A Monte Carlo comparison of three consistent bootstrap procedures

R. Pino Mejías[*†], M.D. Jiménez Gamero[†], A. Enguix González[†]

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

Since bootstrap samples are simple random samples with replacement from the original sample, the information content of some bootstrap samples can be very low. To avoid this fact, some authors have proposed several variants of the classical bootstrap. In this paper we consider two of them: the sequential or Poisson bootstrap and the reduced bootstrap. Both of them, like ordinary bootstrap, can yield second order accurate distribution estimators, that is, the three bootstrap procedures are asymptotically equivalent. The question that naturally arises is which of them should be used in a practical situation, in other words, which of them should be used for finite sample sizes. To try to answer this question, we have carried out a simulation study. Although no method was found to exhibit best performance in all the considered situations, some recommendations are given.

1 Introduction

The bootstrap, introduced by Efron [1], is a powerful tool for nonparametric estimation of sampling distributions and standard errors. It may be described as follows. Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ 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 $X_1, X_2, \ldots, X_n$ and let $\mathbf{X}^* = (X_1^*, X_2^*, \ldots, X_n^*)$ 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 $X_1, X_2, \ldots, X_n$.

Bootstrap samples are simple random samples of size n selected with replacement from the original sample. Thus, all bootstrap samples are not equally informative, when the information content of $\mathbf{X}^*$ is measured through the number of distinct original observations in it, ν_n. As Rao *et al.* [2] assert, this

---

*Corresponding author. E-mail: rafaelp@us.es

variability is neither necessary nor desirable. To avoid it, these authors have introduced a sequential sampling method that keeps constant the information content of bootstrap samples, the sequential or Poisson bootstrap (originally, they named it sequential bootstrap; the name Poisson bootstrap was given latter by Babu *et al.* [3]). It consists of sampling sequentially (with replacement from the original sample) until m distinct original observations appear. To keep the information constant, this method introduces another source of variability: the size of the sequential bootstrap samples.

To reduce the variability of ν_n, keeping the size of bootstrap samples fixed, Jiménez-Gamero *et al.* [4] have introduced a modified bootstrap that consists of only considering those bootstrap samples satisfying $k_1 \leq \nu_n \leq k_2$, for some $1 \leq k_1 \leq k_2 \leq n$. They call it the reduced bootstrap, since it only uses a portion of the set of all possible bootstrap samples. The method is an extension of the one introduced by Muñoz-García *et al.* [5], that takes $k_2 = n$. Note that ordinary bootstrap is a particular case of reduced bootstrap with $k_1 = 1$ and $k_2 = n$.

Another method to reduce the variability of ν_n, keeping the size of bootstrap samples fixed, can be found in Stromberg [6]. In order to reduce the breakdown point of the bootstrap variance estimator, which is $1/n$ regardless of the breakdown point of the estimator due to the fact that $P(\nu_n = n) > 0$, Stromberg [6] proposed to use a "limited replacement bootstrap", in which all bootstrap samples selecting the same element more than m^* times are excluded, that is, it only considers those bootstrap samples with $N_i \leq m^*$, $1 \leq i \leq n$, where $N_i = card\{j : X_j^* = X_i\}$, $1 \leq i \leq n$. As before, note that ordinary bootstrap is a particular case of limited replacement bootstrap with $m^* = n$.

Singh [7] has shown that the bootstrap estimator of the distribution of the sample mean is consistent and that it can be more accurate than the approximation given by the central limit theorem when higher-order population moments exist. Babu and Singh [8] have extended these results to functions of multivariate means and their studentized versions. Babu *et al.* [3] have shown that they are also true for the sequential bootstrap when $m = E(\nu_n)$. Jiménez-Gamero *et al.* [9] have also proven them for the reduced bootstrap, for certain choices of k_1 and k_2. Therefore, the results in Babu *et al.* [3] and Jiménez-Gamero *et al.* [9] reveal that the usual bootstrap high variability of ν_n is not necessary. To our knowledge there is no similar result for the limited replacement bootstrap of Stromberg [6] so, from now on we will not consider this method.

In summary, the above cited results tell us that we have three methods to consistently estimate the distribution function of statistics in an important class (the functions of multivariate means and their studentized versions), that is, all of them behave similarly for large samples. The question that naturally

arises is which of them should be used in a practical situation, that is, for finite sample sizes. In order to try to answer this question, we have carried out a simulation study. Although it is not exhaustive (it is impossible to simulate all cases), we can conclude some general recommendations from the obtained results.

The paper is organized as follows. In Section 2 we briefly review the methods to be compared. In Section 3 we describe the simulations and display the obtained results. Section 4 concludes giving some practical recommendations in the light of the obtained results.

2   The methods

In this Section we first study the variability reduction properties of the considered methods, and then we review some properties of the corresponding distribution estimators. Here we only review the second order correctness properties because they just make the considered distribution estimators to have an edge over the approximation by the limiting normal distribution.

2.1   *Variability reduction properties*

As in the Introduction, let ν_n denote the number of distinct observations in a bootstrap sample. Then under ordinary bootstrap,

$$E(\nu_n) = n \left[1 - \left(1 - \frac{1}{n} \right)^n \right] \simeq 0.632n,$$

$$var(\nu_n) = n \left(1 - \frac{1}{n} \right)^n + n(n-1) \left(1 - \frac{2}{n} \right)^n - n^2 \left(1 - \frac{1}{n} \right)^{2n} \simeq 0.233n,$$

To reduce the variability of ν_n, the sequential bootstrap draws observations from the original sample with replacement until there are $m \simeq E(\nu_n)$ distinct observations in the bootstrap sample. This way, the number of distinct observations is no longer a random variable. Nevertheless, the method yields bootstrap samples with random size N. The expectation and the variance of N are (see Rao *et al.* [2])

$$E(N) = n \left(\frac{1}{n} + \frac{1}{n-1} + ... + \frac{1}{n-m+1} \right) = n + O(1),$$

$$var(N) = \sum_{j=1}^{m} \frac{n(j-1)}{(n-j+1)^2} = n(e-1) + O(1).$$

Next result gives the expectation and the variance of ν_n for the reduced bootstrap (see Appendix C for a proof).

PROPOSITION 2.1 *Let $\mathbf{X}^*$ be a bootstrap sample, $\nu_n = \nu_n(\mathbf{X}^*)$ and let k_1, k_2 be two positive integers with $1 \le k_1 \le k_2 \le n$. Then,*

$$E\left(\nu_n \mid k_1 \le \nu_n \le k_2\right) = k_2 - (k_2 - k_1)\left(1 - \frac{1}{k_2}\right)^{n-k_1}$$

and

$$\begin{aligned}
var\left(\nu_n \mid k_1 \le \nu_n \le k_2\right) &= (k_2 - k_1)(k_2 - k_1 - 1)\left(1 - \frac{2}{k_2}\right)^{n-k_1} + \\
&\quad + (k_2 - k_1)\left(1 - \frac{1}{k_2}\right)^{n-k_1} - (k_2 - k_1)^2\left(1 - \frac{1}{k_2}\right)^{2(n-k_1)}.
\end{aligned} \tag{1}$$

From (1) we have that if $k_1 = k_1(n)$ and $k_2 = k_2(n)$ are such that $(n-k_1)/k_2$ converges to a real limit $l \ne 0$, as $n \to \infty$, then $var\left(\nu_n \mid k_1 \le \nu_n \le k_2\right) = O(k_2 - k_1)$. Hence, as $var(\nu_n) \simeq 0.233n$, if $k_2 - k_1 = o(n)$, the variability of ν_n in the reduced bootstrap will be less than in the usual bootstrap. If $(n - k_1)/k_2 \to \infty$, as $n \to \infty$, this implies that $k_2 = o(n)$. As it is shown in Jiménez-Gamero *et al.* [9], in this case the reduced bootstrap is not consistent and thus this choice of k_1 and k_2 is not advisable. A similar situation occurs when $(n - k_1)/k_2 \to 0$, as $n \to \infty$. Hence, for adequate choices of k_1 and k_2, the reduced bootstrap variability of ν_n is less than that of ordinary bootstrap.

2.2 *Second order accuracy of the distribution estimators*

Let $X_1, X_2, ..., X_n$ be i.i.d. random variables having common distribution function F, mean μ and variance σ^2. Let

$$\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j, \quad s_n^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2 \text{ and } \bar{X}_n^* = \frac{1}{n}\sum_{j=1}^{n} X_j^*.$$

Throughout this paper P_* will denote the bootstrap conditional probability law, given $X_1, X_2, \ldots, X_n$.

If $E|X_1|^3 < \infty$ and F satisfies Cramér's condition, Theorem 1.D in Singh [7] shows that

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

This implies that the ordinary bootstrap approximation to $P\{n^{1/2}(\bar{X}_n - \mu)/\sigma \leq x\}$ is better than the normal approximation, which has an error of order $O(n^{-1/2})$. Theorem 3.1 in Babu *et al.* [3] gives the analogue of the above result for the Poisson bootstrap, with $m = E(\nu_n)$. Corollary 1 in Jiménez-Gamero *et al.* [9] states conditions on k_1 and k_2 $(1 \leq k_1 \leq k_2 \leq n)$ for (2) to hold for the reduced bootstrap distribution estimator. These conditions are the following:

CONDITION C.1.a $\Phi(w_2) - \Phi(w_1) \geq \alpha$, $\forall n \geq n_0$, for some $n_0 \in \mathbb{N}$ and some fixed constant $\alpha > 0$ and $w_1 + w_2 = o(1)$ or $w_1 \to -\infty$ and $w_2 \to +\infty$,

CONDITION C.2.a w_1, $w_2 \to 0$,
where $w_1 = (k_1 - 1 - np)n^{-1/2}\sigma_0^{-1}$, $w_2 = (k_2 - np)n^{-1/2}\sigma_0^{-1}$, $\sigma_0{}^2 = pq - q^2$, $p = 1 - e^{-1}$, $q = 1 - p$ and Φ denotes the standard normal cumulative distribution function. Since $P(\nu_n \leq k) = \Phi(w) + o(1)$ with $w = (k - np)/\sqrt{n}\sigma_0$ (Johnson and Kotz [10, p. 318]), if k_1 and k_2 are such that some of the above conditions hold, then the proportion of bootstrap samples not to be considered because they have a small number of different elements $(\nu_n < k_1)$ is (almost) the same as the proportion of bootstrap samples not to be considered because they have a big number of different elements $(\nu_n > k_2)$.

The above results can be extended to the multivariate case and to statistics which can be expressed as smooth functions of multivariate means. Let $X_1, X_2, ..., X_n$ be n i.i.d. d-dimensional random vectors with common distribution function F, mean μ and variance matrix Σ, where d is a fixed positive integer. Let $\bar{X}_n$ be the sample mean and let Σ_n be the sample variance matrix. Let H be a real valued Borel measurable function on $\mathbb{R}^d$. If H is differentiable at y, we denote by $h(y)$ the vector of first order partial derivatives of H at y. Let $\tau^2 = h(\mu)'\Sigma h(\mu)$ and $\tau_n^2 = h(\bar{X}_n)'\Sigma_n h(\bar{X}_n)$, where the prime denotes transpose.

If $E\|X_1\|^3 < \infty$, F satisfies Cramér's condition, H is three times continuously differentiable in a neighborhood of μ, $h(\mu) \neq 0$ and $\tau > 0$, Corollary 2 in Babu and Singh [8] shows that

$$
\sup_x n^{1/2} \left| P_* \left[n^{1/2} \left\{ H(\bar{X}_n^*) - H(\bar{X}_n) \right\} / \tau_n \leq x \right] - \right.
$$
$$
\left. - P \left[n^{1/2} \left\{ H(\bar{X}_n) - H(\mu) \right\} / \tau \leq x \right] \right| = o(1). \tag{3}
$$

Theorem 4.1 in Babu *et al.* [3] and Theorem 4(b) in Jiménez-Gamero *et al.*
[9] give the analogues of the above result for the Poisson bootstrap (with
$m = E(\nu_n)$) and the reduced bootstrap (with k_1 and k_2 satisfying conditions
C.1.a and C.2.a), respectively.

A class of statistics frequently used as pivots or approximate pivots is the
studentized version of the above considered statistics. The result in (3) is
also true for this class of statistics. Let $\tau_n^{*2} = h(\bar{X}_n^*)'\Sigma_n^* h(\bar{X}_n^*)$, where Σ_n^*
is the sample variance matrix of the bootstrap sample. For a d-dimensional
random vector $X' = (X_{(1)}, X_{(2)}, ..., X_{(d)})$, let W be the $d(d+3)/2$-vector ob-
tained from X by adjoining it the products $X_{(i)}X_{(j)}$, $1 \leq i \leq j \leq d$, that is,
$W' = (X_{(1)}, X_{(2)}, ..., X_{(d)}, X_{(1)}^2, X_{(1)}X_{(2)}, ..., X_{(1)}X_{(d)}, X_{(2)}^2, X_{(2)}X_{(3)}, ..., X_{(d)}^2)$.
If $E\|W_1\|^3 < \infty$, F satisfies Cramér's condition, H is three times continuously
differentiable in a neighborhood of μ, $h(\mu) \neq 0$ and $\tau > 0$, then Theorem 4 in
Babu and Singh [8] shows that

$$
\begin{aligned}
\sup_x n^{-1/2} \Big| P_* \Big[n^{1/2} & \left\{ H(\bar{X}_n^*) - H(\bar{X}_n) \right\} / \tau_n^* \leq x \Big] - \\
& - P \Big[n^{1/2} \left\{ H(\bar{X}_n) - H(\mu) \right\} / \tau_n \leq x \Big] \Big| = o(1).
\end{aligned}
\tag{4}
$$

Theorem 4.3 in Babu *et al.* [3] and Theorem 5(b) in Jiménez-Gamero *et al.*
[9] give the analogues of the result in (4) for the Poisson bootstrap (with
$m = E(\nu_n)$) and the reduced bootstrap (with k_1 and k_2 satisfying conditions
C.1.a and C.2.a), respectively.

As we said in the Introduction, Rao *et al.* [2] asserted that the variability of
ν_n is not desirable. An example of this fact emerges when one considers statis-
tics which are a studentized version of a smooth function of sample means, as
the ones involved in (4), since bootstrap samples with all their components
equal yield a degenerate value of the considered statistics. This undesirable
situation may be avoided by using the sequential bootstrap or the reduced
bootstrap with some $k_1 > 1$.

3   Simulations

In Section 2 we have seen that the ordinary bootstrap, the sequential bootstrap
with $m = E(\nu_n)$ and the reduced bootstrap for adequate choices of k_1 and k_2
yield distribution estimators having all of them the same asymptotic accuracy.
To study and compare the finite sample performance of the corresponding esti-
mators we have carried out a simulation experiment. The considered methods
are displayed in Table 1, where $[x]$ denotes the greatest integer less or equal
than x. The Appendices A and B describe the employed algorithms to gen-
erate reduced bootstrap samples and Poisson bootstrap samples, respectively.

Table 1. Methods.

method	description
1	ordinary bootstrap
2	reduced bootstrap with $k_1 = [np - 2\sqrt{npq}] + 1$, $k_2 = [np + 2\sqrt{npq}]$
3	reduced bootstrap with $k_1 = [np - \sqrt{npq}] + 1$, $k_2 = [np + \sqrt{npq}]$
4	reduced bootstrap with $k_1 = k_2 = [np] + 1$
5	sequential bootstrap with $m = [np] + 1$

To study the corresponding distribution estimators of the statistic

$$t_1(\mathbf{X}) = \sqrt{n}\frac{\bar{X}_n - \mu}{s_n},$$

we have generated $M = 1000$ samples of size $n = 10$ from a standard normal population, N(0,1). For each method in Table 1 and from each sample $\mathbf{X}^m$, $1 \leq m \leq M$, we have generated $B = 1000$ bootstrap samples, $\mathbf{X}^{*m,1}, ..., \mathbf{X}^{*m,B}$. We have considered the following intervals

$$(-\infty, -3), \ [-3, -2.9), \ [-2.9, -2.8), \ [-2.8, -2.7), \ ..., \ [2.9, 3), \ [3, \infty) \quad (5)$$

and, for each method, we have calculated $f(I, m)$, the fraction of $t_1(\mathbf{X}^{*m,b})$, $1 \leq b \leq B$, falling on interval I. As a global measure of the bias of each distribution estimator we have considered

$$BS = \sum_{I \in \mathcal{I}} bias^2(I),$$

where $\mathcal{I}$ is the set of the intervals in (5), $bias(I) = \frac{1}{M}\sum_{m=1}^{M} f(I, m) - f(I)$ and $f(I)$ gives the fraction of $T(\mathbf{X}^m)$, $1 \leq m \leq M$, falling on the interval I; and as a global measure of the mean squared error of each distribution estimator we have considered

$$MS = \sum_{I \in \mathcal{I}} mse(I),$$

where $mse(I) = \frac{1}{M}\sum_{m=1}^{M}\{f(I, m) - f(I)\}^2$.

We have repeated the above experiment for $n = 20, 50, 100, 200, 500, 1000$ and also for the statistics

$$t(\mathbf{X}) = \sqrt{n}\frac{H(\bar{X}_n) - H(\mu)}{|h(\bar{X}_n)|s_n},$$

with $H(x) = \exp(x)$ (statistic t_2) and $H(x) = x^2$ (statistic t_3). For statistic t_3,

Table 2.   Results for statistic t_1 with data from a $N(0, 1)$ population.

method		10	20	50	100	200	500	1000
					sample size			
1	$BS \times 10^3$	0.61	0.98	0.63	1.26	1.11	0.66	1.05
	$MS \times 10^3$	2.07	2.00	1.61	2.24	2.09	1.64	2.01
2	$BS \times 10^3$	0.54	0.96	0.64	1.28	1.10	0.66	1.06
	$MS \times 10^3$	2.07	2.00	1.63	2.23	2.07	1.64	2.04
3	$BS \times 10^3$	0.53	0.96	0.64	1.28	1.10	0.65	1.03
	$MS \times 10^3$	2.07	2.00	1.61	2.25	2.07	1.63	2.01
4	$BS \times 10^3$	8.44	3.69	2.58	3.34	2.85	2.30	2.59
	$MS \times 10^3$	10.96	4.68	3.56	4.30	3.80	3.26	3.55
5	$BS \times 10^3$	0.64	0.98	0.63	1.26	1.12	0.66	1.06
	$MS \times 10^3$	1.94	2.00	1.60	2.23	2.10	1.65	2.03

Table 3.   Results for statistic t_2 with data from a $N(0, 1)$ population.

method		10	20	50	100	200	500	1000
					sample size			
1	$BS \times 10^3$	1.00	0.92	0.99	0.92	0.99	1.26	0.53
	$MS \times 10^3$	2.98	2.02	1.98	1.90	1.98	2.23	1.51
2	$BS \times 10^3$	1.17	0.85	0.99	0.92	0.97	1.24	0.54
	$MS \times 10^3$	3.29	2.01	1.99	1.90	1.94	2.21	1.51
3	$BS \times 10^3$	1.19	0.85	0.99	0.91	0.98	1.27	0.53
	$MS \times 10^3$	3.34	2.00	2.00	1.88	1.95	2.23	1.51
4	$BS \times 10^3$	9.62	6.59	3.67	3.92	3.36	2.21	2.21
	$MS \times 10^3$	12.53	7.61	4.65	4.87	4.31	3.17	3.18
5	$BS \times 10^3$	0.66	0.99	0.98	0.94	0.98	1.24	0.53
	$MS \times 10^3$	2.50	2.10	1.98	1.91	1.96	2.21	1.49

Table 4.   Results for statistic t_3 with data from a $N(5, 1)$ population.

method		10	20	50	100	200	500	1000
					sample size			
1	$BS \times 10^3$	1.02	0.63	1.45	0.69	0.86	0.84	0.92
	$MS \times 10^3$	2.52	1.66	2.42	1.65	1.83	1.82	1.89
2	$BS \times 10^3$	1.19	0.65	1.42	0.70	0.84	0.82	0.90
	$MS \times 10^3$	2.76	1.68	2.40	1.67	1.83	1.79	1.88
3	$BS \times 10^3$	1.17	0.65	1.45	0.71	0.86	0.84	0.89
	$MS \times 10^3$	2.73	1.69	2.41	1.68	1.84	1.82	1.86
4	$BS \times 10^3$	6.39	3.40	3.70	2.30	2.79	2.18	2.15
	$MS \times 10^3$	8.93	4.38	4.67	3.27	3.75	3.14	3.11
5	$BS \times 10^3$	0.96	0.62	1.45	0.69	0.83	0.82	0.89
	$MS \times 10^3$	2.27	1.65	2.42	1.66	1.79	1.80	1.87

to avoid $h(\mu) = 0$, the data comes from a N(5, 1) population. The obtained results are displayed in Tables 2, 3 and 4. Tables 5, 6 and 7 present the results of a similar experiment with data from a mixture of two normal populations, $0.2N(0, 1) + 0.8N(5, 2^2)$.

All computations in this paper have been performed using programs written in the R language (R Development Core Team [11]).

Table 5. Results for statistic t_1 with data from a $0.2N(0,1) + 0.8N(5,2^2)$ population.

method		sample size						
		10	20	50	100	200	500	1000
1	$BS \times 10^3$	1.05	0.83	0.99	1.18	0.65	1.06	1.24
	$MS \times 10^3$	2.51	1.83	1.96	2.16	1.62	2.02	2.21
2	$BS \times 10^3$	0.91	0.86	0.92	1.17	0.66	1.02	1.22
	$MS \times 10^3$	2.46	1.87	1.90	2.14	1.64	1.98	2.19
3	$BS \times 10^3$	0.93	0.86	0.93	1.16	0.67	1.03	1.24
	$MS \times 10^3$	2.47	1.87	1.92	2.12	1.63	2.01	2.21
4	$BS \times 10^3$	8.95	3.22	3.80	3.47	2.05	3.12	3.10
	$MS \times 10^3$	11.71	4.21	4.78	4.43	3.02	4.09	4.05
5	$BS \times 10^3$	1.06	0.82	0.98	1.20	0.64	1.02	1.25
	$MS \times 10^3$	2.34	1.83	1.94	2.18	1.61	2.00	2.22

Table 6. Results for statistic t_2 with data from a $0.2N(0,1) + 0.8N(5,2^2)$ population.

method		sample size						
		10	20	50	100	200	500	1000
1	$BS \times 10^3$	0.67	0.92	0.72	0.99	1.00	1.11	0.74
	$MS \times 10^3$	2.75	2.17	1.73	1.97	1.98	2.07	1.71
2	$BS \times 10^3$	0.75	1.02	0.80	1.08	1.00	1.09	0.72
	$MS \times 10^3$	3.04	2.31	1.82	2.06	1.95	2.05	1.69
3	$BS \times 10^3$	0.78	0.98	0.82	1.07	1.00	1.10	0.73
	$MS \times 10^3$	3.00	2.28	1.83	2.04	1.97	2.06	1.71
4	$BS \times 10^3$	16.18	6.77	3.54	3.15	3.44	3.38	2.24
	$MS \times 10^3$	19.51	7.89	4.54	4.11	4.39	4.36	3.21
5	$BS \times 10^3$	0.95	0.90	0.72	0.98	1.02	1.09	0.73
	$MS \times 10^3$	2.88	2.14	1.74	1.97	1.98	2.06	1.71

Table 7. Results for statistic t_3 with data from a $0.2N(0,1) + 0.8N(5,2^2)$ population.

method		sample size						
		10	20	50	100	200	500	1000
1	$BS \times 10^3$	1.03	0.60	0.96	0.79	1.08	1.37	0.80
	$MS \times 10^3$	2.87	1.71	1.96	1.76	2.03	2.35	1.75
2	$BS \times 10^3$	1.07	0.61	0.98	0.79	1.07	1.34	0.81
	$MS \times 10^3$	3.03	1.74	1.97	1.76	2.04	2.32	1.77
3	$BS \times 10^3$	1.07	0.62	0.97	0.78	1.07	1.33	0.81
	$MS \times 10^3$	3.03	1.73	1.96	1.74	2.05	2.30	1.77
4	$BS \times 10^3$	11.66	4.19	3.51	3.34	3.05	4.01	3.07
	$MS \times 10^3$	14.65	5.19	4.48	4.30	4.01	4.97	4.03
5	$BS \times 10^3$	1.44	0.55	0.97	0.81	1.05	1.36	0.81
	$MS \times 10^3$	3.11	1.65	1.97	1.79	2.01	2.34	1.78

4   Conclusions

Looking at Tables 2–7 we see that method 4 has the biggest bias for all the considered cases. For the rest of the methods in Table 1, we observe that there is no method behaving much better than the others for all statistics, all sample sizes and all distributions. As expected, the differences in BS and MS decrease as the sample size increases. Although the performance of methods 1, 2, 3, 5

and 6 is very similar, some recommendations can be given: use sequential bootstrap for small sample sizes and use methods 2 and 3 for larger samples.

References

[1] Efron, B., 1979, Bootstrap Methods: Another look at the Jackknife. *Annals of Statistics,* **7**, 1–26.
[2] Rao, C.R., Pathak, P.K. and Koltchinskii, V.I., 1997, Bootstrap by sequential resampling. *Journal of Statistical Planning and Inference,* **64**, 257–281.
[3] Babu, G.J., Pathak, G.J. and Rao, G.J., 1999, Second-order correctness of the Poisson bootstrap. *Annals of Statistics,* **27**, 1666–1683.
[4] Jiménez-Gamero, M.D., Muñoz-García, J. and Pino-Mejías, R., 2004, Reduced bootstrap for the median. *Statistica Sinica,* **14**, 1179-1198.
[5] Muñoz-García, J., Pino-Mejías, R., Muñoz-Pichardo, J. and Cubiles-de-la-Vega, M.D., 1997, Identification of Outlier Bootstrap Samples. *Journal of Applied Statistics,* **24**, 333–342.
[6] Stromberg, A.J., 1997, Robust covariance estimates based on resampling. *Journal of Statistical Planning and Inference,* **57**, 321–334.
[7] Singh, K., 1981, On the asymptotic accuracy of Efron's bootstrap. *Annals of Statistics,* **9**, 1187–1195.
[8] Babu, G.J. and Singh, K., 1984, On the term Edgeworth correction by Efron's bootstrap. *Sankhyā Series A,* **46**, 219–232.
[9] Jiménez-Gamero, M.D., Muñoz-García, J. and Cubiles-de-la-Vega, M.D., 2005. Consistency of the reduced bootstrap for sample means, *Statistics and Probability Letters,* to appear.
[10] Johnson, N.L. and Kotz, S., 1977, *Urn Models and Their Application* (Wiley).
[11] R Development Core Team, 2005, R: A language and environment for statistical computing, Vienna: R Foundation for Statistical Computing. Available on line at: www.r-project.org (accessed 4 November 2005).
[12] Smyth G., 2003, Rtrunpois: random number generation for the truncated Poisson distribution, S-archive. Available on line at: www.statsci.org/s/trunpois.html (accessed 4 November 2005).

Appendix A: Algorithm 1

We have followed the next steps to generate reduced bootstrap samples:

1. Select a random sample of size k_2 without replacement from $\{1, 2, ..., n\}$, say $\mathcal{I}_1$.
2. Select a random sample of size k_1 without replacement from $\mathcal{I}_1$, say $\mathcal{I}_2 = \{i_1, i_2, ..., i_{k_1}\}$.
3. Select a random sample of size $n - k_1$ with replacement from $\mathcal{I}_1$, say $\mathcal{I}_3 = (j_1, j_2, ..., j_{n-k_1})$.
4. Let $(l_1, l_2, ..., l_n)$ be an n-vector whose components are obtained by randomly permuting the string $(i_1, i_2, ..., i_{k_1}, j_1, j_2, ..., j_{n-k_1})$.
5. The n-vector $(X_{l_1}, X_{l_2}, ..., X_{l_n})$ is a bootstrap sample satisfying $k_1 \leq \nu_n \leq k_2$.

Appendix B: Algorithm 2

For the poisson bootstrap, next algorithm has been used:

1 Select a random sample of size m without replacement from $\{1, 2, ..., n\}$, say $\mathcal{I} = \{i_1, i_2, ..., i_m\}$.
2 Generate a random sample of size m from a truncated Poisson with origin equal to 1, say $V = (V_1, V_2, ..., V_m)$.
3 The Poisson bootstrap sample is defined by selecting V_j times each position of the original sample i_j, appearing in the set $\mathcal{I}$.

The random number generation for the truncated Poisson distribution has been realized by the function `rtrunpois` available in the S-archive of statsci.org (Smyth, [12]).

Appendix C: Proof of Proposition 2.1

Algorithm 1 in Appendix A is useful not only to get reduced bootstrap samples, but to obtain the mean and the variance of ν_n, given that $k_1 \leq \nu_n \leq k_2$.

Proof Let D be the number of components of vector $\mathcal{I}_3$ in $\mathcal{I}_1 - \mathcal{I}_2$ and $\nu_D = \sum_{j/j \in \mathcal{I}_1 - \mathcal{I}_2} \delta(N_j)$, where $\delta(u) = 1$ if $u > 0$ and $\delta(u) = 0$ otherwise, that is, ν_D is the number of distinct elements in $\mathcal{I}_3$ not belonging to $\mathcal{I}_2$. We have that $D \sim Bin\left(n - k_1, \frac{k_2 - k_1}{k_2}\right)$ and that the conditional distribution of the vector $(N_j)_{j \in \mathcal{I}_1 - \mathcal{I}_2}$, given that $D = d$, is $\mathcal{M}\left(d; \frac{1}{k_2 - k_1}, \frac{1}{k_2 - k_1}, ..., \frac{1}{k_2 - k_1}\right)$. Hence (Johnson and Kotz [10, p. 114]),

$$E(\nu_D \mid D = d) = (k_2 - k_1) - (k_2 - k_1)\left(1 - \frac{1}{k_2 - k_1}\right)^d \qquad \text{and}$$

$$var(\nu_D \mid D = d) = (k_2 - k_1)(k_2 - k_1 - 1)\left(1 - \frac{2}{k_2 - k_1}\right)^d + \qquad \text{(C1)}$$

$$+ (k_2 - k_1)\left(1 - \frac{1}{k_2 - k_1}\right)^d - (k_2 - k_1)^2\left(1 - \frac{1}{k_2 - k_1}\right)^{2d}.$$

Finally, the result follows from (C1) and

$$E\left(\nu_n \mid k_1 \leq \nu_n \leq k_2\right) = k_1 + E(\nu_D) = k_1 + E_D\left\{E(\nu_D \mid D)\right\},$$
$$var\left(\nu_n \mid k_1 \leq \nu_n \leq k_2\right) = var(\nu_D) = E_D\left\{var(\nu_D \mid D)\right\} + var_D\left\{E(\nu_D \mid D)\right\}.$$

$\square$