The Annals of Probability
1993, Vol. 21, No. 2, 917–925

ALMOST SURE BOOTSTRAP OF THE MEAN UNDER RANDOM NORMALIZATION[1]

BY STEVEN J. SEPANSKI

Texas A & M University

We consider the problem of when the bootstrap sample mean, appropriately normalized and centered, converges in distribution along almost every sample path. We allow the normalizing sequence to be an arbitrary sequence of positive random variables. It is proved that the only possible normalizing sequence is essentially $(\sum_{i=1}^n X_i^2)^{1/2}$. Furthermore, if the bootstrap sample mean converges along almost every sample path, then either the variance is finite or else the distribution of X is extremely heavy tailed. In the latter case, the distribution of the bootstrap sample mean is completely determined by how many times the maximum order statistic from the original random sample is repeated in the bootstrap sample. The necessary condition on how heavy the tails must be is $(\sum_{i=1}^n |X_i|^p)^{1/p}/(\sum_{i=1}^n X_i^2)^{1/2} \to 1$ almost surely for all $p \in (0, \infty]$. Furthermore, we show that in this case the limit of the bootstrap sample mean normalized by $(\sum_{i=1}^n X_i^2)^{1/2}$ is Poisson with mean 1.

1. Introduction. Let X, X_i, $i \in N$, be independent identically distributed (iid) real valued random variables defined on some complete probability space $(\Omega, \mathscr{F}, P)$. For $\omega \in \Omega$ and $n \in N$, let $P_n(\omega) = n^{-1}\sum_{i=1}^n \delta_{X_i(\omega)}$ denote the empirical measure and let $\{\hat{X}_{nj}^\omega\}_{j=1}^n$ be iid random variables with laws $P_n(\omega)$. Giné and Zinn (1989) have shown that $EX^2 < \infty$ is necessary for there to exist positive scalars $a_n \uparrow \infty$, centerings $c_n(\omega)$, and a random limiting measure $\mu(\omega)$ [assumed to be nondegenerate with positive probability (henceforth abbreviated w.p.p.)] such that $\mathscr{L}(a_n^{-1}\sum_{j=1}^n \hat{X}_{nj}^\omega - c_n(\omega)) \Rightarrow \mu(\omega)$ a.s. (Here $\Rightarrow$ denotes weak convergence.) Furthermore, it is then necessary that $a_n \sim \sqrt{n}$. We consider a related question, allowing normalization by an arbitrary positive random variable, which we denote by A_n. We show (Theorem 1) that if

$$\mathscr{L}\left(A_n^{-1}(\omega)\sum_{j=1}^n \hat{X}_{nj}^\omega - c_n(\omega)\right) \Rightarrow \mu(\omega) \quad \text{a.s.,} \tag{1.0}$$

then $(\sum_{i=1}^n X_i^2)^{1/2}/A_n \to \xi$ a.s. for some nonnegative random variable ξ supported on the set where μ is nondegenerate.

This is not surprising in the finite variance case. However, in the infinite variance case, we show that if (1.0) holds, then for some random variable ξ as above, $A_n^{-1}(\sum_{i=1}^n |X_i|^p)^{1/p} \to \xi$ a.s. for all $p \in (0, \infty]$. With $p = \infty$ in the expression $(\sum_{i=1}^n |X_i|^p)^{1/p}$, we mean $M_n =: \max_{i \le n}|X_i|$.

Received March 1991; revised November 1991.
[1]Research partially supported by NSF Grant DMS-86-01250.
AMS 1991 *subject classifications*. Primary 62E20, 62F12; secondary 60F05.
Key words and phrases. Bootstrap, central limit theorem, slowly varying tails, maximum order statistic.

Therefore, in the case where the variance is infinite and the bootstrap sum is known to converge for some centering and scaling, one may as well normalize by M_n. The question then arises as to what the limit is. We show that it is Poisson with mean 1 and that the necessary condition $\Sigma_{i=1}^n |X_i|/M_n \to 1$ a.s. is also essentially sufficient (see Theorem 2). We thus have a link between bootstrapping and the work of Maller and Resnick (1984) and of Pruitt (1987) on the contribution to the sum of the summand of maximum modulus.

Before proceeding, let us introduce some notation to be used in addition to that already spelled out above. We will write $\hat{P}$ and $\hat{E}$ for probabilities and expectations, respectively, with respect to the variables $\{\hat{X}_{nj}^{\omega}\}_{j,n}$ with ω fixed. $M_{n,r}$ will denote the rth largest among $|X_1|, \dots, |X_n|$. For simplicity we will write M_n instead of $M_{n,1}$. $X_n^{(r)}$ is the value among $\{X_i\colon i \le n\}$ which yields $M_{n,r}$. That is, $X_n^{(r)} = X_i$ if and only if $M_{n,r} = |X_i|$. For a probability measure μ on R, we define $\bar{\mu}(A) =: \mu(-A)$, where for a set A, we define $-A =: \{-x\colon x \in A\}$.

2. Results.

THEOREM 1. *Let $A_n\colon \Omega \to (0,\infty)$, $n \in N$, be a sequence of ($\mathscr{F}$-measurable) random variables. If there exist random centerings $c_n(\omega)$, $n \in N$, and a random probability measure $\mu(\omega)$, nondegenerate with positive probability, such that* (1.0) *holds, then, with $\Omega_\mu = [\omega \in \Omega\colon \mu(\omega)$ is nondegenerate$]$, there is a nonnegative random variable ξ with $\xi(\omega) > 0$ for almost every $\omega \in \Omega_\mu$ such that either*: (i) $0 < EX^2 < \infty$ *and* $(\Sigma_{i=1}^n X_i^2)^{1/2}/A_n \to \xi I_{\Omega_\mu}$ *a.s.*; *or* (ii) $EX^2 = \infty$ *and* $\forall\, p \in (0,\infty]$, $(\Sigma_{i=1}^n |X_i|^p)^{1/p}/A_n \to \xi I_{\Omega_\mu}$ *a.s.*

Note that in the case of (ii) in the theorem we can conclude from the Hewitt–Savage Zero–One law that $(\Sigma_{i=1}^n |X_i|^p)^{1/p}/M_n \to 1$ a.s. $\forall\, p \in (0,\infty]$. Moreover, it is known [see Maller and Resnick (1984), Theorem 2.3] that this is sufficient to conclude that the tail of the distribution is slowly varying at infinity. In particular, all moments are infinite.

The theorem indicates a dichotomy. (1.0) holds in only one of two cases. Either the tails are light enough for the existence of a second moment or so heavy that all moments are infinite. The finite variance situation has been studied extensively. We are interested primarily in the infinite variance case.

One consequence of Theorem 1 is that, in the case of infinite variance, if the bootstrap of the mean converges weakly almost surely to some random measure $\mu(\omega)$ with some random normalizers A_n, then $A_n \sim M_n$ almost surely on the set Ω_μ. Restricting attention to Ω_μ, the bootstrap of the mean converges weakly almost surely when normalized by the random normalizers M_n. Since $P(\Omega_\mu) > 0$, this is enough to guarantee that the bootstrap of the mean converges weakly almost surely on the whole space to a nondegenerate limit when normalized by M_n. The following theorem demonstrates this and shows that the limiting distribution is Poisson.

THEOREM 2. *Let ρ denote the Poisson law with mean* 1. *The following are equivalent*:

(i) $EX^2 = \infty$ *and there exist centerings* $c_n(\omega)$, *normalizers* $A_n(\omega)$ *and a random measure* $\mu(\omega)$, *nondegenerate with positive probability such that* (1.0) *holds*.

(ii) $\sum_{i=1}^n |X_i|/M_n \to 1$ *a.s. and either* $X_n^{(1)}/M_n \to 1$ *a.s. or* $X_n^{(1)}/M_n \to -1$ *a.s.*

(iii) *Either* $\mathscr{L}(\sum_{j=1}^n \hat{X}_{nj}^{\omega}/M_n) \Rightarrow \rho$ *a.s. or* $\mathscr{L}(\sum_{j=1}^n \hat{X}_{nj}^{\omega}/M_n) \to \bar{\rho}$ *a.s.*

Pruitt (1987) gives several conditions which are equivalent to the first condition in (ii) of Theorem 2. Also, more in the spirit of Pruitt's work, if one is willing to replace M_n by the possibly negative normalizers $X_n^{(1)}$, then one can do away with the second condition in (ii), and in (iii) the limit will always be Poisson with mean 1. Thus, almost sure convergence in law to Poisson with mean one of $\sum_{j=1}^n \hat{X}_{nj}^{\omega}/X_n^{(1)}$ can be added as another equivalent condition in Theorem 1 of Pruitt (1987).

Theorem 2 is the almost sure analogue of Theorem 2.1 of Hall (1990), which deals with convergence in probability. We might point out that in his paper, Hall discusses the case of almost sure convergence briefly but only in the case where the limit is assumed to be a nonrandom normal law (e.g., Proposition 2.1).

3. Proofs. Before proceeding directly to the proof of Theorem 1, we first dispense with some technical questions on the measurability of the random measure μ. We do not assume that this map is measurable. However, the crucial set where μ is nondegenerate is measurable. Indeed, consider the space $\mathscr{P}(R)$ of all probability measures on R with the weak (star) topology and Borel sigma field $\mathscr{B}$. Let $\mu_n(\omega) = \mathscr{L}(A_n^{-1}(\omega)\sum_{j=1}^n \hat{X}_{nj}^{\omega})$. It can be shown that for each $n \in N$, both μ_n and $\mu_n * \bar{\mu}_n$ are measurable maps from $(\Omega, \mathscr{F})$ to $(\mathscr{P}(R), \mathscr{B})$.

Since $\mu * \bar{\mu}$ is the almost sure limit of $\mathscr{F}$-measurable elements, it is $\mathscr{F}$-measurable as well. Because of symmetrization, if $\mu * \bar{\mu}$ is degenerate, it must be δ_0. Hence, $[\mu * \bar{\mu}$ is degenerate$] = (\mu * \bar{\mu})^{-1}(\{\delta_0\}) \in \mathscr{F}$. Also, $\mu * \bar{\mu}$ is nondegenerate if and only if μ is degenerate and so $[\mu$ is nondegenerate$] \in \mathscr{F}$.

If c_n is an $\mathscr{F}$-measurable random variable, then $\mu_n * \delta_{-c_n}$ is also an $\mathscr{F}$-measurable random element of $\mathscr{P}(R)$. Symmetrization allows us to eliminate any measurability hypotheses on c_n.

One of the key steps in the proof of Theorem 1 of Giné and Zinn (1989) was showing that the Lévy measure associated to μ was the zero measure. This is not necessarily the case when we allow random normalization. However, it will be shown that for almost every ω, the function $\lambda \mapsto \pi(\omega)[|x| > \lambda]$ has at most one jump and if there is a jump, then it is one unit. Moreover, for almost all those ω where $\mu(\omega)$ is nondegenerate, the random variable ξ appearing in part (ii) of Theorem 1 is the absolute value of the point where the jump takes place. The next lemma will be useful in proving this.

LEMMA 3. *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables. If there exists $\lambda \in (0,1)$ such that*

$$\lim_{n\to\infty} P\left(\sum_{i=1}^{n} I_{[|X_i| > \lambda M_n]} \geq 2\right) = 1,$$

then X has all moments.

PROOF. Denote by F the distribution function of $|X|$. Without loss of generality we may assume that $F(x) < 1 \ \forall\, x \in R$, since if not, X clearly has all moments. First assume that F is continuous. We will later show how to modify the argument when F is not necessarily continuous.

Let λ be as in the hypothesis. By the continuity of F there exist a_n satisfying $nP(|X| > \lambda a_n) = 1 \ \forall n$. We then have that

$$\begin{aligned} &nP(|X| > a_n)(1 - P(|X| > \lambda a_n))^{n-1} \\ &\quad = nP(F(\lambda|X|) > F(\lambda a_n))F^{n-1}(\lambda a_n) \\ &\quad \leq nEF^{n-1}(\lambda|X|) = P\left(\sum_{i=1}^{n} I_{[|X_i| > \lambda M_n]} = 1\right) = o(1). \end{aligned}$$

The last identity follows from the hypothesis of the lemma. Therefore, $nP(|X| > a_n) \to 0$.

Now since $F(x) < 1 \ \forall\, x$, the sequence a_n is strictly increasing to infinity and therefore for $t > 0$, there exists n such that $a_n < t \leq a_{n+1}$. We then have

$$\lim_{t\to\infty} \frac{P(|X| > t)}{P(|X| > \lambda t)} \leq \lim_{n\to\infty} \frac{P(|X| > a_n)}{P(|X| > \lambda a_{n+1})} = \lim_{n\to\infty} (n+1)P(|X| > a_n) = 0.$$

We now show that $E|X|^p < \infty \ \forall p > 0$. Fix $p > 0$ and let $A > 0$ be such that $t \geq A$ implies $P(|X| > t)/P(|X| > \lambda t) \leq \lambda^p/2$. Next, fix $B > A$; then,

$$\begin{aligned} \int_0^B pt^{p-1}P(|X| > t)\, dt &\leq A^p + \frac{\lambda^p}{2}\int_0^B pt^{p-1}P(|X| > \lambda t)\, dt \\ &\leq A^p + \frac{1}{2}\int_0^B pt^{p-1}P(|X| > t)\, dt. \end{aligned}$$

Letting $B \to \infty$ yields $E|X|^p \leq 2A^p < \infty$.

This proves the desired result in the case when F is continuous. We now handle the case for general F. We use the fact that if Y has a continuous distribution function and X and Y are independent, then $|X| + Y$ also has a continuous distribution function. For the purposes of this proof, we will take $\{Y_i\}_{i=1}^{\infty}$ to be a sequence of i.i.d. random variables which are uniformly distributed over the interval $(\lambda/2, 1)$ and which are independent of $\{X_i\}_{i=1}^{\infty}$. Let $M_n^* =: \max_{i\leq n}(|X_i| + Y_i)$. Now suppose $\sum_{i=1}^{n} I_{[|X_i| > \lambda M_n]} \geq 2$. For some i_1, i_2,

$\min\{|X_{i_1}|, |X_{i_2}|\} > \lambda M_n$ and therefore

$$\frac{\lambda}{2} M_n^* \leq \frac{\lambda}{2}(M_n + 1) \leq \lambda M_n + \frac{\lambda}{2} \leq \lambda M_n + Y_{i_j} < |X_{i_j}| + Y_{i_j}.$$

Hence, $[\Sigma_i^n I_{[|X_i| > \lambda M_n]} \geq 2] \subset [\Sigma_i^n I_{[|X_i| + Y_i > (\lambda/2) M_n^*]} \geq 2]$. By hypothesis, the probability of the former tends to one, hence so does the probability of the latter. So, $\{|X_i| + Y_i\}_{i=1}^\infty$ satisfies the hypotheses of the lemma and we apply the continuous case to conclude that $|X| + Y$ has all moments. It is then immediate that X has all moments as well.

PROOF OF THEOREM 1. The proof will be separated into two cases. First, suppose that $EX^2 < \infty$. In this case it is known [Bickel and Freedman (1981), Theorem 2.1] that $\mathscr{L}(n^{-1/2}\Sigma_{j=1}^n(\hat{X}_{nj}^\omega - \overline{X}_n)) \Rightarrow N(0, \operatorname{Var} X)$ a.s. Moreover, since μ is nondegenerate with positive probability, $\operatorname{Var} X \neq 0$. By the convergence of types theorem we have that for each ω such that $\mu(\omega)$ is nondegenerate, there exists a positive number $\eta(\omega)$ satisfying $n^{-1/2}A_n(\omega) \to \eta(\omega)$. For those ω such that $\mu(\omega)$ is degenerate, necessarily $\sqrt{n}/A_n(\omega) \to 0$. Also, by the strong law of large numbers, $(n^{-1}\Sigma_{i=1}^n X_i^2)^{1/2} \to (EX^2)^{1/2}$ a.s. Then $\xi =: \eta^{-1}(EX^2)^{1/2} I_{\Omega_\mu}$ has the desired properties. This proves (i).

The remainder of the proof is devoted to analyzing what happens when $EX^2 = \infty$. First, a word about the proof. Often certain statements will be made which hold for almost every ω. To avoid technical difficulties we will assume that subsequent almost sure statements satisfy all previous almost sure statements. This does not create a problem since only finitely many such statements will occur.

We first prove that for almost every ω, the array $\{\hat{X}_{nj}^\omega / M_n(\omega)\colon j \leq n\}_{n=1}^\infty$ is infinitesimal. Fix $\varepsilon > 0$ and a positive integer M:

$$\begin{aligned} \max_{j \leq n} \hat{P}\Big(|\hat{X}_{nj}^\omega| / M_n(\omega) > \varepsilon\Big) &= n^{-1} \sum_{i=1}^n I_{[|X_i| > \varepsilon M_n]} \\ &\leq I_{[M_n \leq M]} + n^{-1} \sum_{i=1}^n I_{[|X_i| > \varepsilon M]} \to P(|X| > \varepsilon M) \quad \text{a.s.} \end{aligned}$$

The last line follows from the strong law of large numbers and from the fact that since $EX^2 = \infty$, $M_n \to \infty$ a.s. Letting $M \to \infty$ yields almost sure infinitesimality as needed.

Next, we shall show that a necessary condition for (1.0) is that A_n must converge to infinity sufficiently fast. More precisely,

$$P\big(\limsup(M_n/A_n) = \infty\big) = 0.$$

Suppose, on the contrary, that both $\limsup(M_n/A_n) = \infty$ w.p.p. and (1.0) holds. There then exists an ω and a subsequence n' such that the following three facts hold: The array $\{\hat{X}_{n'j}^\omega / M_{n'}(\omega)\colon j \leq n'\}_{n'}$ is infinitesimal, $\mathscr{L}(A_{n'}^{-1}\Sigma_{j=1}^{n'}\hat{X}_{n'j}^\omega - c_{n'}(\omega)) \Rightarrow \mu(\omega)$ and $A_{n'}(\omega)/M_{n'}(\omega) \to 0$.

The last two of these imply that

$$\frac{\sum_{j=1}^{n'} \hat{X}_{n'j}^{\omega}}{M_{n'}(\omega)} - \frac{c_{n'}(\omega) A_{n'}(\omega)}{M_{n'}(\omega)} \Rightarrow 0.$$

Under the infinitesimality stated in (iii), a necessary condition for this is $\sum_{j=1}^{n'} I_{[|X_i(\omega)| > \varepsilon M_{n'}(\omega)]} \to 0$ [see Araujo and Giné (1980), Corollary 4.8]. This is impossible, at least for $\varepsilon \in (0, 1)$. Hence, we conclude that $\limsup(M_n/A_n) < \infty$ a.s.

Since $M_n \to \infty$ a.s., $A_n \to \infty$ a.s. and we can prove in the same way as we did earlier for M_n that for almost every ω, the array $\{\hat{X}_{nj}^{\omega}/A_n(\omega)\colon j \le n\}_{n=1}^{\infty}$ is infinitesimal. Hence, $\mu(\omega)$ is almost surely an infinitely divisible measure.

Now, let $\pi(\omega)$ denote the Lévy measure of $\mu(\omega)$. By the converse C.L.T. [see, e.g., Araujo and Giné (1980), Theorem 4.7], we have that for almost every ω,

$$\sum_{j=1}^{n} \mathscr{L}\left(\frac{\hat{X}_{nj}^{\omega}}{A_n}\right)\Bigg|_{[|x| > \lambda]} \Rightarrow \pi(\omega)|_{[|x| > \lambda]} \tag{3.1}$$

for all λ which satisfy

$$\pi(\omega)\{-\lambda, \lambda\} = 0. \tag{3.2}$$

In particular,

$$\sum_{i=1}^{n} I_{[|X_i| > \lambda A_n]} \to \pi(\omega)[|x| > \lambda]. \tag{3.3}$$

Note that for almost every fixed ω all but countably many λ satisfy (3.2) and, in particular, Lebesgue almost every λ. However, the set of Lebesgue measure zero may depend on ω.

A simple measurability argument shows that the function $(\omega, \lambda) \mapsto \pi(\omega)[|x| > \lambda]$ is jointly measurable.

The joint measurability allows us to legitimately apply Fubini's theorem to conclude that there is a countable dense set $D \subset (0, \infty)$ such that $\forall\, \lambda \in D$, (3.3) holds almost surely. For each fixed $\lambda \in D$, we write $\xi_\lambda(\omega) = \pi(\omega)[|x| > \lambda]$. ξ_λ is necessarily almost surely a nonnegative integer. We use Lemma 3 to show that ξ_λ cannot be two or greater.

Suppose there exists $\lambda_0 \in D$ such that $\xi_{\lambda_0} \ge 2$ w.p.p. For α large enough, $P([\sup_n(M_n/A_n)\alpha] \cap [\sum_{i=1}^{n} I_{[|X_i| > \lambda_0 A_n]} \to \xi_{\lambda_0}] \cap [\xi_{\lambda_0} \ge 2]) > 0$. For ω in this set we see that eventually

$$2 \le \xi_{\lambda_0} = \sum_{i=1}^{n} I_{[|X_i| > \lambda_0 A_n]} \le \sum_{i=1}^{n} I_{[|X_i| > \lambda_0 M_n/\alpha]} \le \sum_{i=1}^{n} I_{[|X_i| > \lambda M_n]}$$

for some $\lambda \in D \cap (0, \min\{\lambda_0/\alpha, 1\})$. For this λ, $P(\sum_{i=1}^{n} I_{[|X_i| > \lambda M_n]} \ge 2$ eventually$) > 0$. The Hewitt–Savage zero–one law guarantees that the probability is 1. Applying Lemma 3 we conclude that $EX^2 < \infty$, a contradiction.

We conclude that $\xi_\lambda \in \{0,1\}$ a.s. $\forall\ \lambda \in D$. There are two possibilities. (i) $\xi_\lambda = 0\ \forall\ \lambda \in D$ or (ii) there is a $\lambda \in D$ such that $\xi_\lambda = 1$.

Here the proof splits into cases according to whether ω satisfies (i) or (ii) and whether $\mu(\omega)$ is degenerate or not.

For those ω satisfying (i), (3.3) under the hypothesis that (i) holds is equivalent to $\Sigma_{j=1}^n \hat{P}(|\hat{X}_{nj}^\omega| > \lambda A_n) \to 0$ a.s. $\forall\ \lambda > 0$. In this case the limiting distribution must be normal [see Araujo and Giné (1980), Corollary 4.8]. Thus, $\mu(\omega) = N(a(\omega), \sigma^2(\omega))$ for some random parameters $a(\omega)$ and $\sigma^2(\omega)$. By the converse C.L.T. (suppressing ω from the right side below),

$$(3.4)\qquad \sigma^2(\omega) = \lim_{n\to\infty}\left(\frac{\Sigma_{i=1}^n X_i^2}{A_n^2} - \frac{n^{-1}(\Sigma_{i=1}^n X_i)^2}{A_n^2}\right) = \lim_{n\to\infty}\frac{\Sigma_{i=1}^n X_i^2}{A_n^2}.$$

The last equality is obtained from Giné and Zinn [(1989), Lemma 2].

To essentially eliminate the case of (i) and μ nondegenerate holding simultaneously, we now show that $P([\xi_\lambda = 0\ \forall\ \lambda > 0] \cap \Omega_\mu) = 0$. Assume not. On the set in question $\sigma(\omega) > 0$ and therefore

$$\mathscr{L}\left(\frac{\Sigma_{j=1}^n \hat{X}_{nj}^\omega}{\left(\Sigma_{i=1}^n X_i^2\right)^{1/2}} - \sigma^{-1}(\omega)c_n\right) \Rightarrow N(a(\omega)/\sigma(\omega), 1) \quad \text{w.p.p.}$$

That this implies $EX^2 < \infty$ was essentially shown by Csörgő and Mason (1989). They proved it in the case where the limit is a nonrandom normal distribution, the normalizing sequence is the sample standard deviation and the convergence is almost sure instead of with positive probability, but their argument goes through in our case, too. We give their argument briefly with the appropriate changes. By the converse C.L.T., $P(\Sigma_{i=1}^n I[|X_i| > \lambda(\Sigma_{i=1}^n X_i^2)^{1/2}]) \to 0\ \forall\ \lambda > 0$. This implies $M_n/(\Sigma_{i=1}^n X_i^2)^{1/2} \to 0$ w.p.p. The zero–one law allows us to obtain almost sure convergence. Applying Lemma 4.1 of Maller and Resnick (1984) to the variables $\{X_i^2\}_{i=1}^\infty$ yields $EX^2 < \infty$, giving a contradiction as desired.

Of course, for ω satisfying (i) and $\mu(\omega)$ degenerate, recalling (3.4),

$$(3.5)\qquad \frac{\left(\Sigma_{i=1}^n X_i^2\right)^{1/2}}{A_n} \to 0 \quad \text{for almost every } \omega \in [\mu \text{ is degenerate}].$$

We now consider ω satisfying (ii). We note that (ii) implies $\mu(\omega)$ is nondegenerate and therefore this is the only case left to consider. The set of such ω has positive probability because if not, μ is almost surely degenerate.

Now, let $\lambda^*(\omega) = \sup\{\lambda \in D\colon \xi_\lambda(\omega) = 1\}$. Note that $\lambda^* < \infty$ a.s. since the set $\{\lambda \in D\colon \xi_\lambda(\omega) = 1\}$ is bounded above by $\limsup(M_n/A_n)$. Of course, $\lambda^* > 0$. By monotonicity of $\lambda \mapsto \pi(\omega)[|x| > \lambda]$ and the fact that D is dense, $\Sigma_{i=1}^n I_{[|X_i| > \lambda A_n]} \to 1, \forall\ \lambda < \lambda^*$ and $\Sigma_{i=1}^n I_{[|X_i| > \lambda A_n]} \to 0,\ \forall\ \lambda \geq \lambda^*$. Therefore,

$$\lambda^* \leq \liminf(M_n/A_n) \leq \limsup(M_n/A_n) \leq \lambda^*.$$

We therefore have that for almost every ω satisfying (ii),

$$\frac{M_n(\omega)}{A_n(\omega)} \to \lambda^*(\omega). \tag{3.6}$$

We also have that if ω satisfies (ii), then $M_{n,2}(\omega)/A_n(\omega) \to 0$. When we combine this with (3.6) we may conclude $(M_{n,2}/M_n) \to 0$ w.p.p. As usual, the zero–one law yields almost sure convergence. We then apply Theorem 1 of Pruitt (1987) to the variables $\{|X_i|^p\}_{i=1}^{\infty}$ to obtain

$$\frac{\left(\sum_{i=1}^n |X_i|^p\right)^{1/p}}{M_n} \to 1 \quad \text{a.s.} \ \forall\, p \in (0,\infty]. \tag{3.7}$$

Finally, by (3.5), (3.6) and (3.7), $\xi = \lambda^* I[\xi_\lambda = 1$ for some $\lambda]$ has the desired properties.

Proof of Theorem 2. We will prove that (i) implies (ii) first. Without loss of generality we may assume that $P(\Omega_\mu) = 1$. Indeed, if we can prove that (i) implies (ii) in this case, then arguing relative to Ω_μ gives that the limits in (ii) hold with positive probability and the zero–one law gives almost sure convergence.

The first half of (ii) of Theorem 2 follows directly from (ii) of Theorem 1. Assuming $P(\Omega_\mu) = 1$, we have by Theorem 1 that there exists a random variable $\xi > 0$ a.s. such that $M_n/A_n \to \xi$ a.s. Therefore, $\mathscr{L}(M_n^{-1}\sum_{j=1}^n \hat{X}_{nj}^{\omega} - d_n(\omega)) \Rightarrow \nu(\omega)$ a.s. for some centerings $d_n(\omega)$, and an almost surely infinitely divisible and nondegenerate random measure $\nu(\omega)$.

We again denote the corresponding Lévy measure by $\pi(\omega)$. An application of Lemma 3 shows

$$\sum_{i=1}^n I_{[|X_i| > \lambda M_n]} \to 0 \quad \text{a.s.} \ \forall\, \lambda \geq 1, \tag{3.8}$$

$$\sum_{i=1}^n I_{[|X_i| > \lambda M_n]} \to 1 \quad \text{a.s.} \ \forall\, \lambda \in (0,1). \tag{3.9}$$

We may then conclude that

$$\pi(\omega) = a(\omega)\delta_1 + (1 - a(\omega))\delta_{-1} \quad \text{a.s.}, \tag{3.10}$$

for some random variable a taking values in $[0,1]$.

By (3.1) with M_n in place of A_n, and by Fubini's theorem, there is a countable dense set $D \subset (0,\infty)$ such that $\forall\ \lambda \in D$, $\sum_{i=1}^n I_{[X_i < -\lambda M_n]} \to \pi(\omega)(-\infty, -\lambda)$ a.s. In particular, there exists $\lambda \in (0,1) \cap D$ for which this holds. However, for $\lambda \in (0,1)$, the limit is equal to $1 - a(\omega)$ almost surely by (3.10). By the Hewitt–Savage zero–one law, we then conclude that a and π are constant almost surely. Since the sequence is integer valued, so is the limit. Thus, in fact, $a \in \{0,1\}$. By (3.10), either $\pi = \delta_1$ a.s. or $\pi = \delta_{-1}$ a.s. In the former case, $X_n^{(1)}/M_n \to 1$ a.s. In the latter case, $X_n^{(1)}/M_n \to -1$ a.s.

Next, we prove (ii) implies (iii). Suppose, to be specific in (ii), that we have $X_n^{(1)}/M_n \to 1$ a.s.

The first condition in (ii) implies $M_{n,2}/M_n \to 0$ a.s., which, in turn, implies (3.8). Also, (3.9) clearly holds. These two conditions plus the condition $X_n^{(1)}/M_n \to 1$ a.s. yield

$$(3.11) \qquad \sum_{j=1}^{n} \mathscr{L}\left(\frac{\hat{X}_{nj}^{\omega}}{M_n}\right)\Big|_{[|x|>\lambda]} \Rightarrow \delta_1\Big|_{[|x|>\lambda]} \quad \text{a.s.}\ \forall\, \lambda \neq 1.$$

$M_n \uparrow \infty$ a.s. is necessary for the first condition in (ii) and therefore the array under investigation is infinitesimal a.s.

Next, observe that the truncated (at $\delta < 1$) means and variances converge to zero almost surely. Indeed, directly from the first hypothesis of (ii),

$$\sum_{j=1}^{n} \widehat{\mathrm{Var}}\left(\frac{\hat{X}_{nj}^{\omega}}{M_n} I_{[|\hat{X}_{nj}^{\omega}| \le \delta M_n]}\right)$$
$$= \frac{\sum_{i=1}^{n} X_i^2 I_{[|X_i| < \delta M_n]}}{M_n^2} - \left(\frac{\sum_{i=1}^{n} X_i I_{[|X_i| < \delta M_n]}}{M_n \sqrt{n}}\right)^2 \to 0 \quad \text{a.s.},$$
$$\sum_{j=1}^{n} \hat{E}\left(\frac{\hat{X}_{nj}^{\omega}}{M_n} I_{[|\hat{X}_{nj}^{\omega}| \le \delta M_n]}\right) = \frac{\sum_{i=1}^{n} X_i I_{[|X_i| < \delta M_n]}}{M_n} \to 0 \quad \text{a.s.}$$

Under infinitesimality these and (3.11) are sufficient for $\mathscr{L}(\sum_{j=1}^{n} \hat{X}_{nj}^{\omega}/M_n) \Rightarrow \rho$ a.s. [Araujo and Giné (1980), Corollary 4.8].

Finally, (iii) implies (i) is trivial.

Acknowledgments. The results in this paper are contained in the author's Ph.D. dissertation which was completed at Texas A & M University under the supervision of Professor Joel Zinn. His guidance and encouragement were greatly appreciated. The comments of the referee and an Associate Editor were quite helpful in making the presentation more readable.

REFERENCES

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

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. (1989). Bootstrapping empirical functions. *Ann. Statist.* **17** 1447–1471.

GINÉ, E. and ZINN, J. (1989). Necessary conditions for the bootstrap of the mean. *Ann. Statist.* **17** 684–691.

HALL, P. (1990). Asymptotic properties of the bootstrap for heavy-tailed distributions. *Ann. Probab.* **18** 1342–1360.

MALLER, R. A. and RESNICK, S. I. (1984). Limiting behaviour of sums and the term of maximum modulus. *Proc. London Math. Soc. (3)* **49** 385–422.

PRUITT, W. E. (1987). The contribution to the sum of the summand of maximum modulus. *Ann. Probab.* **15** 885–896.

DEPARTMENT OF MATHEMATICS
ALBION COLLEGE
ALBION, MICHIGAN 49224