

EXTRAPOLATION AND THE BOOTSTRAP

By PETER J. BICKEL

University of California, Berkeley, USA

and

ANAT SAKOV

Technion, Israel Institute of Technology, Haifa, Israel

SUMMARY. The m out of n bootstrap, with or without replacement, where $m \rightarrow \infty$ and $m/n \rightarrow 0$ has been proposed on two grounds: (i) As a way of ensuring consistency when the classical bootstrap is not consistent. (ii) When it is consistent, then in conjunction with extrapolation, as a way of obtaining behaviour equivalent to that of the classical bootstrap, to second or higher order, with reduced computation time. In this paper we shall discuss a partial taxonomy of higher order behaviour of the m out of n bootstrap and introduce a general form of extrapolation.

1. Introduction

The m out of n bootstrap, with or without replacement, where $m \rightarrow \infty$ and $m/n \rightarrow 0$ has been proposed on two grounds:

- (i) When the classical bootstrap is not consistent, as a way of ensuring consistency under minimal conditions (Bickel et al., 1997, Politis and Romano 1994, Politis et. al, 1999). We note that what we call the m out of n bootstrap is referred to as subsampling by Politis et. al, 1999 and others.
- (ii) When the classical bootstrap is consistent, then in conjunction with extrapolation, as a way of obtaining behaviour equivalent to that of the classical bootstrap, to second or higher order, with reduced computation time (Bickel and Yahav 1988, Bickel et al., 1997, Bertail 1997, Sakov, 1998).

Paper received January 2002; revised April 2002.

AMS (2000) subject classification. Primary 62G09; secondary 62G05.

Keywords and phrases. Bootstrap, extrapolation, subsampling, Edgeworth expansion.

In this paper we shall,

1. Review, briefly, the established second order results, mentioned in (ii). This is done in Section 2.
2. In Section 3 we present a taxonomy of possible situations and introduce a general form of extrapolation whose use we demonstrate in the next section.
3. Three examples of categories of the taxonomy are studied in Sections 4 and 5. This is done in the context of (i) and (ii) above.

Extrapolation is useful only if an expansion for the distribution of a statistic is known to exist, or may be conjectured. If this is not the case, and the classical bootstrap is not consistent, then the m out of n bootstrap can be used, where m can be chosen using the adaptive rule discussed in Bickel and Sakov (2002), Sakov (1998), Sakov and Bickel (1999) and Götze and Račkauskas (2001). This choice of m , basically guarantees first-order accuracy.

2. Review of previous uses of extrapolation

Let X_1, \dots, X_n be an iid sample from a distribution $F \in \mathcal{F}_0$, and let $T_n = T(X_1, \dots, X_n, F)$ be a random variable of interest. Let $L_n(x) = P(T_n \leq x)$ and assume it converges weakly to a nondegenerate limiting distribution L . Let X_1^*, \dots, X_n^* be a bootstrap sample drawn from \hat{F}_n , the empirical distribution function. The classical bootstrap (Efron, 1979) estimates the unknown L using $L_{n,n}^*$, where

$$L_{n,n}^*(x) = P^*(T_n^* \leq x) = P\left(T_n^* \leq x \mid \hat{F}_n\right), \text{ where } T_n^* = T\left(X_1^*, \dots, X_n^*, \hat{F}_n\right).$$

The bootstrap approximation is known to be consistent in many examples (Shao and Tu, 1995), and to fail in others (Bickel and Freedman, 1981, Hall and Wilson, 1991, Bickel and Ren, 1996, Bickel et al., 1997). The m out of n bootstrap, with or without replacement, where $m \rightarrow \infty$ and $m/n \rightarrow 0$ has been proposed in such cases, and it has been shown to be consistent under minimal conditions (Bickel et al., 1997, Politis and Romano, 1994, Politis et al., 1999). We use T_m^* and $L_{m,n}^*$ as the analogs of T_n^* and $L_{n,n}^*$ when the bootstrap sample size is m . Under this set-up it is possible to choose m , adaptively, in a way that guarantees consistency, and first order rate behaviour (Bickel and Sakov, 2002).

When the classical bootstrap is consistent, then the m out of n bootstrap can be used in conjunction with extrapolation, as a way of obtaining behaviour equivalent to that of the classical bootstrap, to second or higher order, with reduced computation time. We now review this result.

Assume that for some $\alpha > 0$, the following expansions hold:

$$L_n(x) = A_0(x, F) + \frac{1}{n^\alpha} A_1(x, F) + O\left(\frac{1}{n^{2\alpha}}\right) \quad \text{as } n \rightarrow \infty,$$

$$L_{m,n}^*(x) = A_0(x, \hat{F}_n) + \frac{1}{m^\alpha} A_1(x, \hat{F}_n) + O_p\left(\frac{1}{m^{2\alpha}}\right) \quad \text{as } m, n \rightarrow \infty. \quad (1)$$

We assume α is known (typically 1/2 or 1), but the coefficients of the expansion are unknown in general without knowledge of F .

Bickel and Yahav (1988) propose to obtain bootstrap distributions $L_{m_j,n}^*(x)$ for $m_1, m_2 \ll n$. Ignoring the error term in (1), and solving two equations in two unknowns produces \hat{A}_0 and \hat{A}_1 , the estimates of the unknown coefficients in (1). Next estimate $L_n(x)$ by $\hat{L}_n(x) = \hat{A}_0 + (1/n^\alpha)\hat{A}_1$. Under some conditions, this achieves the same second order accuracy for confidence levels as the classical bootstrap. This is done, however, with reduced computation time, since the statistic needs to be evaluated on a sample of size $m_j \ll n$. Bertail (1997) considers using interpolation when the sampling is done without replacement: he assumes an Edgeworth expansion with a known limiting distribution, and proposes to use the known limiting distribution and the bootstrap distribution to improve to second order.

3. A Taxonomy of Situations and General Extrapolation

We study situations in which L_n has an asymptotic expansion in powers of $n^{-1/2}$ to p terms. That is, for $F \in \mathcal{F}_0$,

$$L_n(x) = A_0(x, F) + \sum_{j=1}^{p-1} \frac{1}{n^{j/2}} A_j(x, F) + O\left(n^{-p/2}\right), \quad (2)$$

where, O is interpretable as referring to the sup norm, now and in the sequel. We will also have occasion to consider, when it is defined,

$$\hat{L}_{m,n}(x) = A_0(x, \hat{F}_n) + \sum_{j=1}^{p-1} \frac{1}{m^{j/2}} A_j(x, \hat{F}_n).$$

We will assume that the bootstrap distribution $L_{m,n}^*$ has the expansion,

$$L_{m,n}^*(x) = \hat{L}_{m,n}(x) + O_p\left(m^{-p/2}\right) + O_p\left(\frac{m^{\alpha/2}}{n^{1/2}}\Omega(n)\right), \tag{3}$$

where Ω is a slowly varying function. Note that in the RHS of (3) we have $\hat{L}_{m,n}(x)$, where the target is $L_n(x)$ or $\hat{L}_{n,n}(x)$. We will use extrapolation in order to improve from the former to a good estimate of the latter. We distinguish three cases:

- I) $\alpha = 0$ and we estimate $\hat{L}_{n,n}(x)$. This is the “regular” situation, where the classical n -bootstrap is consistent for $A_0(x, F)$ and the parameters $A_j(x, F)$ are themselves estimable at rate $n^{-1/2}$. The prototypical example is the centered mean,

$$T(\hat{F}_n, F) = \sqrt{n} \int x d(\hat{F}_n - F).$$

Here extrapolation can improve coverage probability accuracy arbitrarily with a suitable choice of pivot $A_j(x, G)$ independent of G , eg. the studentized mean.

- II) $0 < \alpha < 1$ and we estimate $L_n(x)$. These are situations where the n -bootstrap is consistent, but the parameter A_0 is not estimable at a rate $n^{-1/2}$. The example we study in Section 5 is the centered median,

$$T(\hat{F}_n, F) = \sqrt{n} \left(\hat{F}_n^{-1}(1/2) - F^{-1}(1/2) \right).$$

Here, the m -bootstrap can do better than the n -bootstrap (Sakov and Bickel, 2000) and extrapolation improves further.

- III) $\alpha = 1$. Here, the n -bootstrap is inconsistent, but the m -bootstrap works under mild regularity conditions. There are at least two sub-cases of III:

- (a) The $O_p\left(\sqrt{m/n}\right)$ term is purely stochastic and extrapolation gives no improvement. An example of this situation is setting confidence bounds for extrema (Bickel and Sakov, 2002).
- (b) We can write, for some $r \geq 1$

$$O_p\left(\left(\frac{m}{n}\right)^{1/2}\right) = \sum_{j=1}^r B_j\left(x, \hat{F}_n\right) \left(\frac{m}{n}\right)^{j/2} + O_p\left(\left(\frac{m}{n}\right)^{(r+1)/2}\right).$$

An example, is the non-centered mean $T(\hat{F}_n) = \sqrt{n} \int x d\hat{F}_n$ (Sakov, 1998). In this somewhat artificial case, which arises in setting critical value for tests and which we consider in Section 4, extrapolation can lead to approximation to the level which is arbitrarily good if p, r are sufficiently large.

In the remainder of this paper we consider cases I, II and III(b).

In the general form, (3), there are p unknown coefficients, $A_j(x, \hat{F}_n)$, we wish to estimate using extrapolation. In case III(b), we estimate the r unknowns $B_j(x, \hat{F}_n)$, in order to eliminate them.

For testing situations, we wish to estimate the critical value for a level ζ test ($0 < \zeta < 1$ and bounded away from 0 and 1). We assume the null distribution and the bootstrap distribution have the expansions (2) and (3). Then the expansions for the quantiles, $L_n^{-1}(\zeta)$ and $(L_{m,n}^*)^{-1}(\zeta)$ exist and have the same powers and errors as in (2) and (3), and the unknown coefficients depend on the data, F and ζ (this is known as the Cornish-Fisher expansion, see for instance Hall (1992)). Since the form of this expansion is the same as for L_n , we can present, a unified approach to extrapolation, covering cases I, II and III(b) for both estimation and testing.

To simplify the writing, we will denote $\beta_{-j} = A_j(x, \hat{F}_n)$ for $j = 0, \dots, p-1$, and $\beta_j = B_j(x, \hat{F}_n)$ for $j = 1, \dots, r$. Let

$$\beta = (\beta_{-p+1}, \dots, \beta_0, \dots, \beta_r)^t.$$

Note that the β_j coefficients are unknown and (possibly) depend on F and the data, but *not* on m .

The general extrapolation method we use, is the following. Let, $m_k = \lambda_k n^q$ for $k = 1, \dots, K$ and $\lambda_1 < \dots < \lambda_K$ fixed. For case I, $0 < q < 1$; For case II, $1/(p+\alpha) \leq q \leq 1-p/((r+1)(p+\alpha))$ and for case III(b), $q \leq (r+1)/(p+r+1)$. For $k = 1, \dots, K$, let $Y_k = L_{m_k,n}^*(x)$ for estimation and $Y_k = (L_{m_k,n}^*)^{-1}(\zeta)$ for testing. Finally, let $\mathbf{Y} = (Y_1, \dots, Y_K)^t$. We thus propose to fit the regression,

$$Y_k = \sum_{j=-p+1}^0 \beta_j m_k^{j/2} + \sum_{j=1}^r \beta_j \left(\frac{m_k}{n}\right)^{j/2} + O_p(n^{-\gamma/2}), \quad k = 1, \dots, K,$$

where $\gamma = \min(qp, (1-q)(r+1))$, and the O terms takes the role of the error in regression. It should be clear that the second sum is part of the regression

only when $\alpha = 1$. Thus we obtain,

$$\hat{\beta} = (M^t M)^{-1} M^t \mathbf{Y},$$

where $M^t = D_n \Lambda$ and Λ is a $(p + r)$ by K matrix with entries $\lambda_j^{t/2}$ for $-p + 1 \leq t \leq r$ and $1 \leq j \leq K$. D_n is a $p + r$ by $p + r$ diagonal matrix with entries $n^{qt/2}$ for $-p + 1 \leq t \leq 0$ and $n^{(q-1)t/2}$ for $1 \leq t \leq r$. Λ^t is a portion of a non-singular Vandermonde matrix, V . Hence, $\Lambda \Lambda^t$ is the top-left matrix of the non-singular VV^t , and hence $\Lambda \Lambda^t$ is non-singular, and the solution is unique.

In cases I and II,

$$\hat{\theta} = \sum_{j=-p+1}^0 \hat{\beta}_j n^{j/2}$$

estimates $\theta = L_n(x)$ or $L_n^{-1}(\zeta)$. In case III(b), $\hat{\theta}$ estimates $L_{n,n}(x)$ or $(L_{n,n})^{-1}(\zeta)$.

THEOREM 1 *Under the above set-up, $\hat{\theta} = \theta + O_p(n^{-\gamma/2})$, where γ is as before.*

PROOF:

$$\begin{aligned} \hat{\beta} &= (M^t M)^{-1} M^t \mathbf{Y} = (D_n \Lambda \Lambda^t D_n)^{-1} D_n \Lambda [(\Lambda^t D_n) \beta + O_p(n^{-\gamma/2})] \\ &= \beta + D_n^{-1} (\Lambda \Lambda^t)^{-1} \Lambda O_p(n^{-\gamma/2}). \end{aligned}$$

λ_j are fixed so $(\Lambda \Lambda^t)^{-1} \Lambda$ is a matrix with all entries being $O(1)$. Hence,

$$\hat{\beta}_j = \beta_j + O_p(n^{-\frac{1}{2}(\gamma+qj)}) \quad \text{for } j = 0, \dots, -p + 1,$$

and

$$\hat{\theta} = \sum_{j=-p+1}^0 \hat{\beta}_j n^{j/2} = \theta + O_p(n^{-\gamma/2}). \quad \square$$

4. Cases I and III(b)

Case I has essentially been dealt with in past work. It is of interest particularly in constructing confidence bounds.

The methods by which Cornish-Fisher expansion of the bootstrap distribution of $\sqrt{n}(\hat{\mu}_n - \mu(F))$ can be used to construct perturbations Δ_n of order $1/\sqrt{n}$ to $\hat{\mu}_n$ such that $\sqrt{n}(\hat{\mu}_n + \Delta_n - \mu(F))$ tends to Gaussianity more quickly than $\sqrt{n}(\hat{\mu}_n - \mu(F))$ are discussed in the papers by Hall (1998) and Efron (1987). Extrapolation is not used there, and computations are done explicitly, evidently becoming more and more complex for larger p . This is evaded by our method for case I which is also computationally faster. Unfortunately, the choice of the design M_n remains somewhat arbitrary.

The situation of case III(a) is not amenable to extrapolation but good selection rules for m can be devised – see Bickel and Sakov (2002). The prototypical case for situation III(b) is testing $H_0 : \mu(F) = 0$ using the test statistic $\hat{\mu}_n$ such that under regularity conditions $\sqrt{n}(\hat{\mu}_n - \mu(F))$ tends to a Gaussian limit, and further (2) and (3) hold under $F \in \mathcal{F}_0 = \{F \mid \mu(F) = 0\}$.

Using the n -bootstrap to set critical values gives inconsistency (Hall and Wilson, 1991). Using the m out of n bootstrap with $m \rightarrow \infty$ and $m/n \rightarrow 0$ gives consistency but typically slow convergence to the correct limit and poor power behaviour. Extrapolation even for $p = 1$ effectively makes the test equivalent to having a critical value based on the bootstrap distribution of $\sqrt{n}(\hat{\mu}_n^* - \hat{\mu}_n)$, and gives much better results. Further expansion yields second order correction if A_0 is independent of $F \in \mathcal{F}_0$. Examples and details are discussed in Sakov (1998). Figures 1 and 2 give some simulation results of the mean and the trimmed mean $\mu(F) = (1 - 2\alpha)^{-1} \int_{\alpha}^{1-\alpha} F^{-1}(s) ds$. For more details on the simulations, see Section 6.

As far as we know, the situation exemplified by case III(b) occurs only in testing $H_0 : \tau(F) = 0$, where τ may even be abstract and the test statistic is of the form $\sqrt{n}||\hat{\tau}_n||$ for some norm. Then as usual, setting critical values using the n bootstrap is foolish, and the m out of n bootstrap and extrapolation can save the day. But these are also cases where the situation may be fixed more straightforwardly by using the bootstrap distribution of $\sqrt{n}||\hat{\tau}_n^* - \hat{\tau}_n||$ (Bickel and Ren, 2000).

5. Extrapolation for Irregular Functionals – Case II

This section deals with a single example, the median, initially investigated without the use of extrapolation in Sakov and Bickel (2000).

It is well known that the n -bootstrap distribution of the median is consistent (Singh, 1981). Let M_n be the median of \hat{F}_n and $\mu = F^{-1}(1/2)$. Then if F

has a positive derivative $f(\mu)$ at μ ,

$$\sup_x |P^* (\sqrt{n} (M_n^* - M_n) \leq x) - \Phi (2f(\mu)x)| \rightarrow 0,$$

in probability (in fact a.s).

Suppose first that $f \in \mathcal{F}_1 \equiv \{f : |f'(\mu)| \leq M\}$ for some fixed finite M . Then Reiss (1976) showed

$$\sup_x |P (\sqrt{n} (M_n - \mu) \leq x) - \Phi (2f(\mu)x)| = O (n^{-1/2}).$$

and Singh (1981) showed that

$$\sup_x |P^* (\sqrt{n} (M_n^* - M_n) \leq x) - \Phi (2f(\mu)x)| = O \left(\frac{\Omega(n)}{n^{1/4}} \right), \tag{4}$$

where $\Omega(n) = \sqrt{\log(\log(n))}$ is a slowly varying function. It is well known that $f(\mu)$ can be estimated at rate $n^{-1/3}$ uniformly on \mathcal{F}_0 (Silverman, 1986), so the bootstrap approximation is poor. Suppose we use the m out of n bootstrap. We argue below that using $m_n = n^{2/3}$ yields an error in (4) which is $n^{-1/3}\Omega(n)$ i.e.

$$\sup_x |P^* (\sqrt{m} (M_{m_n}^* - M_n) \leq x) - \Phi (2f(\mu)x)| = O_p (n^{-1/3}\Omega(n)).$$

Note the improvement over the classical n -bootstrap. This is the best that can be achieved with the m out of n bootstrap without further refinement. On the other hand, if $\mathcal{F}_p \equiv \{f : |f^{(p)}(\mu)| \leq M\}$, it is well known that f and hence the distribution of $\sqrt{n} (M_n - \mu)$ can be approximated at rate $n^{-\frac{p}{2p+1}}$. We show below that if we combine extrapolation with the m out of n bootstrap, we can achieve the same rate up to $\Omega(n)$.

Here is the argument for the median. Let M_n be the upper end point of the median interval and define M_m^* similarly. We shall assume, without loss of generality, that f has compact support, since the probability that M_n or M_m^* fall outside is exponentially small. The event $\{M_m^* \leq t\}$ is equivalent to the event $\{\sum_{i=1}^m 1(X_i^* > t) < m/2\}$. Then

$$P^* (\sqrt{m} (M_m^* - M_n) \leq x) = P \left(\text{Bin} \left(m, \bar{F}_n \left(M_n + \frac{x}{\sqrt{m}} \right) \right) < \frac{m}{2} \right),$$

where $\text{Bin}(m, p)$ is a Binomial (m, p) variable and $\bar{F}_n(y) = 1 - \hat{F}_n(y-)$. Let $\pi_m(x) = \bar{F}_n (M_n + x/\sqrt{m})$ and Z_m be Binomial $(m, \pi_m(x))$ variable. Using

$|\bar{F}(M_n) - 1/2| \leq 1/n$, the last equation can be written as

$$\begin{aligned} \mathbf{P}^* (\sqrt{m}(M_m^* - M_n) \leq x) &= \mathbf{P} \left(\sqrt{m} \left(\frac{1}{m} Z_m - \pi_m(x) \right) < \right. \\ &\quad \left. \sqrt{m}(\pi_m(0) - \pi_m(x)) + O_p \left(\frac{\sqrt{m}}{n} \right) \right). \end{aligned} \quad (5)$$

where $O_p(\sqrt{m}/n)$ is uniform. Let $\bar{F} = 1 - F$ and $\pi(x) = \bar{F}(M_n + x/\sqrt{m})$ and write

$$W_n(t) = \sqrt{n} \left(\bar{F}_n(t) - \bar{F}(t) \right).$$

Then by a classical result, see for instance Shorack and Wellner (1986), (Theorem 1, p. 542), and using the compactness of the support it follows that,

$$\begin{aligned} &\sup_x \left| W_n \left(M_n + \frac{x}{\sqrt{m}} \right) - W_n(M_n) \right| \\ &= \sup_x \left| \sqrt{n} ((\pi_m(x) - \pi_m(0)) - (\pi(x) - \pi(0))) \right| \\ &= O_p \left(m^{-1/4} \Omega(m) \right). \end{aligned} \quad (6)$$

Since $m < n$, $\Omega(m) \leq \Omega(n)$. Substituting in (5) we get,

$$\begin{aligned} \mathbf{P}^* (\sqrt{m}(M_m^* - M_n) \leq x) &= \mathbf{P} \left(\sqrt{m} \left(\frac{1}{m} Z_m - \pi_m(x) \right) \leq \sqrt{m}(\pi(0) - \pi(x)) \right. \\ &\quad \left. + O_p \left(\frac{\sqrt{m}}{n} + m^{1/4} \frac{\Omega(n)}{\sqrt{n}} \right) \right). \end{aligned} \quad (7)$$

Next note that since $|f'|$ is bounded, then for $|x| \leq \Omega(n)$,

$$\pi(0) - \pi(x) = \bar{F}(M_n) - \bar{F} \left(M_n + \frac{x}{\sqrt{m}} \right) = f(M_n) \frac{x}{\sqrt{m}} + O_p \left(\frac{1}{m} \right). \quad (8)$$

Since $|M_n - \mu| = O_p(n^{-1/2})$ and for $m < n$, $\sqrt{m}/n < 1/\sqrt{m}$, we finally conclude from (7) and (8) that

$$\begin{aligned} \mathbf{P}^* (\sqrt{m}(M_m^* - M_n) \leq x) &= \mathbf{P} \left(\sqrt{m} \left(\frac{1}{m} Z_m - \pi_m(x) \right) \leq x f(\mu) \right. \\ &\quad \left. + O_p \left(m^{1/4} \frac{\Omega(n)}{\sqrt{n}} + \frac{1}{\sqrt{m}} \right) \right). \end{aligned}$$

Represent $P(\text{Bin}(m, \pi) \leq k)$ as an incomplete beta-function and arguing as in Bickel (1992), we get, uniformly,

$$P^*(\sqrt{m}(M_m^* - M_n) \leq x) = \Phi(2f(\mu)x) + O_p\left(m^{1/4}\frac{\Omega(n)}{\sqrt{n}} + \frac{1}{\sqrt{m}}\right). \quad (9)$$

Equating the two error terms gives $m^{3/4} = \sqrt{n}$ or $m = n^{2/3}$ yielding a rate of $n^{-1/3}\Omega(n)$, as claimed. We are able to bring the O_p out of the probability statement since the incomplete beta-function is smooth.

Now suppose, $f \in \mathcal{F}_p$ where $p \geq 2$. Evidently, (7) still holds. Now we can expand $\pi(0) - \pi(x)$ to obtain a polynomial of order $p - 1$ in $m^{-1/2}$ to obtain,

$$\begin{aligned} P^*(\sqrt{m}(M_m^* - M_n) \leq x) &= \Phi(2xf(\mu)) + \sum_{j=1}^{p-1} \phi(2xf(\mu))\frac{Q_j(x)}{m^{j/2}} \\ &\quad + O_p\left(\frac{\sqrt{m}}{n} + m^{1/4}\frac{\Omega(n)}{\sqrt{n}} + m^{-p/2}\right). \end{aligned}$$

Extrapolate, as we have indicated, to get $\hat{\beta}_j$ of $\phi(2xf(\mu))Q_j(x)$ for $j \geq 1$ and $\hat{\beta}_0$ of $\Phi(2xf(\mu))$. Let,

$$\hat{P}_n(x) = \sum_{j=-p+1}^0 \hat{\beta}_j n^{-j/2}.$$

Then by Theorem 1

$$\hat{P}_n(x) = P(\sqrt{n}(M_n - \mu) \leq x) + O_p\left(m^{-p/2} + \frac{\sqrt{m}}{n} + m^{1/4}\frac{\Omega(n)}{\sqrt{n}}\right).$$

Putting $m^{(2p+1)/4} = \sqrt{n}$ or $m = n^{2/(2p+1)}$ we obtain a rate of $n^{-p/(2p+1)}\Omega(n)$ as desired for the estimate $\hat{P}_n(x)$.

Is this a general phenomenon?

Suppose we have a parameter $\theta_n(P)$ which we wish to estimate and

$$\theta_n(F) = \theta(F) + \sum_{j=1}^{p-1} \frac{\gamma_j(F)}{n^{j/2}} + O\left(n^{-p/2}\right),$$

and

$$\theta_m(\hat{F}_n) = \theta_m(F) + O_p\left(\frac{m^\alpha}{\sqrt{n}}\right).$$

Then it is clear that extrapolation can achieve a rate of $n^{-\frac{p}{2}(2\alpha+p)^{-1}}$. We would expect to see this phenomenon in, for instance, bootstrapping M estimates with influence function vanishing outside a compact. We do not know how general the phenomenon is.

6. Simulations

In Figures 1 and 2 we present results of testing whether the mean (one and two-sided) and the trimmed mean (trimming proportion 0.05) are zero. For more simulation see Sakov (1998). The distributions and sample sizes are indicated for each case. The plots show the asymptotic power function (labelled as exact), the power of extrapolation and the power when the null is imposed so that bootstrap samples of size n are consistent (labelled centered).

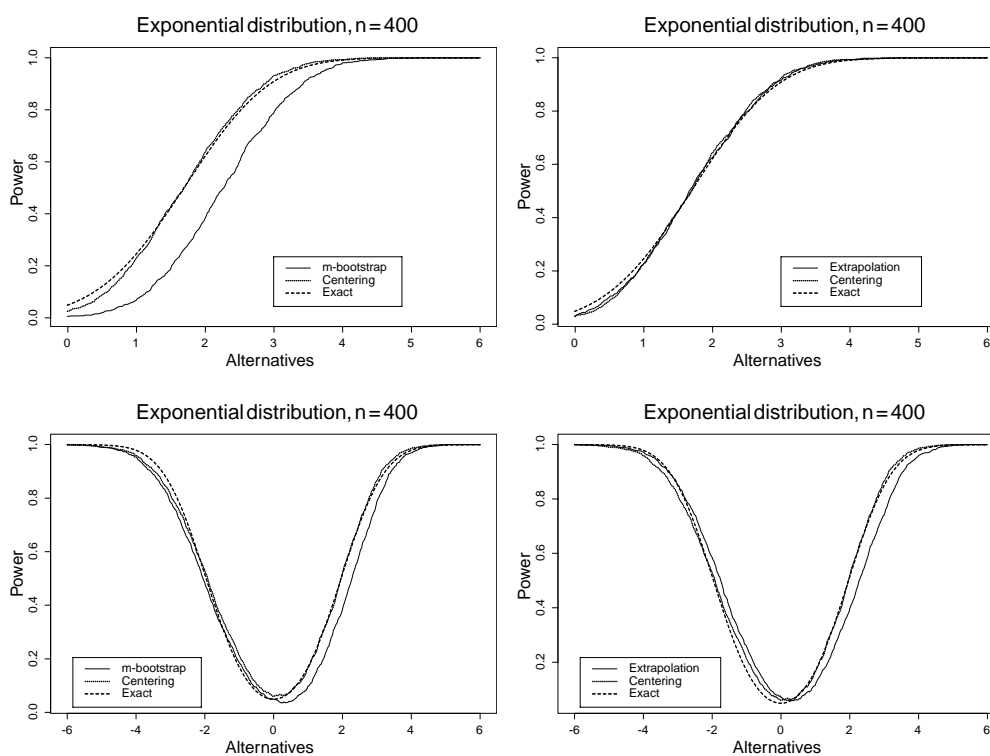


Figure 1. Testing for the mean (one and two-sided)

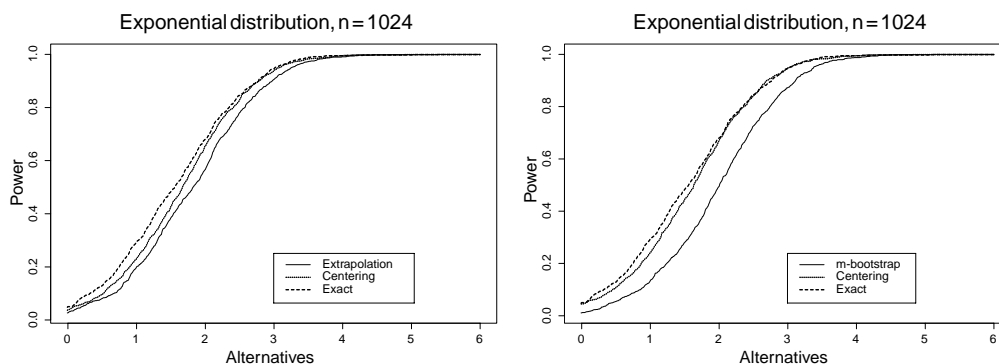


Figure 2. Testing the trimmed mean

For comparison, we present the power for a bootstrap sample of size $m = \sqrt{n}$. The alternatives are on the \sqrt{n} -scale. Clearly, the extrapolation performance is compatible to using bootstrap samples of size n , but extrapolation is faster. The improvement over using $m = \sqrt{n}$ was true for other choices of m as well.

The m -bootstrap in the middle pair seems better than the m -bootstrap in the first pair. This is so since it is a two-sided test. For a detailed analysis see Sakov (1998). The m -bootstrap in the last pair is better than the first since the sample size is larger.

References

- BERTAIL, P. (1997). Second-order properties of an extrapolated bootstrap without replacement under weak assumptions, *Bernoulli*, **3**, 149-179.
- BICKEL, P. (1992). Inference and auditing: the stringer bound, *Internat. Statist. Rev.*, **60**, 197-209.
- BICKEL, P. and FREEDMAN, D. (1981). Some asymptotic theory for the bootstrap, *Ann. Statist.*, **9**, 1196-1217.
- BICKEL, P. and REN, J. (1996). The m out of n bootstrap and goodness of fit tests with doubly censored data. In *Robust Statistics: data, analysis and computer intensive methods*, H. Rieder, ed., Springer-Verlag, New York, 35-47.
- BICKEL, P. and REN, J. (2000). The bootstrap in hypothesis testing. In *State of the Art in Probability and Statistics: a festschrift for W.R. van Zwet*, M. de Gunst, C. Klaassen and A. van der Vaart, eds., IMS lectures notes **36**, Institute of Mathematical Statistics, Beachwood, OH, 91-112.
- BICKEL, P. and SAKOV, A. (2002). On the choice of m in the m out of n bootstrap in estimation problems, in preparation.
- BICKEL, P. and YAHAV, J. (1988). Richardson extrapolation and the bootstrap, *J. Amer. Statist. Assoc.*, **83**, 387-393.
- BICKEL, P., GÖTZE, F. and VAN ZWET, W. (1997). Resampling fewer than n observations: gains, losses and remedies for losses, *Statist. Sinica*, **7**, 1-31.

- EFRON, B. (1979). Bootstrap methods: another look at the jackknife, *Ann. Statist.*, **7**, 1-26.
- EFRON, B. (1987). Better bootstrap confidence intervals (with discussion), *J. Amer. Statist. Assoc.*, **82**, 171-200.
- GÖTZE, F. and RAČKAUSKAS, A. (2001). Adaptive choice of bootstrap sample sizes. In *State of the Art in Probability and Statistics: a festschrift for W.R. van Zwet*, M. de Gunst, C. Klaassen and A. van der Vaart, eds., IMS lectures notes **36**, Institute of Mathematical Statistics, Beachwood, OH, 286-309.
- HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussion), *Ann. Statist.*, **16**, 917-953.
- HALL, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.
- HALL, P. and WILSON, S. (1991). Two guidelines for bootstrap hypothesis testing, *Biometrics*, **47**, 757-762.
- POLITIS, D. and ROMANO, J. (1994). Large sample confidence regions on subsamples under minimal assumptions., *Ann. Statist.*, **22**, 2031-2050.
- POLITIS, D., ROMANO, J. and WOLF, M. (1999). *Subsampling*, Springer-Verlag, New York.
- REISS, R. (1976). Asymptotic expansions for sample quantiles, *Ann. Probab.*, **4**, 249-258.
- SAKOV, A. (1998). *Using the m out of n Bootstrap in Hypothesis Testing*. PhD thesis, University of California, Berkeley.
- SAKOV, A. and BICKEL, P. (1999). Choosing m in the m out of n bootstrap, *Proc. Amer. Statist. Assoc.* (Section on Bayesian Statistical Science), 125-128.
- SAKOV, A. and BICKEL, P. (2000). An Edgeworth expansion for the m out of n bootstrapped median, *Statist. Probab. Lett.*, **49**, 217-223.
- SHAO, J. and TU, D. (1995). *The Jackknife and Bootstrap*, Springer-Verlag, New York.
- SHORACK, G. and WELLNER, J. (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.
- SILVERMAN, B. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.
- SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap, *Ann. Statist.*, **9**, 1187-1195.

PETER J. BICKEL
 DEPARTMENT OF STATISTICS
 UNIVERSITY OF CALIFORNIA, BERKELEY
 367 EVANS HALL
 BERKELEY, CA 94720-3860
 USA
 E-mail: bickel@stat.berkeley.edu

ANAT SAKOV
 FACULTY OF INDUSTRIAL ENGINEERING
 AND MANAGEMENT
 TECHNION – ISRAEL INSTITUTE
 OF TECHNOLOGY
 HAIFA 32000, ISRAEL
 E-mail: sakov@ie.technion.ac.il