance of Z_n it is equivalent to the usual bootstrap estimator obtained by winsorizing the original sample, that is, replacing $X_{i:n}$ by $X_{[n\varepsilon_1]:n}$ for $1 \leq i \leq [n\varepsilon_1]$ and by $X_{[n(1-\varepsilon_2)]:n}$ for $[n(1-\varepsilon_2)] \leq i \leq n$, for some $0 < \varepsilon_1, \varepsilon_2 < 1/2$. We have taken $\varepsilon_1 = \varepsilon_2 = 0.10$. We refer to it as estimator 9.

To study the corresponding variance and distribution estimators (estimators 1 to 9 for the variance and estimators 1 to 6 for the distribution) we generated $M = 10.000$ samples of size $n = 21$ from a standard normal population, N(0,1). For each method (except for the variance estimator 7) and from each sample $\mathbf{X}^m$, $1 \leq m \leq M$, we have first generated $B = 1000$ bootstrap samples, $\mathbf{X}^{*m,1}, \ldots, \mathbf{X}^{*m,B}$, and then

- We estimated $\sigma_n^2 = var\{\sqrt{n}(\hat{\theta}_n - \theta)\}$ by $\hat{\sigma}_n^2(m)$, the sample variance of $\sqrt{n}(\hat{\theta}_n^{*m,b} - \hat{\theta}_n^m)$, $1 \leq b \leq B$, where $\hat{\theta}_n^m$ is the sample median of $\mathbf{X}^m$ and $\hat{\theta}_n^{*m,b}$ is the sample median of $\mathbf{X}^{*m,b}$, $1 \leq b \leq B$.
- We considered the following intervals

$$(-\infty, -3),\ [-3, -2.9),\ [-2.9, -2.8),\ [-2.8, -2.7),\ \ldots,\ [2.9, 3),\ [3, \infty) \quad (10)$$

and, for each interval, we calculated $f(I,m)$, the fraction of $\sqrt{n}(\hat{\theta}_n^{*m,b} - \hat{\theta}_n^m)$, $1 \le b \le B$, falling in the interval I.

Finally, we approximated the bias and the mean squared error (mse) of the corresponding variance estimator by

$$\text{bias} \approx \frac{1}{M}\sum_{m=1}^{M} \hat{\sigma}_n^2(m) - \sigma_n^2 \quad \text{and} \quad \text{mse} \approx \frac{1}{M}\sum_{m=1}^{M} \{\hat{\sigma}_n^2(m) - \sigma_n^2\}^2,$$

respectively. To evaluate the corresponding distribution function estimator we considered the following global measure of the bias, $BS = 100 \times \sum_{I\in\mathcal{I}} bias^2(I)$, where $\mathcal{I}$ is the set of the intervals in (10), $bias(I) = (1/M)\sum_{m=1}^{M} f(I,m) - f(I)$ and $f(I)$ gives the fraction of $\sqrt{n}(\hat{\theta}_n^m - \theta)$, $1 \le m \le M$, falling in the interval I; and as a global measure of the mean squared error of each distribution estimator we have considered $MS = 100 \times \sum_{I\in\mathcal{I}} \text{MSE}\,(I)$, where $\text{MSE}\,(I) = (1/M)\sum_{m=1}^{M} \{f(I,m) - f(I)\}^2$.

We repeated the above experiment for $n = 31, 41$ and also for samples from a exponential negative distribution with mean 1, Exp(1), and for samples from a standard Cauchy population, C(0,1). Since for all the considered populations $E\{\log(1 + |X_1|)\} < \infty$, by the theorem in Babu (1986), in all cases the ordinary bootstrap variance estimator is consistent. Nevertheless, as consistency does not guarantee asymptotic efficiency, we also repeated the above experiment for $n = 500, 1000$ to study empirically the asymptotic bias and the asymptotic mean squared error of the considered estimators. Tables 2 to 5 show the obtained results.

Note that we employed the same number of bootstrap replications to approximate the bootstrap variance estimators and to approximate the bootstrap distribution estimators. According to Efron and Tibshirani (1993, Chap.6), to approximate the bootstrap variance estimator, $B = 100$ usually gives quite satisfactory results. In fact, we repeated the whole experiment, where by whole experiment we mean for all the considered methods, for all the considered sample sizes and for all the considered populations generating the original samples, with $B = 100$ for the variance estimators and obtained almost the same results as for $B = 1000$, the differences being quite negligible.

Looking at Table 2 we see that Methods 1, 3, 7, 8 and 9 yield variance estimators with positive bias, while the rest of the methods underestimate the true variance. The sign of the bias remains constant for all the considered sample sizes ($n = 21, 31, 41$). With respect to the mean squared error, Methods 2 and 4 give, in general, the best results.

Table 2. Results for the variance estimators and $n = 21, 31, 41$.

		N(0,1)		Exp(1)		C(0,1)	
n	method	bias	mse	bias	mse	bias	mse
21	1	0.2650	1.3026	0.2965	1.0320	1.8436	20.6449
	2	-0.3673	0.8562	-0.1640	0.5358	-0.0822	5.7361
	3	0.1902	1.2260	0.2374	0.9547	1.4969	16.0646
	4	-0.5882	0.9033	-0.3213	0.4840	-0.6459	4.2185
	5	-0.8458	1.1001	-0.5021	0.5030	-1.2604	3.7511
	6	-0.8846	1.1013	-0.5024	0.5033	-1.2591	3.7606
	7	0.5878	1.3430	0.5354	1.2296	5.2311	134.2880
	8	0.4931	5.5071	0.4408	3.3981	1.4048	27.0004
	9	0.2639	1.3026	0.2965	1.0320	1.8139	19.4772
31	1	0.2569	1.0952	0.2469	0.7729	1.0129	7.4404
	2	-0.3031	0.7256	-0.1486	0.4352	-0.2709	3.2844
	3	0.3721	1.2687	0.3270	0.8998	1.2631	8.7490
	4	-0.5127	0.7647	-0.2943	0.4028	-0.7171	2.8869
	5	-0.7603	0.9311	-0.4642	0.4320	-1.2138	3.0376
	6	-0.7588	0.9318	-0.4638	0.4328	-1.2136	3.0438
	7	0.4218	0.8302	0.3719	0.6779	2.8621	24.4037
	8	0.5817	4.2666	0.4320	2.4865	1.1660	17.4788
	9	0.2515	1.0920	0.2455	0.7718	0.9860	7.3207
41	1	0.2629	0.9972	0.2086	0.6028	0.7827	4.9280
	2	-0.3071	0.6531	-0.1773	0.3539	-0.3895	2.4137
	3	0.3581	1.1296	0.2742	0.6880	0.9717	5.6365
	4	-0.4639	0.6811	-0.2817	0.3445	-0.6934	2.3066
	5	-0.7031	0.8299	-0.4409	0.3810	-1.1510	2.5726
	6	-0.7027	0.8286	-0.4405	0.3813	-1.1516	2.5765
	7	0.4000	0.6752	0.3279	0.5112	2.2113	12.5698
	8	0.1066	2.4930	0.0931	1.3473	0.2281	8.3937
	9	0.2559	0.9928	0.2067	0.6016	0.7589	4.8719

From Table 3 we observe that the bias of all methods, except for Method 9, has the same sign as for small samples. As expected from the theoretical results in Bloch and Gastwirth (1968) and Hall and Martin (1988), the estimator 7 has less mean squared error than the usual bootstrap estimator. In fact, the Bloch and Gastwirth estimator is the one having the smallest mean squared error. If we only pay attention to the RB variance estimators we see that, with respect to mean squared error, its behaviour is opposite to that for small samples ($n = 21, 31, 41$): for large samples Method 1 and, especially, Method 3 give the best results, while for small samples these methods have the largest mean squared errors.

Table 3. Results for the variance estimators and $n = 500, 1000$.

		N(0,1)		Exp(1)		C(0,1)	
n	method	bias	mse	bias	mse	bias	mse
500	1	0.0962	0.2696	0.0457	0.1155	0.1557	0.6941
	2	-0.5162	0.3995	-0.3416	0.1737	-0.8129	0.9989
	3	0.0935	0.2688	0.0432	0.1141	0.1553	0.6986
	4	-0.5609	0.4391	-0.3694	0.1904	-0.8822	1.0977
	5	-0.6021	0.4794	-0.3969	0.2076	-0.9488	1.1988
	6	-0.6030	0.4813	-0.3968	0.2077	-0.9494	1.2009
	7	0.1305	0.0765	0.0791	0.0398	0.4640	0.4355
	8	0.1171	0.7386	0.0512	0.2989	0.1753	1.8132
	9	-0.0372	0.2460	-0.0280	0.1079	-0.0579	0.6274
1000	1	0.0935	0.1996	0.0295	0.0811	0.1110	0.4775
	2	-0.5244	0.3706	-0.3577	0.1680	-0.8551	0.9597
	3	0.0748	0.1949	0.0189	0.0800	0.0817	0.4634
	4	-0.5569	0.4013	-0.3787	0.1814	-0.9065	1.0401
	5	-0.5881	0.4331	-0.3975	0.1944	-0.9547	1.1197
	6	-0.5887	0.4334	-0.3982	0.1949	-0.9553	1.1209
	7	0.1091	0.0458	0.0446	0.0199	0.3172	0.2134
	8	0.1085	0.5313	0.0379	0.2122	0.1134	1.2619
	9	-0.4235	0.3133	-0.2959	0.1427	-0.7013	0.8155

Table 4. Results for the distribution estimators and $n = 21, 31, 41$.

		N(0,1)		Exp(1)		C(0,1)	
n	method	BS	MS	BS	MS	BS	MS
21	1	3.1584	12.6967	3.1897	12.6334	3.2029	13.1130
	2	5.0037	15.7055	5.1364	15.7446	4.8162	15.8918
	3	3.1446	12.8307	3.1925	12.7790	3.1845	13.2508
	4	6.3032	17.5262	6.4806	17.6046	6.0365	17.6149
	5	8.9050	20.8116	9.1384	20.9136	8.5426	20.7728
	6	8.9192	20.8012	9.1491	20.9114	8.5453	20.7787
31	1	2.1169	10.4285	2.0780	10.2432	2.1494	10.8000
	2	3.1933	12.5231	3.2329	12.4241	3.0842	12.7299
	3	1.9334	10.0823	1.8859	9.8846	2.0014	10.5007
	4	3.9592	13.8185	4.0337	13.7489	3.8237	13.9941
	5	5.4450	16.0692	5.5847	16.0280	5.2604	16.1465
	6	5.4349	16.0627	5.5774	16.0204	5.2592	16.1565
41	1	1.5826	9.0571	1.7258	9.0343	1.6411	9.4364
	2	2.4478	10.9570	2.7112	11.0650	2.3942	11.1751
	3	1.4741	8.8117	1.5920	8.7607	1.5582	9.2241
	4	2.9024	11.7851	3.2119	11.9470	2.8270	11.9839
	5	3.9631	13.5869	4.3744	13.8416	3.8586	13.7260
	6	3.9485	13.5842	4.3548	13.8312	3.8593	13.7289

In summary, the mean squared error of the considered estimators is different for small and large samples. An example is estimator 7 that, although the asymptotic theory says it has a faster convergence rate than does estimator 1 (we can also appreciate this looking at Table 3), can have a quite unsatisfactory behaviour for small sample sizes (see the results in Table 2 for the Cauchy population). Another example involves the RB estimators: for small sample sizes, Methods 2 and 4 give, in general, the best results, while for large samples Method 1 and, especially, Method 3 have smaller mean squared error.

Table 5. Results for the distribution estimators and $n = 500, 1000$.

		N(0,1)		Exp(1)		C(0,1)	
n	method	BS	MS	BS	MS	BS	MS
500	1	0.0617	1.4213	0.0724	1.4718	0.0670	1.4651
	2	0.2348	1.9325	0.2806	2.0529	0.2421	1.9023
	3	0.0610	1.4170	0.0709	1.4667	0.0659	1.4599
	4	0.2655	2.0009	0.3169	2.1323	0.2755	1.9724
	5	0.2971	2.0754	0.3585	2.2142	0.3092	2.0421
	6	0.2984	2.0788	0.3571	2.2143	0.3091	2.0450
1000	1	0.0346	1.1340	0.0437	1.1961	0.0427	1.1489
	2	0.1762	1.5494	0.2361	1.6904	0.2070	1.5223
	3	0.0345	1.1388	0.0443	1.2002	0.0416	1.1458
	4	0.1950	1.5906	0.2610	1.7394	0.2306	1.5685
	5	0.2163	1.6370	0.2876	1.7871	0.2521	1.6094
	6	0.2160	1.6353	0.2890	1.7917	0.2522	1.6100

Looking at Tables 4 and 5 we see that Method 3 is the one that best approximates the distribution of Z_n, followed by the usual bootstrap, Method 2, Method 4 and Methods 5 and 6 (which have similar behaviours). Unlike the RB variance estimators, the ordering of the methods remains the same for all the considered sample sizes, for all the considered populations and for both BS and MS. This assertion can be seen better by looking at Figures 1 and 2, that for each sample size ($n = 21, 31, 41$ for Figure 1 and $n = 500, 1000$ for Figure 2) and for each considered population, display the values of BS. The graph for MS is quite similar.

The results in Tables 3 and 5 seem to support our initial suspicion that the choice of k_1 and k_2 affects the asymptotic efficiency of the corresponding estimators. Deeper theoretical studies, similar to that in Hall and Martin (1988) for the ordinary bootstrap estimator, and to that in Falk and Reiss (1989) for the usual bootstrap distribution estimator of Z_n, should be carried out to confirm this.

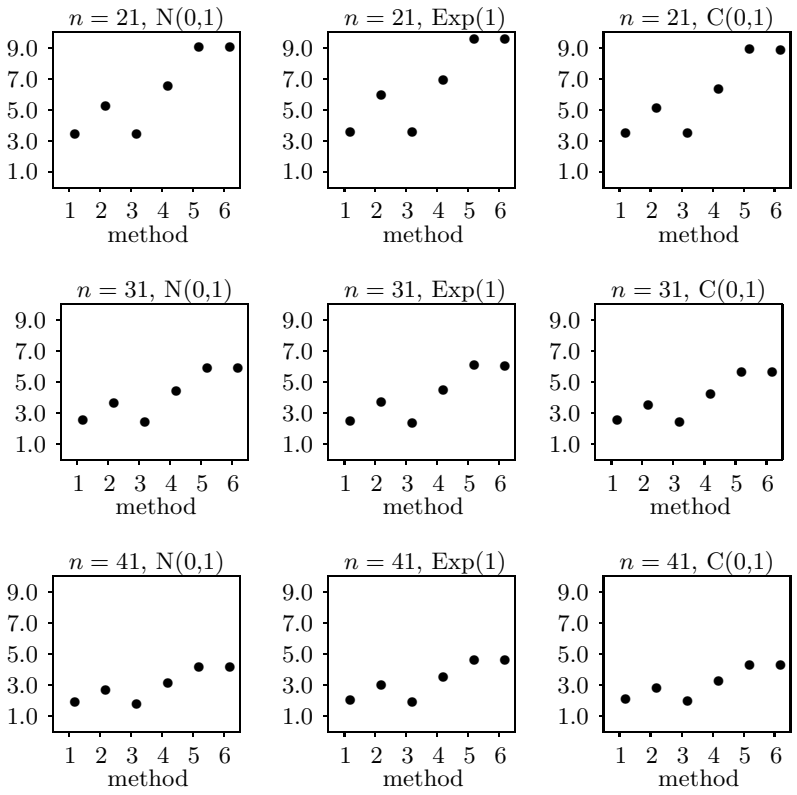

Figure 1. Horizontal axis: method, vertical axis: BS.

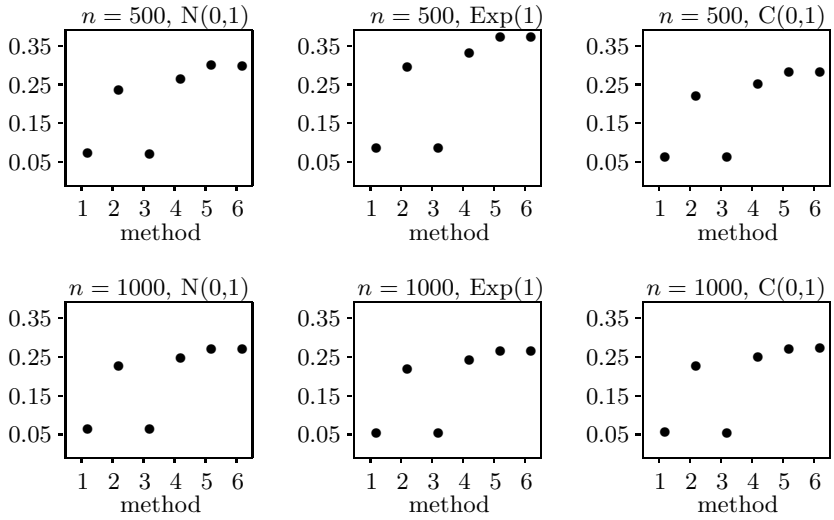

Figure 2. Horizontal axis: method, vertical axis: BS.

Acknowledgements

The authors are grateful to an associate editor for helpful suggestions and comments. This research was partially supported by Spanish MEC grant MTM 2004-0111433.

Appendix

Here we include some auxiliary lemmas used in the proofs of results in Sections 3, 4 and 5. First we introduce some notation. Let $\nu_{m,n} = \sum_{i=1}^{n} I(N_i > 0)$, with $(N_1, \ldots, N_n) \sim \mathcal{M}(m; 1/n, \ldots, 1/n)$. Note that $\nu_n = \nu_{n,n}$. Let ϕ and ϕ_0 denote the standard normal density function and the density of a normal law with mean zero and variance σ_0^2, respectively. Let $Y_1, Y_2, \ldots$ be a sequence of independent and identically distributed Poisson variables with mean 1, $S = \sum_{j=1}^{n} Y_j$ and $T = \sum_{j=1}^{n} I(Y_j > 0)$. With this notation we have that

$$P(k_1 \leq \nu_{m,n} \leq k_2) = \frac{P(S = m,\, k_1 \leq T \leq k_2)}{P(S = m)}. \tag{11}$$

Lemma 7.1. (a) *Uniformly in m and k, $nP(S = m, T = k) = \phi_{0,V}(u, v)\{1 + n^{-1/2}Q_1(u, v)\} + O(n^{-1})$, where $\phi_{0,V}$ is the bivariate normal density with zero mean and dispersion matrix*

$$V = \begin{pmatrix} 1 & q \\ q & pq \end{pmatrix},$$

$u = (m - n)n^{-1/2}$, $v = (k - np)n^{-1/2}$ and Q_1 is a third degree polynomial in u and v whose coefficients are bounded. In particular, if $m = n$, then uniformly in k, $nP(S = n, T = k) = \phi(0)\phi_0(v)\{1 + n^{-1/2}Q(v)\} + O(n^{-1})$, where $Q(v) = [(3q^3 - 3pq^2 + p^2q)/2\sigma_0^4]v - [(4q^3 - 4pq^2 + p^2q)/6\sigma_0^6]v^3$.

(b) *Uniformly in m, k_1 and k_2,*

$$\begin{aligned} &n^{1/2}P(S = n,\, k_1 \leq T \leq k_2) \\ &= \phi(0)\{\Phi(w_2) - \Phi(w_1)\} + n^{-1/2}\phi(0) \int_{w_1\sigma_0}^{w_2\sigma_0} Q(w)\phi_0(w)dw \\ &\quad -n^{-1/2}\phi(0)\left\{S_1(k_2)\phi_0(w_2\sigma_0) - S_1(k_1)\phi_0(w_1\sigma_0)\right\} + O(n^{-1}), \end{aligned}$$

where $S_1(x)$ is a periodic function of period 1, with $S_1(x) = x - 0.5$, if $0 \leq x < 1$.

Proof. Part (a) follows from Theorem 22.1 in Bhattacharya and Ranga Rao (1986). Part (b) follows from (a) and Theorem A.4.3 in Bhattacharya and Ranga Rao (1986).

Lemma 7.2. *If k_1 and k_2 satisfy one of* C.1, C.2, C.3, C.4, *then $(nP(k_1 \leq \nu_n \leq k_2))^{-1} = O(1)$.*

Proof. To prove the result we distinguish two cases: $|w_2 - w_1| \geq \varepsilon$, for some $\varepsilon > 0$, and $|w_2 - w_1| \to 0$. Suppose first that $|w_2 - w_1| \geq \varepsilon$, for some $\varepsilon > 0$. From (11), Lemma 7.1, and taking into account that $\sqrt{n}P(S = n) = \phi(0) + O(n^{-1/2})$, we get $P(k_1 \leq \nu_n \leq k_2) = \Phi(w_2) - \Phi(w_1) + O(n^{-1/2})$. Hence, if C.1 holds, then $(nP(k_1 \leq \nu_n \leq k_2))^{-1} = O(n^{-1})$. If k_1 and k_2 satisfy C.3 or C.4 and $w_1 \to \infty$ (analogously if $w_2 \to -\infty$), then

$$\Phi(w_2) - \Phi(w_1) \geq \Phi(w_1 + \varepsilon) - \Phi(w_1) = \phi(w_1)\varepsilon + O(\varepsilon^2) > \phi(\log^{1/2} n)\varepsilon + O(\varepsilon^2) \geq \frac{1}{n},$$

for all large n, and hence $(nP(k_1 \leq \nu_n \leq k_2))^{-1} = O(1)$. Now, suppose that $|w_2 - w_1| \to 0$. As before, by Lemma 7.1 and taking into account that $\sqrt{n}P(S = n) = \phi(0) + O(n^{-1/2})$, we obtain

$$P(k_1 \leq \nu_n \leq k_2) = \frac{\sum_{k=k_1}^{k_2} P(S = n, T = k)}{P(S = n)} = \frac{k_2 - k_1 + 1}{\sqrt{n}} \left\{\phi(w_1) + o(1)\right\}.$$

Therefore, if C.2 holds, $(nP(k_1 \leq \nu_n \leq k_2))^{-1} \leq (\sqrt{n}\{\phi(l) + o(1)\})^{-1} = O(n^{-1/2})$. If k_1 and k_2 satisfy C.3 or C.4 and $w_1 \to \infty$ (analogously if $w_2 \to -\infty$), then $\phi(w_1) > \phi(\log^{1/2} n)$ and hence $(nP(k_1 \leq \nu_n \leq k_2))^{-1} = O(1)$.

Lemma 7.3. *Let r be a fixed positive integer. If k_1 and k_2 satisfy one of* C.1, C.2, C.3, C.4, *then as $n \to \infty$, $P(k_1 \leq \nu_{n-r,n} \leq k_2)/P(k_1 \leq \nu_{n,n} \leq k_2) = 1 + o(1)$.*

Proof. From (11) we have that

$$\frac{P(k_1 \leq \nu_{n-r,n} \leq k_2)}{P(k_1 \leq \nu_{n,n} \leq k_2)} = \frac{P(S = n - r,\, k_1 \leq T \leq k_2)}{P(S = n,\, k_1 \leq T \leq k_2)} \frac{P(S = n)}{P(S = n - r)}. \tag{12}$$

Since r is fixed,

$$\frac{P(S = n)}{P(S = n - r)} = \frac{n^r}{n^{(r)}} = 1 + o(1). \tag{13}$$

For the first factor on the right side of (12) we apply Lemma 7.1. To do this, we distinguish two cases: $|w_2 - w_1| \to 0$, and $|w_2 - w_1| \geq \varepsilon$ for some $\varepsilon > 0$. If $|w_2 - w_1| \to 0$, we consider

$$\begin{aligned}
\frac{n}{k_2 - k_1 + 1} P(S = n - r, k_1 \leq T \leq k_2) &= \frac{1}{k_2 - k_1 + 1} \sum_{k=k_1}^{k_2} \phi_{0,V}(u_r, k_k), + O(n^{-1/2}), \\
\frac{n}{k_2 - k_1 + 1} P(S = n, k_1 \leq T \leq k_2) &= \frac{1}{k_2 - k_1 + 1} \sum_{k=k_1}^{k_2} \phi_{0,V}(0, v_k) + O(n^{-1/2}) \\
&= \frac{1}{k_2 - k_1 + 1} \phi(0) \sum_{k=k_1}^{k_2} \phi_0(v_k) + O(n^{-1/2}),
\end{aligned}$$

where $u_r = -rn^{-1/2}$ and $v_k = (k - np)n^{-1/2}$. If $|w_2 - w_1| \geq \varepsilon$ for some $\varepsilon > 0$, we consider

$$\begin{aligned} & n^{1/2}P(S = n - r, k_1 \leq T \leq k_2) \\ &= \int_{w_1\sigma_0}^{w_2\sigma_0} \phi_{0,V}(u_r, v)dv + O(n^{-1/2}) \\ &= \phi(u_r)\{\Phi(w_2 + qu_r\sigma_0^{-1}) - \Phi(w_1 + qu_r\sigma_0^{-1})\} + O(n^{-1/2}), \end{aligned}$$

and $n^{1/2}P(S = n, k_1 \leq T \leq k_2) = \phi(0)\{\Phi(w_2) - \Phi(w_1)\} + O(n^{-1/2})$. Hence,

$$\frac{P(S = n - r, k_1 \leq T \leq k_2)}{P(S = n, k_1 \leq T \leq k_2)} = 1 + O(t_n^{-1}), \tag{14}$$

where t_n is as defined in (3). From the assumptions on k_1 and k_2, the right side of (14) is $1 + o(1)$. This, together with (12) and (13), yield the result.

Lemma 7.4. *Let $Y_1, \ldots, Y_n$ be i.i.d. Poisson random variates with mean 1. If $E(|X_1|^3) < \infty$, then uniformly in x and k*

$$n^{1/2}P_*\Big(\sum_{i=1}^{n} D_iY_i \leq n^{1/2}x,\ S = n,\ T \leq k\Big) = \Phi(x)\phi(0)\Phi(v/\sigma_0) + O(n^{-1/2}),$$

$$nP_*\Big(\sum_{i=1}^{n} D_iY_i \leq n^{1/2}x,\ S = n,\ T = k\Big) = \Phi(x)\phi(0)\phi_0(v) + O(n^{-1/2}),$$

where $v = (k - np)n^{-1/2}$.

The proof of Lemma 7.4 follows along the lines of the proof of Proposition 1 in Babu, Pathak and Rao (1999), and so we omit it.

Lemma 7.5. *If k_1 and k_2 satisfy one of* C.1, C.2, C.3, C.4, *then $E_*([\sqrt{n}\{F_n^*(x) - F_n(x)\}]^4 \,/\, k_1 \leq \nu_n \leq k_2) = O(1)$.*

Proof. Let $\mu(k_1, k_2) = E_*\left([\sqrt{n}\{F_n^*(x) - F_n(x)\}]^4 \,/\, k_1 \leq \nu_n \leq k_2\right)$ and $a_i = I(X_i \leq x)$, $1 \leq i \leq n$. With this notation we have that

$$\mu(k_1, k_2) = \frac{1}{n^2}E_*\Big[\Big\{\sum_{i=1}^{n}(N_i - 1)a_i\Big\}^4 \,/\, k_1 \leq \nu_n \leq k_2\Big]. \tag{15}$$

Since

$$\begin{aligned} & E(\cdot \ / \ k_1 \leq \nu_n \leq k_2) \\ &= \frac{P(k_1 \leq \nu_n)}{P(k_1 \leq \nu_n \leq k_2)}E(\cdot \ / \ k_1 \leq \nu_n) - \frac{P(k_2 + 1 \leq \nu_n)}{P(k_1 \leq \nu_n \leq k_2)}E(\cdot \ / \ k_2 + 1 \leq \nu_n), \end{aligned}$$

from Corollary 7.1 in Jiménez Gamero, Muñoz García, Muñoz Reyes and Pino Mejías (1998) and (15), we get

$$\mu(k_1, k_2) = (1 - \pi_1)\mu(1, n) + O(\pi_2) + O(\pi_3/n), \tag{16}$$

$\forall x \in \mathbb{R}$, where

$$\pi_1 = 1 - \frac{P(k_1 \le \nu_{n-1,n} \le k_2)}{P(k_1 \le \nu_{n,n} \le k_2)}, \quad \pi_2 = 1 - \pi_1 - \frac{P(k_1 \le \nu_{n-2,n} \le k_2)}{P(k_1 \le \nu_{n,n} \le k_2)},$$
$$\pi_3 = 1 - \pi_1 - \pi_2 - \frac{P(k_1 \le \nu_{n-3,n} \le k_2)}{P(k_1 \le \nu_{n,n} \le k_2)}.$$

By Lemma 7.3, the right side of (16) is $O(1)$. This completes the proof.

Lemma 7.6. *Let c be a positive constant. If F satisfies* (4)*, then for all $t \in [1, c\log^{1/2} n]$ and all $1 \le k_1 \le k_2 \le n$ we have that*

(a) $$P_*\left\{\sqrt{n}(\theta_n^* - \theta_n) > t \,/\, k_1 \le \nu_n \le k_2\right\}$$
$$\le P_*\{F_n^*(\theta_n + t/\sqrt{n}) - F_n(\theta_n + t/\sqrt{n}) \le -\varepsilon t/\sqrt{n} \,/\, k_1 \le \nu_n \le k_2\},$$

(b) $$P_*\left\{\sqrt{n}(\theta_n - \theta_n^*) > t \,/\, k_1 \le \nu_n \le k_2\right\}$$
$$\le P_*\{F_n(\theta_n - t/\sqrt{n}) - F_n^*(\theta_n - t/\sqrt{n}) \le -\varepsilon t/\sqrt{n} \,/\, k_1 \le \nu_n \le k_2\},$$

for some $\varepsilon > 0$, for all $n \ge n_0$, for some $n_o \in \mathbb{N}$.

Proof. Inequality (a) is proved in the proof of Theorem 1 in Ghosh, Parr, Singh and Babu (1984). Inequality (b) can be derived similarly.

References

Babu, G. J. (1986). A note on bootstrapping the variance of sample quantiles. *Ann. Inst. Statist. Math.* **38**, 439-443.

Babu, G. J., Pathak, P. K. and Rao, C. R. (1999). Second-order correctness of the Poisson bootstrap. *Ann. Statist.* **27**, 1666-1683.

Bhattacharya, R. N. and Ranga Rao, R. (1986). *Normal Approximation and Asymptotic Expansions*, Krieger, Malabar, Florida.

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

Bloch, D. A. and Gastwirth, J. L. (1968). On a simple estimate of the reciprocal of the density function. *Ann. Math. Statist.* **39**, 1083-1085.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall.

Falk, M. and Reiss, R. D.(1989). Weak convergence of smoothed and nonsmoothed bootstrap quantile estimates. *Ann. Probab.* **17**, 362-371.

Ghosh, M., Parr, W. C., Singh, K. and Babu, G. J. (1984). A note on bootstrapping the sample median. *Ann. Statist.* **12**, 1130-1135.

Hall, P. and Martin, M. C. (1988). Exact convergence rate of bootstrap quantile variance estimation. *Probab. Theory Related Fields* **80**, 261-268.

Jiménez-Gamero, M. D., Muñoz-García, J. and Muñoz-Reyes, A. (1998). Bootstrapping the sample median. *Comm. Statist. Theory Method* **27**, 1979-1990.

Jiménez-Gamero, M. D., Muñoz-García, J., Muñoz-Reyes, A. and Pino-Mejías, R. (1998). On Efron's method II with identification of outlier bootstrap samples. *Comput. Statist.* **13**, 301-318.

Johnson, N. L. and Kotz, S. (1977). *Urn Models and Their Application*. Wiley, New York.

Muñoz-García, J., Pino-Mejías, R., Muñoz-Pichardo, J. and Cubiles-de-la-Vega, M.D. (1997). Identification of Outlier Bootstrap Samples. *J. Appl. Stat.* **24**, 333-342.

Rao, C. R., Pathak, P. K. and Koltchinskii, V. I. (1997). Bootstrap by sequential resampling. *J. Statist. Plann. Inference* **64**, 257-281.

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

Shao, J. (1990). Bootstrap estimation of the asymptotic variance of statistical functionals. *Ann. Inst. Statist. Math.* **42**, 737-752.

Shao, J. (1992). Bootstrap variance estimators with truncation. *Statist. Probab. Lett.* **15**, 95–101.

Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9**, 1187-1195.

Stromberg, A. J. (1997). Robust covariance estimates based on resampling. *J. Statist. Plann. Inference* **57**, 321-334.

Dpto. de Estadística e Investigación Operativa, Facultad de Matemáticas, C/ Tarfia s.n., 41.012 Sevilla, Spain.

E-mail: dolores@us.es

Dpto. de Estadística e Investigación Operativa, Facultad de Matemáticas, C/ Tarfia s.n., 41.012 Sevilla, Spain.

E-mail: joaquinm@us.es

Dpto. de Estadística e Investigación Operativa, Facultad de Matemáticas, C/ Tarfia s.n., 41.012 Sevilla, Spain.

E-mail: rafaelp@us.es

(Received April 2003; accepted December 2003)