The Annals of Statistics
1989, Vol. 17, No. 2, 684–691

NECESSARY CONDITIONS FOR THE BOOTSTRAP OF THE MEAN

BY EVARIST GINÉ[1] AND JOEL ZINN[2]

CUNY, College of Staten Island and Texas A&M University, and Texas A & M University

It is proved that $EX^2 < \infty$ is necessary for a very mild form of the bootstrap of the mean to work a.s. and that X must be in the domain of attraction of the normal law if a.s. is weakened to "in probability."

1. Introduction. Let X be a real valued random variable and let X_i, $i \in \mathbb{N}$, be independent, identically, distributed copies of X, with $\mathscr{L}(X) = P$. Let

$$P_n(\omega) = n^{-1} \sum_{i=1}^{n} \delta_{X_i(\omega)}, \qquad n \in \mathbb{N}, \omega \in \Omega,$$

be the empirical measures associated to the sequence $\{X_i(\omega)\}$. For $n \in \mathbb{N}$ and $\omega \in \Omega$, let $\{\hat{X}^{\omega}_{nj}\}_{j=1}^{n}$ be i.i.d. random variables with law $P_n(\omega)$ and let $\overline{X}_n(\omega)$ be the sample mean of $\{X_i(\omega)\}_{i=1}^{n}$, $n \in \mathbb{N}$. Since $P_n(\omega)$ is close to P, we expect that for many statistics H_n, $H_n(X_1, \dots, X_n; P)$ is close in distribution to the bootstrap statistic $\hat{H}_n(\omega) = H_n(\hat{X}^{\omega}_{n1}, \dots, \hat{X}^{\omega}_{nn}; P_n(\omega))$ ω-almost surely or at least in probability. This is, very roughly, the idea of the bootstrap. [See Efron (1979), where this nice idea is made explicit and where it is substantiated with several important examples.] The "bootstrap principle" does not always hold true and it is important to determine its domain of validity. In this note we do precisely this for the simplest of all statistics, $H_n(X_1, \dots, X_n; P) = \sum_{i=1}^{n}(X_i - EX)/a_n$. Bickel and Freedman (1981) and Singh (1981) showed that H_n and $\hat{H}_n(\omega)$ are asymptotically close a.s. with $a_n = n^{1/2}$ and $EX^2 < \infty$. [See Giné and Zinn (1988) for an analogous result for empirical processes.] Athreya (1986) proved that they are asymptotically close in probability if X is in the domain of attraction of the normal law and a_n are the normalizing constants in the clt for X. [See Csörgő and Mason (1988) for an empirical process analogue.] In this note we show that *even* for the existence of a sequence $a_n \to \infty$, random variables $c_n(\omega)$ and a random measure $\mu(\omega)$ such that

$$\mathscr{L}\left(\sum_{j=1}^{n} \hat{X}^{\omega}_{nj}/a_n - c_n(\omega) \right) \to_w \mu(\omega) \quad \text{a.s.},$$

it is necessary that $EX^2 < \infty$, and that if a.s. is relaxed to "in probability," then it is necessary that X be in the domain of attraction of the normal law. This shows for instance that Athreya's (1986) result for $EX^2 = \infty$, X in the domain

Received June 1988; revised August 1988.
[1]Research partially supported by NSF Grant DMS-86-19411.
[2]Research partially supported by NSF Grant DMS-86-01250.
AMS 1980 *subject classifications.* Primary 62E20, 62F12; secondary 60F05.
Key words and phrases. Bootstrap (of the mean), central limit theorem.

of a normal law, can not be improved to an almost sure statement. Similarly, it can be shown that in the p-stable domain of attraction with $p < 2$, Athreya's (1987) result on weak convergence of the distribution function of the bootstrap, cannot be improved to convergence in probability (see Remark 4). Our results are also related to some of the comments in Hartigan (1986).

In what follows we will use the notation set up at the beginning of this introduction.

2. Results and proofs.

THEOREM 1. *If there exist random variables $c_n(\omega)$, $n \in \mathbb{N}$, a strictly increasing sequence $a_n \to \infty$ and a random probability measure $\mu(\omega)$ nondegenerate with positive probability, such that*

$$\mathscr{L}\left(\sum_{j=1}^{n} \hat{X}_{nj}^{\omega}/a_n - c_n(\omega)\right) \to_w \mu(\omega) \qquad \omega\text{-}a.s.,$$

then

$$a_n \simeq n^{1/2}, \qquad EX^2 < \infty$$

and

$$\mathscr{L}\left(\sum_{j=1}^{n} \left(\hat{X}_{nj}^{\omega} - \overline{X}_n(\omega)\right)/n^{1/2}\right) \to_w N(0, \operatorname{Var} X) \qquad \omega\text{-}a.s.$$

PROOF. We show first that

$$\mathscr{L}\left(\sum_{j=1}^{n} \left(\hat{X}_{nj}^{\omega} - \overline{X}_n(\omega)\right)/a_n\right) \to_w N(0, \sigma^2) \qquad \omega\text{-a.s.} \tag{1}$$

for some nonrandom $\sigma^2 > 0$. The system $\{\hat{X}_{nj}^{\omega}/a_n,\ j = 1, \dots, n\}_{n=1}^{\infty}$ is infinitesimal with probability 1 since for all $\varepsilon > 0$,

$$P_n\{|\hat{X}_{n1}^{\omega}/a_n| > \varepsilon\} = \sum_{i=1}^{n} I(|X_i| > \varepsilon a_n)/n \to 0 \quad \text{a.s.}$$

by the law of large numbers. Hence, $\mu(\omega)$ is a.s. an infinitely divisible measure. Let $\pi(\omega)$ be the Lévy measure of $\mu(\omega)$ and for each $\delta > 0$, $\delta \in \mathbb{Q}$, let h_δ be a bounded continuous function on $\mathbb{R}^+$, zero on $[0, \delta/2]$ and one on $[\delta, \infty)$. Let $\pi_\delta(dx, \omega) = h_\delta(x)\pi(dx, \omega)$. By the converse central limit theorem in $\mathbb{R}$ [e.g., Araujo and Giné (1980), Chapter 2] we have

$$\sum_{i=1}^{n} h_\delta(X_i(\omega)/a_n)\delta_{X_i(\omega)/a_n} \to_w \pi_\delta(\omega) \qquad \omega\text{-a.s.}$$

Let $\mathscr{F}$ be a countable measure determining set of bounded continuous functions, e.g., $\mathscr{F} = \{x \to e^{itx},\ t \in \mathbb{Q}\}$. Then

$$\sum_{i=1}^{n} h_\delta(X_i(\omega)/a_n) f(X_i(\omega)/a_n) \to \int f \pi_\delta(dx, \omega) \qquad \forall f \in \mathscr{F}, \forall \delta \in \mathbb{Q}^+,\ \omega\text{-a.s.}$$

But since $a_n \to \infty$, $\int f \pi_\delta(dx, \omega)$ is a tail random variable with respect to $\{X_i\}$. Hence $\int f \pi_\delta(dx, \omega) = \text{constant} \ \forall\ f \in \mathscr{F}$, $\forall\ \delta \in \mathbb{Q}^+$, ω-a.s. Since $\mathscr{F}$ is measure determining there is a fixed nonrandom measure π_δ such that $\pi_\delta(\omega) = \pi_\delta$, $\delta \in \mathbb{Q}^+$, ω-a.s. This shows that there is a nonrandom Lévy measure π such that

$$\pi(\omega) = \pi \qquad \omega\text{-a.s.}$$

Then, again by the converse clt, there is a countable set $D \subset \mathbb{R}^+$ such that

$$\sum_{i=1}^{n} I(|X_i(\omega)| > \lambda a_n) \to \pi(\lambda, \infty) \qquad \forall\ \lambda \in D,\ \omega\text{-a.s.} \tag{2}$$

$\pi(\lambda, \infty)$ takes on only nonnegative integer values. Assume that, for some $\lambda \in D$, $\pi(\lambda, \infty) = r \neq 0$. Then

$$\sum_{i=1}^{n} I(|X_i(\omega)| > \lambda a_n) = r \quad \text{eventually a.s.} \tag{3}$$

by (2). Therefore

$$\lim_{n \to \infty} P\left\{ \sum_{i=1}^{n} I(|X_i| > \lambda a_n) = r \right\} = 1 \tag{4}$$

[since $P\{\cup_{m=1}^{\infty} \cap_{n=m}^{\infty} \{\sum_{i=1}^{n} I(|X_i| > \lambda a_n) = r\}\} = 1$]. On the other hand,

$$\begin{aligned} &\lim_{n \to \infty} P\left\{ \sum_{i=1}^{n} I(|X_i| > \lambda a_n) = r \right\} \\ &\quad = \lim_{n \to \infty} \binom{n}{r} P\{|X| > \lambda a_n\}^r P\{|X| \le \lambda a_n\}^{n-r} \\ &\quad = r^r e^{-r}/r! < 1. \end{aligned} \tag{5}$$

To see this recall that if the partial sums of a triangular array of row-wise independent uniformly bounded random variables converge in law, then the expected values of the partial sums converge to the expected value of the limit [since, e.g., Hoffmann-Jørgensen's (1974) inequality provides uniform integrability; see de Acosta and Giné (1979), Theorem 2.1 or 3.3]. Then, this remark applied to (3) gives $nP\{|X| > \lambda a_n\} \to r$ and it follows from this that $(1 - P\{|X| > \lambda a_n\})^n \to e^{-r}$. These two limits yield (5).

The limits (4) and (5) are in contradiction. Therefore $\pi(\lambda, \infty) = 0\ \forall\ \lambda \in D$, that is,

$$\pi = 0. \tag{6}$$

Also, (3) becomes

$$\sum_{i=1}^{n} I(|X_i| > \lambda a_n) = 0 \quad \text{eventually, a.s.} \tag{3'}$$

for all $\lambda > 0$, so that for all $\lambda > 0$ and $p \in \mathbb{R}$,

$$\sum_{i=1}^{n} |X_i|^p I(|X_i| > \lambda a_n) = 0 \quad \text{eventually, a.s.} \tag{7}$$

Suppose now that $\sigma(\omega)$ is the (nonzero with positive probability) standard deviation of the normal component of $\mu(\omega)$. The truncated variances necessary condition for the clt [e.g., Araujo and Giné (1980)] becomes, in view of (7), a variance condition, namely

$$\lim_{n\to\infty}\left[\sum_{i=1}^{n}(X_i(\omega))^2/a_n^2-\left(\sum_{i=1}^{n}X_i(\omega)/a_n n^{1/2}\right)^2\right]=\sigma^2(\omega)\qquad \omega\text{-a.s.} \tag{8}$$

In particular $\sigma^2(\omega)$ is a tail random variable and therefore

$$\sigma^2(\omega)=\sigma\neq 0\qquad \omega\text{-a.s.}, \tag{9}$$

for some $\sigma \neq 0$. Then, (2), (6), (8) and (9) give, by the central limit theorem,

$$\mathscr{L}\left(\sum_{j=1}^{n}\left(\hat{X}_{nj}^{\omega}-n^{-1}\sum_{i=1}^{n}X_i(\omega)I(|X_i(\omega)|\leq a_n)\right)/a_n\right)\to_w N(0,\sigma^2)\qquad \omega\text{-a.s.},$$

but by (7) the truncated $P_n(\omega)$-expectation of $\hat{X}_{nj}^{\omega}$ can be replaced by its $P_n(\omega)$-expectation $\overline{X}_n(\omega)$ and (1) is proved.

If $EX^2 < \infty$, then (8) and the law of large numbers yield

$$n/a_n^2\to\sigma^2/\operatorname{Var}X, \tag{10}$$

that is, $a_n \asymp n^{1/2}$. So, we can assume $EX^2 = \infty$. In this case we can simplify (8) by means of the following result about comparison of empirical moments.

LEMMA 2. *If* $E|X|^p = \infty$ $(p > 0)$ *and* $0 < p' < p$, *then*

$$\frac{\left(\sum_{i=1}^{n}|X_i|^{p'}/n\right)^{p}}{\left(\sum_{i=1}^{n}|X_i|^{p}/n\right)^{p'}}\to 0\quad a.s. \tag{11}$$

PROOF OF LEMMA 2. By Hölder's inequality,

$$\begin{aligned}\sum_{i=1}^{n}|X_i|^{p'}/n &\leq a^{p'}+\sum_{i=1}^{n}|X_i|^{p'}I(|X_i|>a)/n\\ &\leq a^{p'}+\left(\sum_{i=1}^{n}|X_i|^{p}/n\right)^{p'/p}\left(\sum_{i=1}^{n}I(|X_i|>a)/n\right)^{1-p'/p}.\end{aligned}$$

Now the lemma follows from the law of large numbers (both for finite and infinite moments) on dividing by $(\sum_{i=1}^{n}|X_i|^p/n)^{p'/p}$ and taking limits first as $n \to \infty$ and then as $a \to \infty$. □

Now, by Lemma 2 with $p = 2$ and $p' = 1$, (8) and (9) become

$$\lim_{n\to\infty}\sum_{i=1}^{n}X_i^2/a_n^2=\sigma^2\neq 0\quad \text{a.s.} \tag{12}$$

By (7) we can truncate in (12) and then take expectations [by boundedness of the summands, as in (5)] to obtain

$$na_n^{-2}EX^2I(|X|\leq a_n)\to\sigma^2\neq 0. \tag{13}$$

We will obtain a contradiction to $\sigma^2 \neq 0$ from (13) and (3′). By (3′), $|X_n|/a_n \leq 1$ eventually a.s. and therefore the Borel–Cantelli lemma gives

$$\sum_{n=1}^{\infty} P\{|X| > a_n\} < \infty. \tag{14}$$

Since $na_n^{-2} \to 0$ by (13) (recall that we are assuming $EX^2 = \infty$) there is $r_n \to \infty$, $r_n < n$, such that

$$\delta_n := na_n^{-2} \max_{k \leq r_n} a_k^2 \to 0.$$

Hence, using (13) once more we obtain

$$\begin{aligned}\sigma^2 &\leq \lim_{n\to\infty} na_n^{-2} \sum_{k=1}^{n} a_k^2 P\{a_{k-1} < |X| \leq a_k\}\\ &\leq \lim_{n\to\infty} \delta_n + \limsup_{n\to\infty} na_n^{-2}\Big[\max_{k\leq n}\big(a_k^2/k\big)\Big] \sum_{k=r_n}^{n} kP\{a_{k-1} < |X| \leq a_k\}\\ &= 0\end{aligned}$$

since $\limsup_{n\to\infty} na_n^{-2}[\max_{k\leq n}(a_k^2/k)] < \infty$ by (13) and the hypothesis $a_n \nearrow \infty$, and $\sum_{k=r_n}^{\infty} kP\{a_{k-1} < |X| \leq a_k\} \to 0$ by (14). We thus have a contradiction with (9). Therefore, $EX^2 < \infty$ and $a_n \simeq n^{1/2}$.

Finally, taking $a_n = n^{1/2}$ in (8), the law of large numbers gives $\sigma^2 = \operatorname{Var} X$ and (1) becomes

$$\mathscr{L}\left(\sum_{j=1}^{n}\big(\hat{X}_{nj}^{\omega} - \overline{X}_n(\omega)\big)/n^{1/2}\right) \to_w N(0, \operatorname{Var} X) \quad \text{a.s.} \qquad \square$$

In connection with Theorem 1 it is worth mentioning that Csörgő and Mason (1988) have recently remarked that a.s. weak convergence to $N(0,1)$ of $\mathscr{L}(\sum_{j=1}^{n}(\hat{X}_{nj}^{\omega} - \overline{X}_n(\omega))/s_n(\omega))$, where s_n is the sample s.d., takes place only if $EX^2 < \infty$.

Consider now $(\mathscr{P}(\mathbb{R}), w)$, the set of probability measures on $\mathbb{R}$ with the weak topology. This is a Polish space. We say that $\mu_n(\omega) \to_w \mu(\omega)$ in probability if $d(\mu_n, \mu) \to 0$ in probability for some distance metrizing weak convergence. This definition does not depend on the distance used because $d(\mu_n, \mu) \to 0$ in pr if and only if every subsequence has a further subsequence for which convergence takes place a.s. and if for ω fixed $d(\mu_{n'}(\omega), \mu(\omega)) \to 0$, then $\bar{d}(\mu_{n'}(\omega), \mu(\omega)) \to 0$ for any other distance $\bar{d}$ metrizing weak convergence.

Passing to a.s. convergence along subsequences will allow us to use the methods of Theorem 1 to prove the next result. Before stating it we note that the random measure $\mu_n(\omega) = \mathscr{L}(\sum_{j=1}^{n}\hat{X}_{nj}^{\omega}/a_n - c_n(\omega))$, where c_n is a random variable, is a true random variable with values in $(\mathscr{P}(\mathbb{R}), w)$: The preimage of any weak neighborhood of any probability measure by μ_n is measurable [for this it suffices to check that the random quantity $E_{P_n(\omega)} f(\sum_{j=1}^{n}\hat{X}_{nj}^{\omega}/a_n - c_n(\omega))$ is a random variable for every f bounded and continuous, which is obviously true].

THEOREM 3. *If there exist random variables $c_n(\omega)$, $n \in \mathbb{N}$, a strictly increasing sequence $a_n \to \infty$ and a random measure $\mu(\omega)$ nondegenerate with* *positive probability, such that*

$$\mathscr{L}\left(\sum_{j=1}^{n} \hat{X}_{nj}^{\omega}/a_n - c_n(\omega)\right) \to_w \mu(\omega) \quad \textit{in probability},$$

then

$$a_n = n^{1/2}L(n) \quad \textit{where } L(n) \textit{ is slowly varying}$$

and there exists $\sigma \neq 0$ *such that*

$$\mathscr{L}\left(\sum_{i=1}^{n} (X_i - EX)/a_n\right) \to_w N(0, \sigma^2)$$

and

$$\mathscr{L}\left(\sum_{j=1}^{n} \left(\hat{X}_{nj}^{\omega} - \overline{X}_n(\omega)\right)/a_n\right) \to_w N(0, \sigma^2) \quad \textit{in probability}.$$

PROOF. From the above observation and the proof of Theorem 1 we obtain that for every subsequence n' there is a subsequence n'' such that

$$\sum_{i=1}^{n''} |X_i|^p I(|X_i| > \lambda a_{n''}) = 0 \quad \text{eventually a.s. (for all } \lambda > 0,\ p \in \mathbb{R}),$$

$$\lim_{n'' \to \infty} \left[\sum_{i=1}^{n''} X_i^2/a_{n''}^2 - \left(\sum_{i=1}^{n''} X_i/a_{n''}n''^{1/2}\right)^2\right] = \sigma^2 \quad \text{a.s.},$$

and

$$\mathscr{L}\left(\sum_{j=1}^{n''} \left(\hat{X}_{n''j}^{\omega} - \overline{X}_{n''}(\omega)\right)/a_{n''}\right) \to_w N(0, \sigma^2) \quad \text{a.s.},$$

for some $\sigma \neq 0$, independent of the sequences $\{n''\}$. In particular, $\mu(\omega) = N(0, \sigma^2)$ ω-a.s. Hence,

$$\sum_{i=1}^{n} |X_i|^p I(|X_i| > \lambda a_n) \to 0 \quad \text{in pr, } \lambda > 0,\ p \in \mathbb{R}, \tag{15}$$

$$\lim_{n \to \infty} \left[\sum_{i=1}^{n} X_i^2/a_n^2 - \left(\sum_{i=1}^{n} X_i/a_n n^{1/2}\right)\right] = \sigma^2 \quad \text{in pr}, \tag{16}$$

and

$$\mathscr{L}\left(\sum_{j=1}^{n} \left(\hat{X}_{nj}^{\omega} - \overline{X}_n(\omega)\right)/a_n\right) \to_w N(0, \sigma^2) \quad \text{in pr}. \tag{17}$$

If $EX^2 < \infty$, then (16) and the law of large numbers give

$$na_n^{-2} \to \sigma^2/\operatorname{Var} X,$$

that is, $a_n \simeq n^{1/2}$. If $EX^2 = \infty$, then Lemma 2 and (16) yield

$$\sum_{i=1}^{n} X_i^2/a_n^2 \to \sigma^2 \quad \text{in pr}. \tag{18}$$

But this implies by classical results that

(19) a_n is regularly varying with exponent $1/2$.

Finally we show that

$$\mathscr{L}\left(\sum_{i=1}^{n}(X_i - EX)/a_n\right) \to N(0, \sigma^2). \tag{20}$$

If $EX^2 < \infty$, there is nothing to prove. Assume $EX^2 = \infty$. Then, since as in (13) in the proof of Theorem 1, $na_n^{-2}EX^2I(|X| \le a_n) \to \sigma^2$, we have $na_n^{-2} \to 0$ and, in particular [note that $E|X| < \infty$ by (19) and (15) with $p = 0$; see (22)],

$$na_n^{-2}\operatorname{Var}(XI(|X| \le a_n)) \to \sigma^2. \tag{21}$$

Taking expectations in (15) with $p = 0$ (note that the summands are bounded) we obtain that for $\lambda > 0$,

$$nP\{|X| > \lambda a_n\} \to 0. \tag{22}$$

It follows easily by regular variation that $na_n^{-1}E|X|I(|X| > a_n) \to 0$, and this, (21) and (22) imply (20) by the classical clt [e.g., Araujo and Giné (1980), page 63]. □

REMARK 4. Suppose that X is in the domain of attraction of a p-stable law, $p \in (0, 2)$, and let $a_n(\omega) = a_n(X_1(\omega), \dots, X_n(\omega)) = \max_{i \le n}|X_i(\omega)|$. It is easy to verify that if

$$\mathscr{L}\left(\sum_{i=1}^{n}\hat{X}_{ni}^{\omega}/a_n(\omega) - c_n(w)\right) \to_w \mu(\omega) \quad \text{in probability,}$$

then, as in Theorems 1 and 3, the integrals $\int f\pi_\delta(dx, \omega)$ defined in the proof of Theorem 1 are still tail random variables and therefore, the Lévy measure $\pi(\omega)$ of $\mu(\omega)$ is not random (on a ω-set of probability 1). Since the limiting measure $\mu(\omega)$ in Athreya's (1987) result has a random Lévy measure, this shows that his result cannot be improved to convergence in probability.

Acknowledgment. We thank K. Athreya for suggesting that a result of the type of Lemma 2 would simplify our first proof of $a_n \simeq n^{1/2}$ in Theorem 1.

REFERENCES

ARAUJO, A. and GINÉ, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York.

ATHREYA, K. B. (1986). Bootstrap of the mean in the infinite variance case, II. Technical Report 86-21, Dept. Statistics, Iowa State Univ.

ATHREYA, K. B. (1987). Bootstrap of the mean in the infinite variance case. *Ann. Statist.* **15** 724–731.

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

CSÖRGŐ, S. and MASON, D. M. (1988). Bootstrapping empirical functions. *Ann. Statist.* To appear.

DE ACOSTA, A. and GINÉ, E. (1979). Convergence of moments and related functionals in the general central limit theorem in Banach spaces. *Z. Wahrsch. verw. Gebiete* **48** 213–231.

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

GINÉ, E. and ZINN, J. (1988). Bootstrapping general empirical measures. Preprint.

HARTIGAN, J. A. (1986). Comment on "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," by B. Efron and R. Tibshirani. *Statist. Sci.* **1** 75–77.

HOFFMANN-JØRGENSEN, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52** 159–186.

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

DEPARTMENT OF MATHEMATICS
CUNY, COLLEGE OF STATEN ISLAND
ST. GEORGE CAMPUS
130 STUYVESANT PLACE
STATEN ISLAND, NEW YORK 10301

DEPARTMENT OF MATHEMATICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843