Giuseppe Cavaliere, Iliyan Georgiev,
A.M. Robert Taylor

Wild bootstrap of the mean in the infinite variance case

Quaderni di Dipartimento

Serie Ricerche 2011, n. 5
ISSN 1973-9346

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze Statistiche "Paolo Fortunati"

Wild Bootstrap of the Sample Mean in the Infinite Variance Case*

Giuseppe Cavaliere[a], Iliyan Georgiev[b] and A.M.Robert Taylor[c]

[a]Department of Statistical Sciences, University of Bologna

[b]Faculdade de Economia, Universidade Nova de Lisboa

[c]School of Economics, University of Nottingham

March 2011

Abstract

It is well known that the standard i.i.d. bootstrap of the mean is inconsistent in a location model with infinite variance (α-stable) innovations. This occurs because the bootstrap distribution of a normalised sum of infinite variance random variables tends to a random distribution. Consistent bootstrap algorithms based on subsampling methods have been proposed but have the drawback that they deliver much wider confidence sets than those generated by the i.i.d. bootstrap owing to the fact that they eliminate the dependence of the bootstrap distribution on the sample extremes. In this paper we propose sufficient conditions that allow a simple modification of the bootstrap (Wu, 1986, Ann.Stat.) to be consistent (in a conditional sense) yet to also reproduce the narrower confidence sets of the i.i.d. bootstrap. Numerical results demonstrate that our proposed bootstrap method works very well in practice delivering coverage rates very close to the nominal level and significantly narrower confidence sets than other consistent methods.

Keywords: Bootstrap, stable distributions, random probability measures, weak convergence.

1 Introduction

Let $\{Y_i\}_{i\in\mathbb{N}}$ be a sequence of independent, identically distributed (i.i.d.) symmetric random variables (r.v.'s) with a common distribution function F which is in the domain

*We are grateful to two anonymous referees, an Associate Editor and the Editor for their helpful and constructive comments on a previous version of the paper. Cavaliere and Georgiev also thank the Fundação para a Ciência e a Tecnologia, project n^o PTDC/EGE-ECO/108620/2008 for financial support.
Address correspondence to: Robert Taylor, School of Economics, The Sir Clive Granger Building, University of Nottingham, Nottingham NG7 2RD, U.K. E-mail: Robert.Taylor@nottingham.ac.uk

of attraction of an α-stable law, $\alpha \in (0, 2)$. That is,

$$1 - F(x) \sim px^{-\alpha} L(x) \text{ as } x \to \infty$$

with $p := \lim_{x\to\infty} (1 - F(x)) / (1 - F(x) + F(-x))$ and $L(\cdot)$ a slowly varying function at infinity. The parameter α (the so-called index of stability or characteristic exponent), which controls the thickness of the tails of the distribution of Y_i, will be treated as unknown in this paper. For further details see Chapter XVII of Feller (1971).

Consider now the location model $X_i = \theta + Y_i$ with location parameter θ. As shown in, for example, Feller (1971) there exists an increasing sequence a_n with $nP(|Y_1| > a_n x) \to x^{-\alpha}$ as $n \to \infty$ such that

$$S_n := a_n^{-1} \sum_{i=1}^{n} (X_i - \theta) = a_n^{-1} \sum_{i=1}^{n} Y_i \xrightarrow{w} S(\alpha), \text{ as } n \to \infty, \tag{1}$$

where $\xrightarrow{w}$ denotes weak convergence and $S(\alpha)$ is a stable random variable with index α. Well-known special cases of $S(\alpha)$ are the Gaussian ($\alpha = 2$) and Cauchy ($\alpha = 1$) distributions. Notice that α controls the existence or otherwise of the moments of $S(\alpha)$; in particular, $S(\alpha)$ has finite k-th order moment if and only $\alpha > k$ (or $\alpha = 2$). For $1 < \alpha < 2$, infinite variance but finite mean therefore obtain, while for $\alpha \leq 1$ both the mean and the variance are infinite. It is worth noting that having infinite variance does not necessarily imply that Y_i is in the domain of attraction of $S(\alpha)$ with $\alpha < 2$. For example, if Y_i is Student t with 2 degrees of freedom then, despite the fact that $V(Y_i) = \infty$, the convergence in (1) is satisfied for $\alpha = 2$ and $a_n := (n \log n)^{1/2}$; cf. Abadir and Magnus (2003).

Now, let $\{X_i^*\}_{i=1,..,n}$ be an i.i.d. sample from the empirical distribution function of $\{X_i\}_{i=1,..,n}$. Then, the distribution of the bootstrap analogue of S_n, say

$$S_n^* := a_n^{-1} \sum_{i=1}^{n} \left(X_i^* - \bar{X}_n\right), \; \bar{X}_n := n^{-1} \sum_{i=1}^{n} X_i$$

is dominated by the sample extremes (see, Athreya, 1987; Knight, 1989; Hall, 1992), with two important consequences. First, the distribution of the bootstrap version of the mean fails to estimate consistently the asymptotic (*unconditional*) distribution of the sample mean; instead, it tends to a random asymptotic distribution. Second, it frequently exhibits lower dispersion than the asymptotic distribution of the sample mean.

Work on the bootstrap of the sample mean for $\alpha \in (0, 2)$ has focused on new resampling algorithms for consistent estimation of the asymptotic distribution of the mean. In particular, Arcones and Giné (1989, 1991) establish that consistency can be obtained using the 'm out of n' bootstrap; i.e., a bootstrap based on the statistic

$$S_{m,n}^* := a_m^{-1} \sum_{i=1}^{m} \left(X_i^* - \bar{X}_n\right),$$

provided the bootstrap sample size m is smaller, at an appropriate rate, than the size n of the observed sample. Similarly, the subsampling methods proposed in Politis and Romano (1994), Romano and Wolf (1999) and Bertail *et al.* (1999) all provide a consistent estimator of the true asymptotic distribution; see also Politis *et al.* (1999). However, as noted by Hall and Yao (2003), because these approaches are based on a sample size that is an order of magnitude smaller than the size of the observed sample, the confidence sets they deliver can be very conservative. Indeed, a finite sample numerical comparison of the m out of n and subsampling methods in the present context is provided by Cornea and Davidson (2009), who show both approaches to be quite unreliable in practice. They also provide a p-value based refined bootstrap method. None of these approaches, however, all of which eliminate the dependence of the resampling distribution on the sample extremes, are capable of reproducing the narrow confidence sets of the i.i.d. bootstrap.

In section 2 of this paper we discuss how a simple variant of the standard bootstrap can reproduce the narrow asymptotic confidence sets of the i.i.d. bootstrap and, at the same time, deliver correct (asymptotic) coverage probabilities. The idea we propose stems from the fact that, from a practical perspective, the main shortcoming of the i.i.d. bootstrap is not its convergence to a random distribution (as opposed to the fixed asymptotic distribution of the sample mean), but rather, its failure to deliver correct (asymptotic) coverage probabilities. We will establish that the bootstrap can be used in order to estimate an appropriate *conditional* (random) asymptotic distribution of the sample mean and, consequently, to deliver confidence sets with correct asymptotic coverage probabilities.

The variant of the standard bootstrap that we will consider is a version of the so-called *wild* bootstrap of Wu (1986), Liu (1988) and Mammen (1993). We will show that this bootstrap allows for consistent estimation of the asymptotic distribution of the sample mean conditional on the sequence $|X_1 - \theta|, ..., |X_n - \theta|$. This key property underlies at least four important advantages over existing bootstrap techniques aiming to estimate the (unconditional) asymptotic distribution of the sample mean. First, the wild bootstrap delivers confidence sets for the location parameter which have correct coverage probability in large samples. Second, it preserves the sample extremes, and thus, frequently evaluates the sample mean with respect to a more concentrated distribution than the unconditional one (i.e., bootstrap confidence sets tend to be smaller). Third, the size m of the wild bootstrap sample coincides with the original sample size, n, hence avoiding the issue of how to choose m in practice (and the resulting impact of the choice of m on the finite sample behaviour of the bootstrap; see Cornea and Davidson, 2009) as is required for the subsampling and m out of n bootstrap based approaches. Fourth, unlike, for example, the p-value based method of Cornea and Davidson (2009), it requires neither the knowledge nor a preliminary estimator of the tail index α, which is a considerable advantage given that it is known to be very difficult to estimate the tail index well; see, in particular, Resnick (1997).

2 Theoretical Results

In order to present the main result on the validity of the wild bootstrap in the case of symmetric, infinite variance random variables, let $Z := \{Z_i\}_{i\in\mathbb{N}}$, where $Z_i^{-\alpha}$ $(i \in \mathbb{N})$ are the arrival times of a Poisson process with unit rate, and let $\delta := \{\delta_i\}_{i\in\mathbb{N}}$ be a sequence of independent r.v.'s with $P(\delta_i = 1) = P(\delta_i = -1) = 0.5$, Z and δ being independent. We state first the following fact regarding a conditional asympotic distribution of the sample mean.

Lemma 1 *Let $\{Y_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. symmetric (about zero) r.v.'s in the domain of attraction of an α-stable law with $\alpha \in (0,2)$. Then the conditional distribution of S_n of (1) given the sequence $\{|Y_i|\}_{i\in\mathbb{N}}$ satisfies*

$$\mathcal{L}\left(S_n \middle| \{|Y_i|\}_{i=1}^n\right) = \mathcal{L}\left(a_n^{-1}\sum_{i=1}^n Y_i \middle| \{|Y_i|\}_{i=1}^n\right) \xrightarrow{w} \mathcal{L}\left(\sum_{i=1}^\infty Z_i\delta_i \middle| Z\right) \tag{2}$$

in the sense of weak convergence of random measures. Here $\mathcal{L}(\cdot|A)$ denotes the conditional distribution of $(\cdot)$ w.r.t. the σ-algebra $\sigma(A)$.

The lemma follows from Lemmas 1 and 2 in LePage, Woodroofe and Zinn (1981) by the argument that Knight (1989) developed for his Theorem 2. Note that the series $\sum_{i=1}^\infty Z_i\delta_i$ conditionally on Z converges almost surely (a.s.) by the three-series theorem, since $Z_i \sim i^{-1/\alpha}$ as $i \to \infty$ (a.s.).[1] Although we do not use Lemma 1 in the proof of our main result, the lemma is important in order to understand why the result holds. Namely, because the simple version of the wild bootstrap that we propose will be shown to yield a consistent estimator of $\mathcal{L}(\sum_{i=1}^\infty Z_i\delta_i | Z)$, the asymptotic distribution of the sample mean *conditional* on the absolute values of the correctly centered observations.

As in Wu (1986) and Liu (1988), let the bootstrap sample be generated as

$$X_i^* := \bar{X}_n + (X_i - \bar{X}_n)w_i\ ,\ i = 1, ..., n, \tag{3}$$

with w_i an i.i.d. sequence with $E(w_i) = 0$ and $E(w_i^2) = 1$. It is well-known that conditionally on $X_1, ...X_n$ it holds that X_i^* form an independent sequence with $E^*(X_i^*) = \bar{X}_n$ and $V^*(X_i^*) = (X_i - \bar{X}_n)^2$ (throughout, the associated conditional measure is denoted by P^*, with E^* and V^* denoting respectively expectation and variance under P^*, i.e. conditional on the original sample). The bootstrap analogue of S_n is given by

$$S_n^* := a_n^{-1}\sum_{i=1}^n \left(X_i^* - \bar{X}_n\right) = a_n^{-1}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)w_i \tag{4}$$

and our interest concerns inference on θ based on the distribution of S_n^* under P^*. In what follows, the distribution function of S_n^* under P^* is denoted by F_n^*, and F_n^{*-1}, defined by $F_n^{*-1}(x) := \inf\{y : F_n^*(y) \geq x\}$ $(x \in (0,1))$, is the corresponding (conditional) quantile function.

[1] This can be seen by noting that $Z_i^{-\alpha}$ can be written as the partial sum of the waiting times of the Poisson process, which are i.i.d. Exp(1).

In Theorem 1 we now state our main result, the proof of which can be found in the appendix.

Theorem 1 *Let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. r.v.'s which are symmetric around θ and lie in the domain of attraction of an α-stable law with $\alpha \in (0,2)$. Also let $P(w_i = 1) = P(w_i = -1) = 0.5$ hold in (3). Then S_n^* of (4) satisfies*

$$\mathcal{L}\left(S_n^* \middle| \{X_i\}_{i=1}^n\right) \xrightarrow{w} \mathcal{L}\left(\sum_{i=1}^{\infty} Z_i\delta_i w_i \middle| Z,\delta\right) = \mathcal{L}\left(\sum_{i=1}^{\infty} Z_i\delta_i \middle| Z\right) \tag{5}$$

and

$$P^*\left(S_n^* \le S_n\right) = F_n^*\left(S_n\right) \xrightarrow{w} U[0,1], \tag{6}$$

the equality in (5) meaning that

$$P\left(\sum_{i=1}^{\infty} Z_i\delta_i w_i \le x \middle| Z,\delta\right) = P\left(\sum_{i=1}^{\infty} Z_i\delta_i \le x \middle| Z\right) \ P\text{-a.s.}, \forall x\in\mathbb{R}.$$

Consequently, for $\eta \in (0,1)$, tests of the null hypothesis $H_0 : \theta = \theta_0$ against either: (i) $H_{1,i} : \theta \neq \theta_0$, or (ii) $H_{1,ii} : \theta > \theta_0$, or (iii) $H_{1,iii} : \theta < \theta_0$, when based on the statistic $|\bar{X}_n - \theta_0|$ for $H_{1,i}$, or the statistic $\bar{X}_n - \theta_0$ for $H_{1,ii}$ and $H_{1,iii}$, and using the critical values $a_n n^{-1} F^{*-1}(1-\eta/2)$ for $H_{1,i}$, or $-a_n n^{-1}F^{*-1}(\eta)$ for $H_{1,ii}$, or $a_n n^{-1} F^{*-1}(\eta)$ for $H_{1,iii}$, each have asymptotic size η. Moreover, confidence intervals for θ of the form

$$\left[\bar{X}_n - a_n n^{-1} F_n^{*-1}(1-\eta+\zeta),\, \bar{X}_n - a_n n^{-1} F_n^{*-1}(\zeta)\right]$$

have asymptotic coverage probability $1-\eta$ for any $\zeta \in (0,\eta)$, an obvious example of which could be $\zeta = \eta/2$.

Some remarks are in order.

Remark 1. From (2) and (5) we see that, although the wild bootstrap does not deliver a consistent estimate of the unconditional asymptotic distribution of the sample mean, it does estimate consistently the asymptotic distribution of the sample mean *conditional* on the absolute values of the correctly centered observations. Moreover, the latter distribution turns out to have an a.s. continuous cumulative distribution process. This makes the wild bootstrap a useful device for drawing inferences on the location parameter θ; for instance, as the sample size diverges, it delivers confidence sets with correct coverage probability.

Remark 2. For applications it is important to note that $a_n n^{-1} F_n^{*-1}$ is the distribution function of $n^{-1}\sum_{i=1}^n \left(X_i^* - \bar{X}_n\right) = n^{-1} a_n S_n^*$ under P^*, so no knowledge of a_n is required to obtain the critical values and the confidence bounds proposed in Theorem 1. Normalization by a_n in the first part of the theorem plays a technical role to ensure that a P^*-distribution with non-degenerate limiting behaviour, like F_n^*, is discussed. Self-normalization of the sample mean is an alternative. For example, as in Athreya (1987),

$\tilde{S}_n := (\max_{i=1,...,n} |X_i|)^{-1} \sum_{i=1}^n (X_i - \theta)$ and $\tilde{S}_n^* := (\max_{i=1,...,n} |X_i|)^{-1} \sum_{i=1}^n \left(X_i^* - \bar{X}_n\right)$ could be considered in place of S_n and S_n^*. In this case $\mathcal{L}\left(Z_1^{-1} \sum_{i=1}^\infty Z_i \delta_i \middle| Z\right)$ would appear on the right-hand side of (2) and (5).

Remark 3. The condition $P(w_t = 1) = P(w_t = -1) = 0.5$ is crucial for the results in Theorem 1 to hold. For instance, if w_i are Gaussian then, conditionally on $X_1, ..., X_n$, the bootstrap statistic S_n^* no longer satisfies (5) and is instead exact Gaussian with mean zero and variance $a_n^{-2} \sum_{i=1}^n (X_i - \bar{X}_n)^2$. The results in Theorem 1 also fail to hold for other popular choices for the distribution of the w_i made in the wild bootstrap literature, including the two-point distributions proposed in Liu (1988) and Mammen (1993). Such distributions, while leading to asymptotic refinements in the finite variance case, cannot therefore be successfully applied in the infinite variance case.

Remark 4. As Theorem 1 holds for all $\alpha \in (0, 2)$, the bootstrap allows consistent estimation of the asymptotic conditional distribution of the sample mean even when the data do not have finite mean. However, consistency of the associated bootstrap tests on the location parameter requires $\alpha \in (1, 2)$. This is the case also for tests based on the asymptotic distribution of the sample mean.

Remark 5. Our proposed bootstrap is based upon the statistic $n^{-1} \sum_{i=1}^n \left(X_i^* - \bar{X}_n\right)$; a closely related bootstrap statistic is given by $n^{-1} \sum_{i=1}^n (X_i^* - \theta_0)$ but where the bootstrap data are generated according to the device

$$X_i^* := \theta_0 + (X_i - \theta_0) w_i \ , \ i = 1, ..., n. \tag{7}$$

These two formulations therefore differ in that while in the latter the observations are centred around a hypothesised value θ_0, in the former they are centred around the sample mean $\bar{X}_n$. It can be shown that the version which centres on θ_0 is equivalent to the statistic obtained using the randomisation test approach outlined in Section 15.2 of Lehmann and Romano (2005).[2] As they demonstrate, this delivers an exact test of $H_0 : \theta = \theta_0$ under the conditions given in this paper. In principle, therefore, an exact α-level confidence interval for θ could be obtained by inverting this test statistic. In practice, however, this has to be done numerically and to do so would require the practitioner to calculate the bootstrap/randomisation statistic across a large grid of values of θ_0. This would entail a high degree of computational burden, not required by the bootstrap method we have outlined in this paper. Consequently, while we recognise that a superior method to that outlined in this paper exists in theory, we believe that the formulation of the wild bootstrap we use in (3) is considerably more attractive from a practical perspective.

Remark 6. The following alternative construction of the bootstrap sample may be considered:

$$X_i^* = \bar{X}_n + (X_i - \hat{m}_n(X)) w_i \ , \ i = 1, ..., n, \tag{8}$$

where $\hat{m}_n(X)$ denotes the sample median of $X_1, ..., X_n$. The results in Theorem 1 hold also for this alternative design (the proof is omitted). However, in contrast to

[2] We are very grateful to one of the referees for pointing this out to us.

the standard design (3), where the bootstrap shocks satisfy $(X_i - \bar{X}_n)w_i = (Y_i - \bar{Y}_n)w_i$, where $\bar{Y}_n = O_p(n^{-1}a_n)$ diverges for $\alpha < 1$, in the case of (8) it holds that $(X_i - \hat{m}_n(X))w_i = (Y_i - \hat{m}_n(Y))w_i$, where $\hat{m}_n(Y) = O_p\left(n^{-1/2}\right)$ for any α. Hence, when $\alpha \leq 1$ we would recommend using the bootstrap approximation based on (8) as this would be expected to deliver more precise coverage rates in such cases. This is confirmed by the numerical results presented in the next section.

Remark 7. A further advantage of the wild bootstrap over methods aimed at estimating the (asymptotic) unconditional distribution of the sample mean is its robustness to cases where the data are independent but not identically distributed. For instance, consider the location model $X_i = \theta + c_i Y_i$ where the (fixed and finite) scale parameter c_i varies over i. It can be shown that result (6) of Theorem 1 holds provided that, e.g., there are finitely many regimes for c_i and the proportion of observations in each regime converges as $n \to \infty$. Moreover, the wild bootstrap procedure can also be shown to be robust to certain forms of structural changes in α. In contrast, none of the methods based on estimating the unconditional distribution of the mean are valid in cases where there is heterogeneity in the scale and/or tail thickness parameters.

3 Numerical Results

In this section we briefly investigate the finite sample properties of confidence sets built using the wild bootstrap approach described in the previous section. Accordingly, data are i.i.d. drawn from a symmetric stable distribution with tail index $\alpha \in \{0.5, 0.75, ..., 1.75, 2\}$. Notice that for $\alpha = 2$ the stable distribution reduces to the (light tailed) Gaussian distribution. Random stable variables are generated as in Chambers *et al.* (1976), using the Kiss+Monster function of Gauss 8.0. We consider samples of size $T \in \{20, 50, 100, 200, 500\}$. For all α and T, we report the coverage probability of a $100(1-\eta)\%$ (two-sided) confidence set of the form

$$[\bar{X}_n - (n^{-1}a_n)F_n^{*-1}(1-\eta/2), \bar{X}_n - (n^{-1}a_n)F_n^{*-1}(\eta/2)],$$

where $F_n^{*-1}(\cdot)$ is the bootstrap estimator of the (conditional) quantile function of $a_n^{-1}n\left(\bar{X}_n - \theta\right)$, see the previous section. The confidence level is set to $1 - \eta = 0.95$.

As is standard, for any x, $F_n^{*-1}(x)$ is approximated by repeating the wild bootstrap re-sampling procedure a large number, say B, of times; cf. Hansen (1996) and Andrews and Buchinsky (2000). Specifically, from the doubly independent sequence $\{\{w_{j,i}\}_{i=1}^n\}_{j=1}^B$, B (conditionally) independent bootstrap statistics, say $S_{n,j}^*$, $j = 1, ..., B$, are computed as above but from $X_{j,i}^* := \bar{X}_n + \left(X_i - \bar{X}_n\right)w_{j,i}$, $i = 1, ..., n$ (or from $X_{j,i}^* := \bar{X}_n + (X_i - \hat{m}_n(X))w_{j,i}$ if the modified wild bootstrap of Remark 5 is used). The bootstrap quantile function at x is then approximated by the percentage of bootstrap statistics which do not exceed x; that is, $\hat{F}_n^{*-1}(x) := B^{-1}\sum_{j=1}^B \mathbb{I}\left(S_{n,j}^* \leq x\right)$.

Coverage probabilities of 95% confidence sets ($\eta = 0.05$) are reported for both the wild bootstrap and the i.i.d. bootstrap. Two versions of the wild bootstrap are reported; the first is based on (3), and the second on deviations from the median as in (8) of Remark 5. Finally, for each of the two versions of the wild bootstrap, we also report the quartiles of the distribution of the ratio between the bootstrap confidence set length and the length of an *exact* confidence set based on the unconditional distribution of $\bar{X}_n$. Specifically, and with F_n denoting the unconditional distribution function of $a_n^{-1}n\left(\bar{X}_n - \theta\right)$, we consider the exact interval with coverage $1-\eta$: $[\bar{X}_n - (n^{-1}a_n)F_n^{-1}(1-\eta/2), \bar{X}_n - (n^{-1}a_n)F_n^{-1}(\eta/2)]$. The latter method, which requires the knowledge of the distribution function of X_i or (in order to be asymptotically valid) of the stability index α – neither of which is required by our proposed wild bootstrap procedures – is of course infeasible in practice but is useful in that it represents the best possible performance that could be obtained using any method based on the unconditional distribution; it therefore provides a useful benchmark. The number of Monte Carlo replications is set to 50,000 while the number of bootstrap replications is $B = 999$. The results are reported in Table 1.

We do not report results for consistent methods such as the m out of n or subsampling based approaches (these are reported elsewhere in the literature; see, for example, Cornea and Davidson, 2009). Notice, however, that in large samples confidence sets based on these methods will behave like confidence sets based on the unconditional distribution of the mean, and so the observations which follow regarding the behaviour of the wild bootstrap based method relative to the unconditional distribution will also apply (at least approximately) to the behaviour of the wild bootstrap relative to the m out of n and subsampling methods. Indeed, as noted above, these methods will be inferior to the exact confidence sets based on the unconditional distribution.

A number of observations can be made regarding the results in Table 1.

1. The standard (i.i.d.) bootstrap tends to be too liberal, increasingly so, other things being equal, as the tail index parameter, α, decreases, but is well behaved in the Gaussian case, $\alpha = 2$. For example, in the case where $\alpha = 0.5$ the coverage probability is only 89% even for a sample size of $T = 500$. This reflects the inability of the naive bootstrap to estimate the correct (conditional) distribution of the sample mean under infinite variance errors.

2. In contrast the two wild bootstrap methods deliver confidence sets with considerably more accurate coverage than the i.i.d. bootstrap in the infinite variance case, and give almost identical coverage rates to the i.i.d. bootstrap in the Gaussian case, $\alpha = 2$. The confidence set for the version of the wild bootstrap centred on the mean is slightly too conservative for the sample sizes considered for low values of α, with coverage probabilities increasing as α decreases, other things equal. These coverage rates rise to around 97% for $\alpha = 0.5$, but improve rapidly as α increases, and also improve as the sample size increases. The version of the wild bootstrap centred on the median, as in Remark 5, tends to be more accurate for $\alpha \leq 1$ than the wild bootstrap centred on the mean; indeed for $\alpha = 0.5$ the former has a coverage rate of 95% when $T = 500$. Where $\alpha > 1$ the two variants of the wild bootstrap appear almost equivalent.

3. The length of the wild bootstrap confidence sets are often significantly smaller, even dramatically so, than the length of the corresponding unconditional confidence

Table 1. Estimated Coverage Probabilities and length of Wild bootstrap confidence sets

		bootstrap algorithm								
		i.i.d.	wild (mean)				wild (median)			
α	T	cover.	cover.	conf. set length[a]			cover.	conf. set length[a]		
				25%	50%	75%		25%	50%	75%
0.5	20	0.877	0.975	0.003	0.009	0.046	0.948	0.002	0.008	0.039
	50	0.887	0.977	0.003	0.009	0.045	0.952	0.003	0.008	0.040
	100	0.891	0.977	0.003	0.009	0.043	0.949	0.003	0.008	0.040
	200	0.893	0.974	0.003	0.008	0.040	0.951	0.003	0.008	0.039
	500	0.893	0.971	0.003	0.008	0.041	0.950	0.003	0.008	0.040
0.75	20	0.887	0.955	0.020	0.043	0.118	0.938	0.020	0.042	0.108
	50	0.899	0.963	0.021	0.044	0.115	0.949	0.021	0.043	0.110
	100	0.904	0.964	0.021	0.043	0.113	0.952	0.021	0.043	0.110
	200	0.905	0.962	0.021	0.044	0.113	0.951	0.021	0.044	0.111
	500	0.906	0.96	0.021	0.043	0.111	0.951	0.021	0.043	0.110
1.00	20	0.899	0.943	0.096	0.159	0.306	0.934	0.096	0.158	0.295
	50	0.913	0.954	0.098	0.159	0.306	0.948	0.099	0.159	0.301
	100	0.913	0.953	0.100	0.162	0.308	0.948	0.100	0.161	0.305
	200	0.915	0.953	0.100	0.161	0.310	0.948	0.101	0.161	0.309
	500	0.915	0.952	0.101	0.161	0.303	0.950	0.101	0.161	0.302
1.25	20	0.908	0.936	0.196	0.275	0.439	0.931	0.198	0.277	0.433
	50	0.920	0.947	0.204	0.281	0.437	0.945	0.204	0.281	0.437
	100	0.925	0.950	0.205	0.283	0.440	0.949	0.206	0.284	0.439
	200	0.925	0.951	0.207	0.284	0.441	0.949	0.207	0.285	0.440
	500	0.925	0.950	0.207	0.285	0.439	0.948	0.207	0.284	0.439
1.50	20	0.917	0.931	0.397	0.503	0.687	0.930	0.401	0.506	0.688
	50	0.931	0.946	0.413	0.511	0.691	0.944	0.415	0.513	0.693
	100	0.932	0.948	0.418	0.514	0.694	0.947	0.419	0.515	0.695
	200	0.932	0.946	0.421	0.518	0.695	0.946	0.422	0.518	0.696
	500	0.934	0.948	0.422	0.518	0.691	0.949	0.422	0.518	0.693
1.75	20	0.923	0.927	0.591	0.693	0.835	0.929	0.597	0.699	0.842
	50	0.936	0.943	0.625	0.708	0.840	0.942	0.627	0.711	0.844
	100	0.939	0.946	0.635	0.712	0.841	0.946	0.637	0.714	0.842
	200	0.940	0.946	0.641	0.713	0.840	0.947	0.642	0.714	0.841
	500	0.942	0.949	0.643	0.714	0.840	0.949	0.644	0.714	0.841
2.00	20	0.929	0.927	0.843	0.948	1.057	0.929	0.853	0.959	1.068
	50	0.939	0.939	0.912	0.980	1.052	0.940	0.917	0.986	1.058
	100	0.945	0.945	0.940	0.991	1.044	0.946	0.943	0.994	1.047
	200	0.947	0.948	0.960	0.998	1.038	0.948	0.961	1.000	1.040
	500	0.949	0.950	0.972	1.000	1.031	0.949	0.972	1.002	1.031

Notes: [a] these are the quartiles of the ratio between the length of the bootstrap confidence sets and the length of an exact (unconditional) confidence set.

sets. Taking the case of $\alpha = 0.5$ and $T = 500$, here 50% of the time the length of the wild bootstrap confidence set does not exceed 0.8% of the length of the corresponding exact unconditional confidence set. Similarly, 75% of the time the length of the wild bootstrap confidence set does not exceed 4% of the length of the exact unconditional set. Even for the larger values of α considered the wild bootstrap confidence sets are much narrower than the corresponding unconditional sets. Notice also that the interval length (as a fraction of the length of the unconditional sets) narrows, other things being equal, as α decreases.

The result discussed in 3 is particularly important, since it emphasises that, for infinite variance sequences, conditional inference (in this case, inference conditionally on $|X_i|$, $i = 1, ..., n$) is clearly preferable to unconditional inference, be the latter based on the unconditional distribution or on the m out of n or subsampling methods.

4 Conclusions

In this paper we have focused attention on the problem of bootstrapping the sample mean in a location model with symmetric infinite variance (α-stable) errors. Although the i.i.d. bootstrap delivers narrower confidence sets than those based on the asymptotic (unconditional) distribution (the latter coinciding with those based on either the m out of n or subsampling methods in large samples), it does not deliver the desired nominal coverage probability, even in the limit. We have proposed a new procedure, based on the wild bootstrap method, and have demonstrated analytically that it delivers correct (asymptotic) coverage probabilities. Monte Carlo experiments were also reported which suggest the proposed method performs very well in practice, delivering coverage rates very close to the nominal level even in relatively small samples and displaying, often very substantially, narrower confidence sets than the corresponding (infeasible) sets based on the exact unconditional distribution. This is especially useful in practice in the light of the well-known unreliability of feasible methods based on the m out of n and subsampling procedures in small samples, where coverage rates can often be a very long way from the nominal probability; see Cornea and Davidson (2009).

While bootstrapping the sample mean in a location model may appear to be a somewhat limited problem, it is nonetheless important in that it provides the basis for further research developing corresponding bootstrap procedures in more advanced regression designs in both cross-sectional and time-series models with infinite-variance errors. In the latter case, allowing for temporal dependence in the bootstrap sample, as in the recent generalisation of the wild bootstrap by Shao (2010), may well be a useful approach to investigate. Further research is also warranted to relax the assumption of symmetric errors that we have made in this paper. In such a case the wild bootstrap outlined in this paper is no longer asymptotically valid. However, we conjecture that using random variables w_i in (3) such that $P(w_i = 1) = \hat{\pi}\left(|X_i - \bar{X}_n|\right)$, where $\hat{\pi}(\cdot)$ is a suitable estimator of the function π defined as $\pi(x) = P(X_1 - \theta > 0 | |X_i - \theta| = x)$ for $x \geq 0$, will deliver a version of our proposed bootstrap which performs well. Clearly, in the symmetric case $\pi(x) = 0.5$ identically for $x > 0$, but it is likely to be a non-constant function for asymmetric data. A detailed exploration of such extensions is beyond the scope of the present paper

A Appendix

Proof of Theorem 1

Let $Z_{n1}, ..., Z_{nn}$ denote the order statistics of $|Y_1|, ..., |Y_n|$, and let $\delta_{n1}, ..., \delta_{nn} \in \{-1, 1\}$ be such that $Z_{n1}\delta_{n1}, ..., Z_{nn}\delta_{nn}$ is a permutation of $Y_1, ..., Y_n$. LePage, Woodroofe and Zinn (1981) prove that $a_n^{-1}(Z_{n1}, ..., Z_{nn}, 0, ...) \xrightarrow{w} Z$ and $(\delta_{n1}, ..., \delta_{nn}, 0, ...) \xrightarrow{w} \delta$ in the product space $\mathbb{R}^\infty$, with Z and δ introduced earlier. Using Knight's (1989) method of proof for his Theorem 2, it obtains that

$$\mathcal{L}\left(a_n^{-1}\sum_{i=1}^{n} Y_i w_i \middle| \{Y_i\}_{i=1}^n\right) \xrightarrow{w} \mathcal{L}\left(\sum_{i=1}^{\infty} Z_i\delta_i w_i \middle| Z, \delta\right), \tag{A.1}$$

where the limit is in the sense of weak convergence of random measures (we omit the details as they are very similar to Knight's). By the independence of Z, δ and $\{w_i\}_{i\in\mathbb{N}}$, with both δ and $\{w_i\}_{i\in\mathbb{N}}$ i.i.d. Rademacher sequences, it holds that (i) $(\sum_{i=1}^\infty Z_i\delta_i w_i, Z)$ is independent of δ, and (ii) $\{\delta_i w_i\}_{i\in\mathbb{N}}$ is distributed like δ and the two are jointly independent of Z. These properties justify respectively the first and the second of the following equalities: for every $x \in \mathbb{R}$,

$$P\left(\sum_{i=1}^{\infty} Z_i\delta_i w_i \le x \middle| Z, \delta\right) = P\left(\sum_{i=1}^{\infty} Z_i\delta_i w_i \le x \middle| Z\right) = P\left(\sum_{i=1}^{\infty} Z_i\delta_i \le x \middle| Z\right),$$

P-a.s. Together with (A.1), this proves (5).

We proceed with the proof of (6). Weak convergence of random measures on $\mathbb{R}$ implies weak convergence of the associated cumulative processes on $D(\mathbb{R})$, the Skorokhod space with metric defined by compounding the metrics over the compact sets $[-k, k]$ $(k \in \mathbb{N})$; see the discussion in Daley and Vere-Jones (2008, pp.143-44). Hence, from (A.1) we can conclude that $\tilde{F}_n \xrightarrow{w} F^*$ in $D(\mathbb{R})$, where $\tilde{F}_n$ and F^* are the cumulative processes associated with the random measures $\mathcal{L}(a_n^{-1}\sum_{i=1}^n Y_i w_i | \{Y_i\}_{i=1}^n)$ and $\mathcal{L}(\sum_{i=1}^\infty Z_i\delta_i | Z)$, respectively. Next,

$$a_n^{-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}_n\right) w_i = a_n^{-1}\sum_{i=1}^{n} Y_i w_i - \left(n a_n^{-1}\bar{Y}_n\right)\bar{w}_n = a_n^{-1}\sum_{i=1}^{n} Y_i w_i + o_{P_*}(1) \;\; P\text{-a.s.},$$

since $n a_n^{-1}\bar{Y}_n = o\left(n^{1/(4+\alpha)}\right)$ P-a.s. and $\bar{w}_n = n^{-1}\sum_{i=1}^n w_i = O_{P_*}(n^{-1/2})$ (resp., by Lemma 1 of Chan and Tran, 1989, and the CLT). Therefore, the Skorokhod distance between F_n^* and $\tilde{F}_n$ vanishes P-a.s. Since $\tilde{F}_n \xrightarrow{w} F^*$ in $D(\mathbb{R})$, it follows that $(\tilde{F}_n, F_n^*) \xrightarrow{w} (1, 1)F^*$ in the product space $D(\mathbb{R})^2$. A theorem of Lévy (1931), cf. Theorem 3 of Knight (1989), yields that F^* has continuous sample paths P-a.s. So, by the continuous mapping theorem, $\sup_{x\in K}|\tilde{F}_n(x) - F_n^*(x)| \xrightarrow{P} 0$ for every compact $K \subset \mathbb{R}$. This fact together with $S_n = a_n^{-1}\sum_{i=1}^n Y_i = O_P(1)$, see (1), yields that

$$F_n^*(S_n) = \tilde{F}_n(S_n) + o_P(1). \tag{A.2}$$

Let $\tilde{F}_n^{-1}$ be defined analogously to F_n^{*-1}, so for every $x \in \mathbb{R}$ and $\eta \in (0,1)$,

$$\tilde{F}_n(x) \geq \eta \text{ iff } x \geq \tilde{F}_n^{-1}(\eta). \tag{A.3}$$

For $\eta \in (0,1)$, by iterating expectations,

$$\begin{aligned} P\left(\tilde{F}_n(S_n) \geq \eta\right) &= E\left[P\left(\tilde{F}_n(S_n) \geq \eta \middle| \{|Y_i|\}_{i=1}^n\right)\right] \\ &= E\left[P\left(S_n \geq \tilde{F}_n^{-1}(\eta) \middle| \{|Y_i|\}_{i=1}^n\right)\right]. \end{aligned} \tag{A.4}$$

Note that $\tilde{F}_n$ has a $\sigma(\{|Y_i|\}_{i=1}^n)$-measurable version: for every $x \in \mathbb{R}$, with $\sigma_i := \operatorname{sgn}(Y_i)$ $(i = 1, ..., n)$,

$$\begin{aligned} \tilde{F}_n(x) &= P\left(a_n^{-1}\sum_{i=1}^n |Y_i|\sigma_i w_i < x \middle| \{|Y_i|, \sigma_i\}_{i=1}^n\right) \\ &= P\left(a_n^{-1}\sum_{i=1}^n |Y_i|\sigma_i w_i < x \middle| \{|Y_i|\}_{i=1}^n\right) \; P\text{-a.s.}, \end{aligned}$$

where the latter conditional probability is $\sigma(\{|Y_i|\}_{i=1}^n)$-measurable; $\{\sigma_i\}_{i=1}^n$ was removed from the condition using the independence of $(a_n^{-1}\sum_{i=1}^n |Y_i|\sigma_i w_i, \{|Y_i|\}_{i=1}^n)$ and $\{\sigma_i\}_{i=1}^n$. Considering a $\sigma(\{|Y_i|\}_{i=1}^n)$-measurable version of $\tilde{F}_n$, also $\tilde{F}_n^{-1}(\eta)$ will be $\sigma(\{|Y_i|\}_{i=1}^n)$-measurable by (A.3). Further,

$$\begin{aligned} P\left(S_n \geq \tilde{F}_n^{-1}(\eta) \middle| \{|Y_i|\}_{i=1}^n\right) &= P\left(a_n^{-1}\sum_{i=1}^n |Y_i|\sigma_i \geq \tilde{F}_n^{-1}(\eta) \middle| \{|Y_i|\}_{i=1}^n\right) \\ &\overset{a.s.}{=} P\left(a_n^{-1}\sum_{i=1}^n |Y_i|\sigma_i w_i \geq \tilde{F}_n^{-1}(\eta) \middle| \{|Y_i|\}_{i=1}^n\right) \\ &\overset{a.s.}{=} P\left(a_n^{-1}\sum_{i=1}^n |Y_i|\sigma_i w_i \geq \tilde{F}_n^{-1}(\eta) \middle| \{|Y_i|, \sigma_i\}_{i=1}^n\right) \\ &= P\left(a_n^{-1}\sum_{i=1}^n Y_i w_i \geq \tilde{F}_n^{-1}(\eta) \middle| \{Y_i\}_{i=1}^n\right) \\ &= 1 - \tilde{F}_n(\tilde{F}_n^{-1}(\eta)-), \end{aligned}$$

where the second, third and fifth inequality rely on the $\sigma(\{|Y_i|\}_{i=1}^n)$-measurability of $\tilde{F}_n^{-1}(\eta)$, the second one also on the fact that $\{\sigma_i\}_{i=1}^n$ and $\{\sigma_i w_i\}_{i=1}^n$ are equidistributed and independent of $\{|Y_i|\}_{i=1}^n$, the third one on the independence of $(a_n^{-1}\sum_{i=1}^n |Y_i|\sigma_i w_i,$

$\tilde{F}_n^{-1}(\eta), \{|Y_i|\}_{i=1}^n)$ and $\{\sigma_i\}_{i=1}^n$, and the fifth one on (A.3). By combining the previous display with (A.4), we find that

$$P\left(\tilde{F}_n\left(S_n\right) \geq \eta\right) = 1 - E\left[\tilde{F}_n(\tilde{F}_n^{-1}(\eta)-)\right].$$

It holds that

$$\left|\tilde{F}_n(\tilde{F}_n^{-1}(\eta)-) - \eta\right| \leq \sup_{x\in\mathbb{R}} |\tilde{F}_n(x) - \tilde{F}_n(x-)| \xrightarrow{P} 0$$

by the weak convergence of $\tilde{F}_n$ to the process $\tilde{F}$ with a.s. continuous trajectories. The relation $\tilde{F}_n(\tilde{F}_n^{-1}(\eta)-) \xrightarrow{P} \eta$ can be integrated by the bounded convergence theorem to get $E\left[\tilde{F}_n(\tilde{F}_n^{-1}(\eta)-)\right] \to \eta$, which by the arbitrariness of $\eta \in (0,1)$ implies that $\tilde{F}_n\left(S_n\right) \xrightarrow{w} U[0,1]$. In view of (A.2), this completes the proof of (6).

Let now $\eta \in (0,1)$ be given. Under $H_0 : \theta = \theta_0$, by (A.3) for F_n^* it holds that

$$\begin{aligned} P\left(a_n^{-1} n(\bar{X}_n - \theta_0) < F_n^{*-1}(\eta)\right) &= P\left(S_n < F_n^{*-1}(\eta)\right) \\ &= P(F_n^*(S_n) < \eta) \to \eta, \end{aligned}$$

the convergence by (6). This means that the test of H_0 against $H_1 : \theta < \theta_0$ which rejects for $a_n^{-1} n(\bar{X}_n - \theta_0) < F_n^{*-1}(\eta)$ has asymptotically correct size. Tests against $H_1 : \theta > \theta_0$ and $H_1 : \theta \neq \theta_0$ can be considered similarly.

If $\zeta \in (0, \eta)$, the confidence interval $I = \left[\bar{X}_n - a_n n^{-1} F_n^{*-1}(1-\eta+\zeta), \bar{X}_n - a_n n^{-1} F_n^{*-1}(\zeta)\right]$ for θ has asymptotic confidence level $1-\eta$ because

$$\begin{aligned} P\left(\theta \in I\right) &= P\left(\theta \leq \bar{X}_n - a_n n^{-1} F_n^{*-1}(\zeta)\right) - P\left(\theta < \bar{X}_n - a_n n^{-1} F_n^{*-1}(1-\eta+\zeta)\right) \\ &= P\left(S_n \geq F_n^{*-1}(\zeta)\right) - P\left(S_n > F_n^{*-1}(1-\eta+\zeta)\right), \end{aligned}$$

where $P\left(S_n \geq F_n^{*-1}(\zeta)\right) = P\left(F_n^*(S_n) \geq \zeta\right) \to 1-\zeta$ by (6), and for every $\omega \in (0, \eta-\zeta)$,

$$\begin{aligned} P\left(S_n > F_n^{*-1}(1-\eta+\zeta)\right) &\in \left[P\left(S_n \geq F_n^{*-1}(1-\eta+\zeta+\omega)\right), P\left(S_n \geq F_n^{*-1}(1-\eta+\zeta)\right)\right] \\ &= \left[P\left(F_n^*(S_n) \geq 1-\eta+\zeta+\omega\right), P\left(F_n^*(S_n) \geq 1-\eta+\zeta\right)\right] \\ &\to [\eta-\zeta-\omega, \eta-\zeta] \end{aligned}$$

by (6), so $P\left(S_n > F_n^{*-1}(1-\eta+\zeta)\right) \to \eta - \zeta$. ■

References

Abadir, K. and J. Magnus (2003), The central limit theorem for Student's distribution, *Econometric Theory* 19, 1195–1195; Solution in *Econometric Theory* 20, 1261–1264.

Andrews, D.W.K., and M. Buchinsky (2000), A three-step method for choosing the number of bootstrap repetitions, *Econometrica* 68, 23–51.

Arcones M.A. and E. Giné (1989), The bootstrap of the mean with arbitrary bootstrap sample size, *Annales Institut Henri Poincare Probabilites et Statistiques B* 25, 457–481;

Arcones M.A. and E. Giné (1991), Additions and corrections to "The bootstrap of the mean with arbitrary bootstrap sample size", *Annales Institut Henri Poincare Probabilites et Statistiques B* 27, 583–585.

Athreya K.B. (1987), Bootstrap of the mean in the infinite variance case, *Annals of Statistics* 15, 724–731.

Bertail, P., D.N.Politis and J.P.Romano (1999), On subsampling estimators with unknown rate of convergence, *Journal of the American Statistical Association* 94, 569–579.

Chambers J.M., C. L. Mallows and B.W. Stuck (1976), A method for simulating stable random variables, *Journal of the American Statistical Association* 71, 340–344.

Chan, N.H. and L.T. Tran (1989), On the first-order autoregressive process with infinite variance, *Econometric Theory* 5, 354-362.

Cornea, A. and R. Davidson (2009), A refined bootstrap for heavy tailed distributions, mimeo. Downloadable from `http://russell-davidson.arts.mcgill.ca/articles/`

Daley D.J. and D. Vere-Jones (2008), *An Introduction to the Theory of Point Processes*, Springer.

Feller, W. (1971), *An Introduction to Probability Theory and Its Applications*, Vol. 2. 2nd ed. New York: Wiley.

Hall P. (1992), Asymptotic properties of the bootstrap for heavy-tailed distributions, *Annals of Probability* 18, 1342–1360.

Hall, P. and Q.Yao (2003), Inference in ARCH and GARCH models with heavy-tailed errors, *Econometrica* 71, 285–317.

Hansen, B.E. (1996), Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* 64, 413–430.

Knight K. (1989), On the bootstrap of the mean in the infinite variance case, *Annals of Statistics*, 17, 1168–1175.

Lehmann, E. L. and J.P. Romano (2005), *Testing Statistical Hypotheses*, Springer, New York, Third edition.

LePage R., M. Woodroofe and J. Zinn (1981), Convergence to a stable distribution via order statistics, *Annals of Probability* 9, 624–632.

Lévy P. (1931), Sur les séries dont les termes sont des variables êventuelles indépendanes, *Studia Math.* 3, 119–155.

Liu, R.Y. (1988), Bootstrap procedures under some non i.i.d. models. *Annals of Statistics* 16, 1696–1708.

Mammen, E. (1993), Bootstrap and wild bootstrap for high dimensional linear models, *Annals of Statistics* 21, 255–285.

Politis, D.N. and J.P.Romano (1994), Large sample confidence regions based on subsamples under minimal assumptions, *Annals of Statistics* 22, 2031–2050

Politis, D.N., J.P. Romano and M. Wolf (1999), *Subsampling*, Springer-Verlag, New York.

Resnick, S. (1997), Heavy tailed modeling and teletraffic data. *Annals of Statistics* 25, 1805–1869.

Romano J.P. and M. Wolf (1999), Subsampling inference for the mean in the heavy-tailed case, *Metrika* 50, 55–69.

Shao, X. (2010) The dependent wild bootstrap, *Journal of the American Statistical Association* 105, 218—235.

Wu, C.F.J. (1986), Jackknife, bootstrap, and other resampling methods, *Annals of Statistics* 14, 1261–1295.