The Annals of Probability
1992, Vol. 20, No. 4, 1805–1842

RATIOS OF TRIMMED SUMS AND ORDER STATISTICS

BY HARRY KESTEN AND R. A. MALLER

Cornell University and University of Western Australia

Let X_i be independent and identically distributed random variables with distribution F. Let $M_n^{(n)} \le \cdots \le M_n^{(1)}$ be the sample $X_1, X_2, \dots, X_n$ arranged in increasing order, with a convention for the breaking of ties, and let $X_n^{(n)}, \dots, X_n^{(1)}$ be the sample arranged in increasing order of modulus, again with a convention to break ties. Let $S_n = X_1 + \cdots + X_n$ be the sample sum. We consider sums trimmed by large values,

$$^{(r)}S_n = S_n - M_n^{(1)} - \cdots - M_n^{(r)}, \qquad r = 1, 2, \dots, n, \ ^{(0)}S_n = S_n,$$

and sums trimmed by values large in modulus,

$$^{(r)}\tilde{S}_n = S_n - X_n^{(1)} - \cdots - X_n^{(r)}, \qquad r = 1, 2, \dots, n, \ ^{(0)}\tilde{S}_n = S_n.$$

In this paper we give necessary and sufficient conditions for $^{(r)}\tilde{S}_n / |X_n^{(r)}| \to \infty$ and $^{(r)}S_n / M_n^{(r)} \to \infty$ to hold almost surely or in probability, when $r = 1, 2, \dots$. These express the dominance of the sum over the large values in the sample in various ways and are of interest in relation to the law of large numbers and to central limit behavior. Our conditions are related to the relative stability almost surely or in probability of the trimmed sum and, hence, to analytic conditions on the tail of the distribution of X_i which give relative stability.

1. Introduction. Let X_i, X be independent and identically distributed random variables with distribution F and let $M_n^{(n)} \le \cdots \le M_n^{(1)}$ be the sample $X_1, X_2, \dots, X_n$ arranged in increasing order; more precisely, let $m_n(j)$, $n \ge 1$, $1 \le j \le n$, be the number of X_i satisfying $X_i > X_j$, $1 \le i \le n$, or $X_i = X_j$, $1 \le i \le j$, and let $M_n^{(r)} = X_j$ if $m_n(j) = r$. Thus, ties are broken by priority of index. Let

$$S_n = X_1 + \cdots + X_n, \tag{1.1}$$

and define the trimmed sum

$$^{(r)}S_n = S_n - M_n^{(1)} - \cdots - M_n^{(r)}, \qquad r = 1, 2, \dots, n, \ ^{(0)}S_n = S_n. \tag{1.2}$$

Also let $X_n^{(n)}, \dots, X_n^{(1)}$ denote the sample arranged in increasing order of modulus, with a similar convention as above to specify $X_n^{(r)}$ in the case of ties among $|X_i|$. Define sums trimmed by removing the values of largest modulus:

$$^{(r)}\tilde{S}_n = S_n - X_n^{(1)} - \cdots - X_n^{(r)} \qquad r = 1, 2, \dots, n, \ ^{(0)}\tilde{S}_n = S_n.$$

Consider the behavior of the ratio $S_n / M_n^{(1)}$, when $EX^- < EX^+ < \infty$, where $X^+ = \max(X, 0)$ and $X^- = X^+ - X$. Then by the strong law of large numbers,

Received February 1991; revised December 1991.
AMS 1980 *subject classifications*. Primary 60F15, 60J15; secondary 60F05, 62G30.
Key words and phrases. Trimmed sums, order statistics, relative stability, strong law.

$n^{-1}S_n \to EX > 0$ almost sure (a.s.) and, as is well known, $M_n^{(1)}/n \to 0$ a.s. Also, it is easily shown that $P\{M_n^{(1)} \le 0 \text{ i.o.}\} = 0$, if $F(0) < 1$, which we will assume throughout this paper. Thus $n/M_n^{(1)} \to \infty$ a.s. and it follows that $S_n/M_n^{(1)} \to \infty$ a.s. Equivalently, ${}^{(1)}S_n/M_n^{(1)} \to \infty$ a.s. The condition ${}^{(1)}S_n/M_n^{(1)} \to \infty$ a.s. or in probability, or more generally ${}^{(r)}S_n/M_n^{(r)} \to \infty$ a.s. or in probability, where $r = 1, 2, \dots,$ expresses in one way the dominance of the sum over the large values in the sample. A similar interpretation can be given to ${}^{(r)}\tilde{S}_n/|X_n^{(r)}| \to \infty$. These relations are of interest in connection with the law of large numbers and central limit behavior. Our aim in this paper is to give necessary and sufficient conditions for ${}^{(r)}\tilde{S}_n/|X_n^{(r)}| \to \infty$ and ${}^{(r)}S_n/M_n^{(r)} \to \infty$ a.s. or in probability.

An important role in our calculations is played by the concept of relative stability. S_n is said to be relatively stable (in probability) if $S_n/B_n \to_P +1$ or $S_n/B_n \to_P -1$ (abbreviated to $S_n/B_n \to_P \pm 1$) for a nonstochastic sequence B_n, *which we will always assume is strictly positive and increasing*. When $X_i \ge 0$ a.s. an intimate connection between the condition $S_n/M_n^{(1)} \to_P \infty$ and relative stability was proved by Breiman (1965); he showed in fact that these are equivalent to each other and to an analytic condition due to Feller [(1971), page 236], which we write generally as

$$\frac{|\nu(x)|}{xP(|X| > x)} \to \infty, \qquad x \to \infty, \tag{1.3}$$

or, equivalently,

$$\frac{|A(x)|}{xP(|X| > x)} \to \infty, \qquad x \to \infty, \tag{1.4}$$

where, for $x > 0$,

$$A(x) = \int_0^x [1 - F(y) - F(-y)]\, dy \tag{1.5}$$

and

$$\nu(x) = E[XI(|X| \le x)]. \tag{1.6}$$

We also define

$$A_+(x) = \int_0^x [1 - F(y)]\, dy, \qquad A_-(x) = \int_0^x F(-y)\, dy \tag{1.7}$$

and

$$V(x) = E[X^2 I(|X| \le x)]. \tag{1.8}$$

That conditions (1.3) and (1.4) continue to characterize relative stability of S_n for general X_i was shown by Maller (1978) [see also Rogozin (1976)]. *We will always assume that X_i do not have bounded support*, so $P(|X| > x) > 0$ for all x; thus when (1.4) holds, $|A(x)|$ is strictly positive for large x. Consequently, for sufficiently large x, $A(x)$ and hence $\nu(x)$ are of constant sign; this follows from the continuity of $A(x)$ and then from the fact that $\nu(x) \sim A(x)$ as

$x \to \infty$, which is true under either one of (1.3) and (1.4). Furthermore, Rogozin (1976) shows that $|A(x)|$ and hence $|\nu(x)|$ are slowly varying as $x \to \infty$ [see Bingham, Goldie and Teugels (1987) for definitions of slow and regular variation], and the norming constants B_n for which $S_n/B_n \to_P \pm 1$ satisfy

$$B_n \sim n|\nu(B_n)| \sim n|A(B_n)| \quad \text{as } n \to \infty. \tag{1.9}$$

Since convergence of the type $S_n/B_n \to_P \pm 1$ entails $X_n^{(1)}/B_n \to_P 0$ and $M_n^{(1)}/B_n \to_P 0$ as $n \to \infty$ [e.g., by Gnedenko and Kologorov (1968), page 124], and since $P(M_n^{(j)} \geq 0) \to 1$ as $n \to \infty$, for $j \geq 1$, when $F(0) < 1$, "positive" relative stability of the type $S_n/B_n \to_P +1$ (equivalently, as it turns out, positive relative stability of the trimmed sums ${}^{(r-1)}\tilde{S}_n$ or ${}^{(r-1)}S_n$) implies ${}^{(r)}\tilde{S}_N/|X_n^{(r)}| \to_P \infty$ and ${}^{(r)}S_n/M_n^{(r)} \to_P \infty$, $r = 1, 2, \dots$. It is plausible to conjecture that the converses of these are true, and we prove in Theorem 2.1 that this is indeed the case. Thus this type of behavior does not depend on the order of trimming, that is, on the value of r.

For a.s. convergence, Mori (1976) showed that trimming S_n to ${}^{(r-1)}\tilde{S}_n$ "improves" almost sure behavior in cases where $E|X|$ may be infinite, in the sense that $({}^{(r-1)}\tilde{S}_n - C_n)/n$ may still converge to 0 a.s. for a nonstochastic sequence C_n; in fact this occurs if and only if the integral

$$\int_0^\infty \{xP(|X| > x)\}^r \frac{dx}{x} \tag{1.10}$$

converges. Maller (1984) extended this to give conditions for which ${}^{(r-1)}\tilde{S}_n/B_n \to \pm 1$ a.s. for some B_n and showed that this entails $X_n^{(r)}/B_n \to 0$ a.s., so that $X_n^{(r)}/{}^{(r)}\tilde{S}_n \to 0$ in this case. [Versions of Mori's result for ${}^{(r-1)}S_n$ are also mentioned in Maller (1984).] So once again, relative stability (a.s.) of the trimmed sum appears closely related to the divergence of ${}^{(r)}\tilde{S}_n/|X_n^{(r)}|$. A similar relation will be seen to hold between ${}^{(r-1)}S_n/B_n \to 1$ a.s., for some B_n, and ${}^{(r)}S_n/M_n^{(r)} \to \infty$. For results related to the case $r = 1$ of these relations, see Chow and Robbins (1961), Maller (1978), Kesten (1971), O'Brien (1980) and Maller and Resnick (1984).

The kind of trimming is immaterial when the summed random variables are nonnegative almost surely, a case which is of interest in itself. In fact, let

$${}^{(r)}T_n = X_1^2 + X_2^2 + \cdots + X_n^2 - \left(X_n^{(1)}\right)^2 - \cdots - \left(X_n^{(r)}\right)^2, \qquad r = 1, 2, \dots, n,$$

with ${}^{(0)}T_n$ as the sample sum of squares. Then Theorem 2.2, applied to X_i^2, tells us, for example, that for $r = 1, 2, \dots$ the following are equivalent:

$$\frac{{}^{(r)}T_n}{\left(X_n^{(r)}\right)^2} \to \infty \quad \text{a.s.};$$

$$\frac{{}^{(r-1)}T_n}{C_n} \to 1 \quad \text{a.s., for some nonstochastic sequence } C_n;$$

$$\int_1^\infty \left\{ \frac{x^2 P(|X| > x)}{\int_0^x y P(|X| > y)\, dy} \right\}^r \frac{dx}{x} < \infty. \tag{1.11}$$

[For $r = 1$, (1.11) is equivalent to $E(X)^2 < \infty$.] Likewise Theorem 2.1 gives the equivalence of the following:

$$\frac{{}^{(r)}T_n}{\left(X_n^{(r)}\right)^2} \to_P \infty;$$

$$\frac{{}^{(r-1)}T_n}{C_n} \to_P 1, \quad \text{for some nonstochastic sequence } C_n;$$

$$\frac{x^2 P(|X| > x)}{\int_0^x y P(|X| > y)\, dy} \to 0, \quad \text{as } x \to \infty. \tag{1.12}$$

Condition (1.12) is Lévy's condition for F to be in the domain of attraction of the normal distribution, while (1.11) is a condition on the rate of convergence of this quantity to 0. Thus we have a natural link with the central limit behavior of the sum.

2. Results.

First we state the version of the theorem for convergence in probability. We need the following definitions: when $\varepsilon > 0$, and $P(X < 0) > 0$, let $-L_-(\varepsilon)$ and $L_+(\varepsilon)$ be an ε and a $(1-\varepsilon)$ quantile of F, that is,

$$F(-L_-(\varepsilon)-) \le \varepsilon \le F(-L_-(\varepsilon)) \quad \text{and} \quad F(L_+(\varepsilon)-) \le 1 - \varepsilon \le F(L_+(\varepsilon)).$$

When $P(X < 0) = 0$ take $L_-(\varepsilon) = 0$. $L_+(\varepsilon)$ and/or $L_-(\varepsilon)$ may not be uniquely defined by these, but any choice which satisfies the conditions will work. We also define

$$\mu(\varepsilon) = E\left[XI(-L_-(\varepsilon) \le X \le L_+(\varepsilon))\right] = \int_{[-L_-(\varepsilon),\, L_+(\varepsilon)]} x\, dF(x), \tag{2.1}$$

$$\sigma^2(\varepsilon) = E\left[X^2 I(-L_-(\varepsilon) \le X \le L_+(\varepsilon))\right] = \int_{[-L_-(\varepsilon),\, L_+(\varepsilon)]} x^2\, dF(x). \tag{2.2}$$

THEOREM 2.1. *Let $F(0) < 1$. For $r = 1, 2, \ldots$ the following are equivalent*:

$$\frac{{}^{(r)}S_n}{M_n^{(r)}} \to_P \infty, \qquad n \to \infty; \tag{2.3}$$

$$\frac{\mu(\varepsilon)}{\varepsilon^{1/2}\sigma(\varepsilon)} \to \infty, \qquad \varepsilon \to 0+; \tag{2.4}$$

there is a nonstochastic sequence $B_n > 0$, $B_n \uparrow \infty$, *for which*

$$\frac{{}^{(r-1)}S_n}{B_n} \to_P 1, \qquad n \to \infty, \tag{2.5}$$

$$\frac{\nu(x)}{xP(|X| > x)} \to \infty, \qquad x \to \infty; \tag{2.6}$$

$$\frac{x\nu(x)}{V(x)} \to \infty, \qquad x \to \infty, \text{ [see (1.8) for } V(x)\text{]}; \tag{2.7}$$

$$\frac{\nu(x)}{\{P(|X| \ge x)V(x)\}^{1/2}} \to \infty, \qquad x \to \infty. \tag{2.8}$$

Furthermore, (2.3)–(2.8) *are equivalent to each of the following*:

$$\frac{{}^{(r)}\tilde{S}_n}{|X_n^{(r)}|} \to_P \infty, \qquad n \to \infty; \tag{2.9}$$

$$\frac{{}^{(r-1)}\tilde{S}_n}{B_n} \to_P 1, \qquad n \to \infty; \tag{2.10}$$

where B_n *is the same sequence as in* (2.5). *The same sequence* B_n *can be used in* (2.5) *or* (2.10) *for all* $r \ge 1$.

THEOREM 2.2. *Let* $F(0) < 1$. *The following are equivalent for* $r = 1, 2, 3, \dots$ *and* $n \to \infty$:

$$\frac{{}^{(r)}\tilde{S}_n}{|X_n^{(r)}|} \to \infty \quad a.s.; \tag{2.11}$$

$$\liminf_{n\to\infty} \frac{{}^{(r)}\tilde{S}_n}{|X_n^{(r)}|} > 0 \quad a.s.; \tag{2.12}$$

there is an $x_0 > 0$ *such that* $A(x) > 0$ *for* $x \ge x_0$ *and*

$$\int_{x_0}^{\infty} \left\{ \frac{xP\{|X| > x\}}{A(x)} \right\}^r \frac{dx}{x} < \infty; \tag{2.13}$$

there is a nonstochastic sequence $B_n > 0$, $B_n \uparrow \infty$, *for which*

$$\frac{{}^{(r-1)}\tilde{S}_n}{B_n} \to 1 \quad a.s.; \tag{2.14}$$

there is a nonstochastic sequence $B_n > 0$, $B_n \uparrow \infty$, *for which*

$$(2.15)\qquad 0 < \liminf_{n\to\infty} \frac{{}^{(r-1)}\tilde{S}_n}{B_n} \le \limsup_{n\to\infty} \frac{{}^{(r-1)}\tilde{S}_n}{B_n} < \infty \quad a.s.$$

Finally, (2.11)–(2.15) *hold for* $r = 1$ *if and only if*

$$0 < EX \le E|X| < \infty.$$

THEOREM 2.3. *Let* $F(0) < 1$ *and assume* $E|X| = \infty$. *Then the following are equivalent for* $r = 2, 3, \dots$ *and* $n \to \infty$:

$$(2.16)\qquad \frac{{}^{(r)}S_n}{M_n^{(r)}} \to \infty \quad a.s;$$

$$(2.17)\qquad \liminf_{n\to\infty} \frac{{}^{(r)}S_n}{M_n^{(r)}} > 0 \quad a.s;$$

there is an $x_0 > 0$ *such that*

$$(2.18a)\qquad \int_{x_0}^{\infty} \left\{ \frac{x[1 - F(x)]}{A_+(x)} \right\}^r \frac{dx}{x} < \infty$$

and

$$(2.18b)\qquad \int_0^{\infty} \left\{ \frac{x}{A_+(x)} \right\} |dF(-x)| < \infty;$$

there is a nonstochastic sequence $B_n > 0$, $B_n \uparrow \infty$, *for which*

$$(2.19)\qquad \frac{{}^{(r-1)}S_n}{B_n} \to 1 \quad a.s.;$$

there is a nonstochastic sequence $B_n > 0$, $B_n \uparrow \infty$, *for which*

$$(2.20)\qquad 0 < \liminf_{n\to\infty} \frac{{}^{(r-1)}S_n}{B_n} \le \limsup_{n\to\infty} \frac{{}^{(r-1)}S_n}{B_n} < \infty \quad a.s.$$

When $E|X| < +\infty$, *conditions* (2.16)–(2.20) *are still equivalent, even for* $r = 1$, *if* (2.18) *is replaced by*

$$E(X^-) < E(X^+) < \infty \quad \text{or, equivalently,} \quad E(X) > 0.$$

Finally, (2.16)–(2.20) *can hold for* $r = 1$ *only if* $E|X| < +\infty$.

REMARK 1. For convergence in probability, since (2.4) and (2.6)–(2.8) do not depend on r, we see that (2.3) and (2.5), and (2.9) and (2.10), hold for all r if they hold for $r = 1$ [even with the same sequence B_n for all r in (2.5) and (2.10)]; thus all conditions are equivalent to (positive) relative stability (in probability) ($S_n/B_n \to_P +1$) or to $S_n/M_n^{(1)} \to_P \infty$. This shows that trimming a fixed number of extremes does not affect relative stability in this sense. However, the example in Remark 6 shows that this is not so for a.s. convergence. We note that the sequences B_n in (2.14) and (2.15) are not necessarily the same, although they are of the same order of magnitude. The same comment applies in (2.19) and (2.20).

We note further that if the conditions in Theorem 2.2 (or Theorem 2.3) hold for some r, then they hold for r replaced by any $s \geq r$, as is obvious for (2.18a); for (2.13) it follows from (4.1). We can then use the same B_n for all $s \geq r$ (since ${}^{(r-1)}\tilde{S}_n/B_n \to 1$ a.s. implies ${}^{(r-1)}\tilde{S}_n/B_n \to_P 1$, and we can use the same B_n for stability in probability whatever the value of r, by Theorem 2.1).

REMARK 2. It will be shown in the proof of Theorem 2.1 (see Lemma 3.2) that (2.3) and the other conditions of the theorem are equivalent to either of the statements

$$P\left\{{}^{(r-1)}\tilde{S}_n \geq T\left|X_n^{(r)}\right|\right\} \to 1, \tag{2.21}$$

$$P\left\{{}^{(r-1)}S_n \geq TM_n^{(r)}\right\} \to 1, \tag{2.22}$$

for some $T > 0$. These conditions, apparently weaker than (2.3) or (2.9), are the analogues of (2.12) and (2.17) for convergence in probability.

REMARK 3. In Theorem 2.3, the conditions express the dominance of the positive parts of the X_i over the negative parts. This is seen most clearly from condition (2.18b), which is equivalent to

$$\frac{\sum_{i=1}^n X_i^-}{\sum_{i=1}^n X_i^+} \to 0 \quad \text{a.s.},$$

where $X_i^+ = \max(X_i, 0)$, $X_i^- = X_i^+ - X_i$. This result is due to Pruitt (1981) and Erickson (1973); see Lemma 5.1 for a generalization.

Since $E|X| = \infty$, (2.18b) can be seen to imply $A_-(x)/A_+(x) \to 0$, $x \to \infty$. Also, one can appeal to Proposition 3.6 of Maller and Resnick (1984), which shows that $A_-(x)/A_+(x) \to 0$, $x \to \infty$, is equivalent to $\sum_{i=1}^n X_i^- / \sum_{i=1}^n X_i^+ \to_P 0$. Thus by Theorem 2.3, if ${}^{(r)}S_n/M_n^{(r)} \to \infty$ a.s., or (2.18) holds, and $E|X| = \infty$ (so that $r \geq 2$), then the dominance of X_i^+ over X_i^-, and of $A_+(x)$ over $A_-(x)$, forces $B_n/n \sim A_+(B_n) \to \infty$ as $n \to \infty$ [see (1.9)]. However, Theorem 2.2 describes quite a different phenomenon; (2.13) may hold with $E|X| = \infty$ in cases where the tails are balanced, that is, $(1 - F(x))/F(-x) \to 1$, $A_+(x)/A_-(x) \to 1$, $x \to \infty$. Surprisingly, we can even have ${}^{(r-1)}\tilde{S}_n/B_n \to 1$ a.s. for some $r \geq 2$, yet $B_n/n \to 0$. For example, take

$$1 - F(x) = \frac{1}{x \log x} - \frac{1}{x \log x (\log \log x)^2}$$

when $x \geq c_1$, and

$$F(-x) = \frac{1}{x \log x}$$

when $x \geq c_2$, where $c_1 > e$ and $c_2 > 1$. Then the values of c_1, c_2 and of F on $(-c_2, c_1)$ can be adjusted so that, for large x,

$$A(x) = \frac{1}{\log \log x}.$$

(2.13) holds for this distribution when $r \geq 2$, so ${}^{(1)}\tilde{S}_n/B_n \to 1$ a.s. yet $B_n \sim nA(B_n) = o(n)$, $n \to \infty$. This example is a random walk that is relatively stable in probability, with $S_n/n \to_P 0$, and with $E|X| = \infty$.

${}^{(r-1)}\tilde{S}_n$ may be relatively stable a.s., without ${}^{(r-1)}S_n$ being so, when $r = 2, 3, \dots$. In fact, (2.13) implies (2.18a) since $A(x) \leq A_+(x)$, but not (2.18b), as the preceding example shows. On the other hand, it can be shown that if $E|X| = \infty$, then (2.18a) and (2.18b) together imply (2.13), so a.s. relative stability of ${}^{(r-1)}S_n$ implies that of ${}^{(r-1)}\tilde{S}_n$ (and then the norming sequences B_n may be taken the same, by Theorem 2.1).

REMARK 4. For those results relating to absolute value trimming in Theorems 2.1 and 2.2, we can obtain equivalences for ${}^{(r)}\tilde{S}_n/B_n \to -1$ and for ${}^{(r)}\tilde{S}_n/|X_n^{(r)}| \to -\infty$ (in probability or a.s.), by interchanging X_i^+ and X_i^-. However, this cannot be done in those results where large values are trimmed. We have partial results relating to ${}^{(r)}S_n/B_n \to -1$ and ${}^{(r)}S_n/M_n^{(r)} \to -\infty$ (in probability or a.s.) which we hope to present elsewhere.

REMARK 5. We saw in Remark 3 that

$${}^{(r)}S_n/M_n^{(r)} \to \infty \quad \text{a.s.} \quad \text{implies} \quad \sum_{i=1}^{n} X_i^- \Big/ \sum_{i=1}^{n} X_i^+ \to_P 0.$$

However, this is not implied by ${}^{(r)}S_n/M_n^{(r)} \to_P \infty$. For a counterexample, we only need a relatively stable S_n for which $A_-(x)/A_+(x)$ does not converge to 0. The example in the previous remark satisfies this; for a simpler example, take $F(x) = 1 - 1/x$, $x \geq 2$, and $F(-x) = 1/(2x)$, $x > 1$, $F(x) = \frac{1}{2}$ otherwise, for which $A_+(x) \sim \log x$ and $A_-(x) \sim \frac{1}{2}\log x$, as $x \to \infty$. Here $S_n/(\frac{1}{2}n \log n) \to_P 1$ [as can be seen by means of (1.4) and (1.9)] but the positive and negative tails are sufficiently close that $A_-(x)/A_+(x)$ does not converge to 0. An example of a relatively stable random walk with $E|X| < \infty$ and $E(X) = 0$, and hence with $0 < \lim_{x\to\infty} A_+(x) < \infty$, $0 < \lim_{x\to\infty} A_-(x) < \infty$ and $S_n/n \to 0$ a.s., originally due to Feller and Breiman, is given in Durrett, Kesten and Lawler (1991).

REMARK 6. The function

$$F(x) = \left\{1 - x^{-1}\beta(\log x)^{\beta-1}\exp\left((\log x)^{\beta}\right)\right\}I(x \geq e)$$

is the distribution function of a nonnegative random variable when $0 < \beta < 1$, for which $A(x) = A_+(x)$ is the slowly varying function

$$A(x) = \int_0^x [1 - F(y)]\, dy = \exp\big((\log x)^\beta\big) = \frac{x[1 - F(x)]}{\beta(\log x)^{\beta - 1}}$$

when $x \geq e$. So the integral in (2.18a) is, if $x_0 \geq e$,

$$I_r = \int_{x_0}^\infty \left\{ \frac{x[1 - F(x)]}{A(x)} \right\}^r \frac{dx}{x} = \beta^r \int_{x_0}^\infty (\log x)^{r(\beta - 1)} \frac{dx}{x} + \text{const.},$$

which converges if and only if $r(1 - \beta) > 1$. Thus, given β, r can be chosen large enough, that is, sufficient terms can be trimmed, for $I_r < \infty$. Given $r > 1$, if $\beta = 1 - 1/r$, we have $I_r = \infty$ but $I_{r+1} < \infty$, thus, in general, one less term may not be enough trimming for (2.16) or (2.19).

Remark 7. One way of expressing our results is to say that we are interested in the case when the large values of X are negligible with respect to the (trimmed) sum. By contrast, Maller and Resnick (1984) and Pruitt (1987) studied the case when the large values are comparable in magnitude to the trimmed sum, that is, when ${}^{(r)}\tilde{S}_n/|X_n^{(r)}|$ or ${}^{(r)}S_n/M_n^{(r)}$ or their absolute values have a finite limit or lim sup. This occurs typically when the tails of F are "heavy," for example, when $P\{|X| > x\}$ is slowly varying. By contrast, in the present paper the tails are relatively "light"; the truncated mean $\nu(x)$ is usually slowly varying, hence can increase no faster than x^ε, for any fixed $\varepsilon > 0$. Section 4 of Maller and Resnick (1984) also considers the case $r = 1$ of the present Theorem 2.2. The case $r = 1$ of (2.11) is comparatively easy to analyze, though, because by Kesten (1971) this can occur if and only if $E|X| < \infty$, while for $r > 1$, (2.11)–(2.15) may hold when $E|X| = \infty$ (see the example in Remark 6).

Generally speaking, the *divergence* to one side (e.g., to $+\infty$) of ${}^{(r)}\tilde{S}_n/|X_n^{(r)}|$ or of ${}^{(r)}S_n/M_n^{(r)}$ seems to put more emphasis on the interplay between the positive and negative parts of the X_i than is encountered in the cases of finite limits such as occur in Theorems 2.3 and 3.1 of Maller and Resnick (1984) and Pruitt (1987). The latter results basically only have conditions on $|X|$ [see Pruitt (1987), Remark 2].

3. Proof of Theorem 2.1.

To prove that (2.3) implies (2.4), the following lemma is required.

Lemma 3.1. *For all $T > 0$, $\varepsilon > 0$, $B_1 > 0$, $B_2 > 0$, $r \geq 0$, there exist constants $C = C(T, \varepsilon, B_1, B_2, r) > 0$, $K = K(T)$ and $n_0 = n_0(T, \varepsilon, B_1, B_2, r)$ with the following property*:

Let $W_i^{(N)}$, $i = 1, 2, \ldots, N$, be independent and identically distributed random variables with distributions function $G^{(N)}$, where $N = 1, 2, \ldots$. Let $\delta > 0$ and define $L(N, \delta)$ to be any $(1 - \delta/N)$-quantile of $G^{(N)}$, that is,

$$G^{(N)}(L(N, \delta) -) \leq 1 - \delta/N \leq G^{(N)}(L(N, \delta)),$$

and define the truncated moments

$$m(N,\delta) = E\{W_1^{(N)} I(W_1^{(N)} < L(N,\delta))\} + L(N,\delta)[1 - \delta/N - G^{(N)}(L(N,\delta)-)],$$

$$s^2(N,\delta) = E\{(W_1^{(N)})^2 I(W_1^{(N)} < L(N,\delta))\} + L^2(N,\delta)[1 - \delta/N - G^{(N)}(L(N,\delta)-)].$$

Assume that $G^{(N)}$ *satisfies the inequalities*

$$\int_{x\le 0} |x|^3 \, dG^{(N)}(x) \le B_1^3 \tag{3.1}$$

and

$$1 - G^{(N)}(B_2) \le \tfrac{1}{16} \quad \text{and} \quad s(N,\delta) \ge 4(B_1 + B_2). \tag{3.2}$$

Then, uniformly in $0 < \varepsilon/4 \le \delta \le 4\varepsilon$ *and* $N \ge n_0$,

$$P\left\{\sum_{i=1}^N W_i^{(N)} \ge Nm(N,\delta) + TN^{1/2}s(N,\delta) + TL(N,\delta)\right\} \ge C, \tag{3.3}$$

and

$$P\left\{\sum_{i=1}^N W_i^{(N)} I(W_i^{(N)} \le L(N,\delta)) \le Nm(N,\delta) - TN^{1/2}s(N,\delta) + K(T)L(N,\delta), \text{ and } W_i^{(N)} \ge L(N,\delta) \text{ for at least } r \text{ values of } i \le N, \text{ but } W_i^{(N)} > L(N,\delta) \text{ for at most } r \text{ values of } i \le N\right\} \ge C. \tag{3.4}$$

We shall not prove Lemma 3.1 here since the proof is essentially the same as that of Lemma 2 of Kesten and Lawler (1992). In fact, (3.3) is proved explicitly there. In our application below, the distribution $G^{(N)}$ will be the conditional distribution of $X_i + Z_i(n)$ given $X_i + Z_i(n) \ge 0$, or of $-X_i - Z_i(n)$ given $X_i + Z_i(n) \le 0$. Here the $Z_i(n)$ are, at first, 0, then later, uniform random variables on $[0, n^{-1/2}]$, and all X_i and $Z_i(n)$ are independent. n and N will be quantities of the same order of magnitude. In particular, $G^{(N)}$ will be concentrated on $[0,\infty)$ so that (3.1) holds with $B_1 = 0$. If $E(X_1^+)^2 = \infty$ [respectively, $E(X_1^-)^2 = \infty$], then (3.2) is also trivial for some B_2 for the above choices of $G^{(N)}$. We must check that (3.3) and (3.4) remain valid even when $E(X_1^+)^2 < \infty$ [respectively, $E(X_1^-)^2 < \infty$]. Let us consider the case when $0 < E(X_1^+)^2 < \infty$ and $G^{(N)}$ is the conditional distribution of $X_i + Z_i(n)$ given $X_i + Z_i(n) \ge 0$, with $Z_i(n)$ as before for some n with $|N - \alpha n| \le n^{1/2}$ and $\alpha = P(X + Z_1(n) \ge 0)$. In this case, let $W_1^{(N)}, \dots, W_N^{(N)}$, be independent with distribution $G^{(N)}$. It is now easy to verify the Lindeberg–Feller conditions, so that

$$\frac{\sum_{i=1}^{N}\left[W_i^{(N)} - E\left(W_1^{(N)}\right)\right]}{\left(N \operatorname{Var}\left(W_1^{(N)}\right)\right)^{1/2}}$$

converges in distribution to a standard normal random variable as $N \to \infty$. Also

$$0 < \lim_{N\to\infty} s^2(N,\delta) = E\left\{X_1^2 | X_1 \geq 0\right\} < \infty,$$

$$L(N,\delta) = o(N^{1/2}),$$

and

$$\begin{aligned} E\left(W_1^{(N)}\right) &\geq E\left\{W_1^{(N)} I\left(W_1^{(N)} < L(N,\delta)\right)\right\} \\ &\quad + L(N,\delta)\left[1 - \delta/N - G^{(N)}(L(N,\delta)-)\right] \\ &= m(N,\delta). \end{aligned}$$

The inequality (3.3) is immediate from this (including the uniformity in δ in the interval $[\varepsilon/4, 4\varepsilon]$). A similar argument works for (3.4).

We now commence the proof that (2.3) implies (2.4). Write $\alpha = P(X \geq 0)$. We assume first that F is continuous, in which case $Z_i(n)$ will be taken as 0 in the discussion following Lemma 3.1, and $W_i^{(N)}$ and $G^{(N)}$ will not depend on N, so we can drop the superfluous superscript N. We remove the restriction of continuity later. For any $T > 0$ and $1 \leq \eta \leq 2$ we will estimate $P\{\Gamma\}$, where

$$\begin{aligned} \Gamma &= \Gamma(\eta, T, r) \\ &= \Bigg\{ \sum_{i=1}^{n} X_i I(X_i \leq L_+(\eta/n)) \\ &\qquad \leq n\mu(\eta/n) - Tn^{1/2}\sigma(\eta/n) - TL_-(\eta/n) \\ &\qquad\quad + K(\alpha^{-1}(2T+2))L_+(\eta/n), \text{ and } M_n^{(r)} \geq L_+(\eta/n) \geq M_n^{(r+1)} \Bigg\}, \end{aligned} \tag{3.5}$$

by decomposing with respect to the set of indices Λ, where $X_i \geq 0$, and applying Lemma 3.1. Write

$$\begin{aligned} P(\Gamma) &= \sum_{\Lambda} P\{X_i \geq 0 \text{ for } i \in \Lambda,\ X_i < 0 \text{ for } i \notin \Lambda \text{ and } \Gamma\} \\ &\geq \sum_{|N-\alpha n| \leq n^{1/2}} P\{X_i \geq 0 \text{ for } i \in \Lambda,\ X_i < 0 \text{ for } i \notin \Lambda\} \\ &\qquad \times P\left\{\Gamma | X_i \geq 0 \text{ for } i \in \Lambda,\ X_i < 0 \text{ for } i \notin \Lambda\right\}, \end{aligned}$$

where $\sum_\Lambda$ denotes summation over all subsets Λ of $\{1, 2, \ldots, n\}$, and N is the cardinality of Λ. Since X_i are independent and identically distributed, we have for fixed Λ,

$$\begin{aligned} &P\{\Gamma | X_i \geq 0 \text{ for } i \in \Lambda,\ X_i < 0 \text{ for } i \notin \Lambda\} \\ &\qquad = P\left\{\Gamma | X_i \geq 0 \text{ for } 1 \leq i \leq N,\ X_i < 0 \text{ for } N+1 \leq i \leq n\right\}. \end{aligned} \tag{3.6}$$

Now let $I_+(\varepsilon) = I(0 \le X \le L_+(\varepsilon))$, $I_-(\varepsilon) = I(-L_-(\varepsilon) \le X \le 0)$, and define

$$\mu_+(\varepsilon) = E(XI_+(\varepsilon)), \qquad \mu_-(\varepsilon) = E(XI_-(\varepsilon)),$$
$$\sigma_+^2(\varepsilon) = E(X^2 I_+(\varepsilon)), \qquad \sigma_-^2(\varepsilon) = E(X^2 I_-(\varepsilon)).$$

Then the right-hand side of (3.6) is at least

$$\begin{aligned} P\Bigg\{ &\sum_{i=1}^{N} X_i I(X_i \le L_+(\eta/n)) \\ &\le n\mu_+(\eta/n) - Tn^{1/2}\sigma_+(\eta/n) + K(\alpha^{-1}(2T+2))L_+(\eta/n), \\ &X_i \ge L_+(\eta/n) \text{ for at least } r \text{ values of } i \le N, \\ &\text{but } X_i > L_+(\eta/n) \text{ for at most } r \text{ values of } i \le N \Big| X_i \ge 0, 1 \le i \le N \Bigg\} \\ &\times P\Bigg\{ \sum_{i=1}^{n-N} (-X_i) \ge -n\mu_-(\eta/n) \\ &\qquad + Tn^{1/2}\sigma_-(\eta/n) + TL_-(\eta/n) \Big| X_i < 0, 1 \le i \le n-N \Bigg\} \\ &= \pi_1 \pi_2, \quad \text{say}. \end{aligned}$$

When $\alpha = 1$ we take $\pi_2 = 1$. For $\alpha < 1$ (and $\alpha > 0$, by assumption of the theorem), we will have $1 < N < n$ for sufficiently large n, when $|N - \alpha n| \le n^{1/2}$, so that π_1 and π_2 will be well defined.

To estimate π_1, apply (3.4) of Lemma 3.1 when $W_i = W_i^{(N)}$ has the conditional distribution of X_i, given $X_i \ge 0$, and $\delta = N\eta/(\alpha n)$. For this choice we may take $L(N, \delta) = L_+(\eta/n)$, since

$$P\{W_1 \ge L_+(\eta/n)\} = P\{X \ge L_+(\eta/n) | X \ge 0\} = \eta/(\alpha n) = \delta/N.$$

Moreover, since we assume continuity of F, we have, letting $G(x) = G^{(N)}(x)$,

$$m(N, \delta) = \int_0^{L(N,\delta)} x\, dG(x) = \alpha^{-1} \int_0^{L_+(\eta/n)} x\, dF(x) = \alpha^{-1}\mu_+(\eta/n),$$

while, similarly,

$$s^2(N, \delta) = \alpha^{-1}\sigma_+^2(\eta/n).$$

Suppose $|N - \alpha n| \le n^{1/2}$ and n is sufficiently large that $n^{1/2} \le \alpha n/2$. Then $\alpha n \ge N - n^{1/2}$ and $N \ge \alpha n - n^{1/2} \ge \alpha n/2 \ge \alpha^2 n/4$, or $n^{1/2} \le 2\alpha^{-1}N^{1/2}$. Thus

$$\begin{aligned} n\mu_+(\eta/n) &= \alpha n m(N, \delta) \\ &\ge Nm(N, \delta) - n^{1/2} m(N, \delta) \\ &\ge Nm(N, \delta) - 2\alpha^{-1}N^{1/2} s(N, \delta), \end{aligned}$$

noting that $m(N, \delta) \le s(N, \delta)$ by the Cauchy–Schwarz inequality. Also

$(\alpha n)^{1/2} \le 2\alpha^{-1/2}N^{1/2} \le 2\alpha^{-1}N^{1/2}$ gives

$$-Tn^{1/2}\sigma_+(\eta/n) = -T(\alpha n)^{1/2}s(N,\delta) \ge -2T\alpha^{-1}N^{1/2}s(N,\delta).$$

Therefore, if $|N - \alpha n| \le n^{1/2}$ and $1/2 \le \eta/2 < \delta < 2\eta \le 4$, we have by (3.4):

$$\begin{aligned}\pi_1 \ge P\Bigg\{ &\sum_{i=1}^{N} W_i I(W_i \le L(N,\delta)) \le Nm(N,\delta) - 2\alpha^{-1}(T+1)N^{1/2}s(N,\delta) \\ &+K(\alpha^{-1}(2T+2))L(N,\delta), \\ &\text{and } W_i \ge L(N,\delta) \text{ for at least } r \text{ values of } i \le N, \\ &\text{but } W_i > L(N,\delta) \text{ for at most } r \text{ values of } i \le N \Bigg\} \\ \ge{}& C(2\alpha^{-1}(T+1),1) > 0.\end{aligned}$$

[Here we have abbreviated $C(T,\varepsilon,B_1,B_2,r)$ to $C(T,\varepsilon)$.] Similarly, applying (3.3) of Lemma 3.1 when W_i has the conditional distribution of $-X_i$, given $X_i < 0$, when F is continuous and with N replaced by $n - N$, gives, when $|N - \alpha n| \le n^{1/2}$,

$$\pi_2 \ge C((2T+2)/(1-\alpha),1) > 0.$$

We therefore conclude from (3.5) and (3.6) that

$$P\{\Gamma\} \ge C(2\alpha^{-1}(T+1),1)C((2T+2)/(1-\alpha),1)P\{|N-\alpha n| \le n^{1/2}\}.$$

Since $P\{|N - \alpha n| \le n^{1/2}\}$ is bounded away from 0 as $n \to \infty$ by the central limit theorem (N is binomial with success probability α), the same is true of $P\{\Gamma\}$. However, when $\Gamma(\eta, T, r)$ occurs,

$$\begin{aligned}{}^{(r)}S_n &\le \sum_{i=1}^{n} X_i I(X_i \le L_+(\eta/n)) \\ &\le n\mu(\eta/n) - Tn^{1/2}\sigma(\eta/n) - TL_-(\eta/n) + K(\alpha^{-1}(2T+2))L_+(\eta/n),\end{aligned}$$

and $M_n^{(r)} \ge L_+(\eta/n)$. Since (2.3) tells us that the probability of

$$\left\{{}^{(r)}S_n \le \left[T + K(\alpha^{-1}(2T+2))\right]L_+(\eta/n) \text{ and } M_n^{(r)} \ge L_+(\eta/n)\right\}$$

tends to 0 as $n \to \infty$, we must have

$$\begin{aligned}&n\mu(\eta/n) - Tn^{1/2}\sigma(\eta/n) - TL_-(\eta/n) + K(\alpha^{-1}(2T+2))L_+(\eta/n) \\ &\qquad > \left[T + K(\alpha^{-1}(2T+2))\right]L_+(\eta/n), \quad \text{eventually}.\end{aligned}$$

This means that for any $T > 0$ and $1 \le \eta \le 2$,

$$\liminf_{n\to\infty} \frac{\mu(\eta/n)}{n^{-1/2}\sigma(\eta/n) + n^{-1}[L_+(\eta/n) + L_-(\eta/n)]} \ge T.$$

Since all estimates are uniform in $1 \le \eta \le 2$, we have when F is continuous

$$\frac{\mu(\varepsilon)}{\varepsilon^{1/2}\sigma(\varepsilon) + \varepsilon[L_+(\varepsilon) + L_-(\varepsilon)]} \to \infty, \qquad \varepsilon \to 0+, \tag{3.7}$$

and this certainly implies (2.4).

To remove the assumption of continuity, we use a method similar to one of Pruitt (1987). Suppose (2.3) holds but there is an infinite sequence n' of integers and a constant a such that

$$\mu(1/n') \le a(n')^{-1/2}\sigma(1/n'). \tag{3.8}$$

We shall show that this leads to a contradiction. Define random variables $X_{in}^{\#} = X_i + Z_i(n)$, where, for each n, $Z_i(n)$ are uniform random variables on $[0, z_n]$, $0 \le z_n \le 1$, and all X_i and $Z_i(n)$ are independent. Let $(M^{\#})_n^{(j)}$ be the j-th largest term among $X_{1n}^{\#}, \dots, X_{in}^{\#}$, which is uniquely defined a.s. since $X_{in}^{\#}$ are continuous random variables. Let

$$^{(r)}S_n^{\#} = \sum_{i=1}^{n} X_{in}^{\#} - (M^{\#})_n^{(1)} - \cdots - (M^{\#})_n^{(r)}, \qquad r = 1, 2, \dots, n.$$

Notice that if $X_m \le M_m^{(j)}$, then

$$X_{mn}^{\#} \le M_m^{(j)} + Z_m(n) \le M_m^{(j)} + 1,$$

so the number of $X_{mn}^{\#}$ greater than $M_m^{(j)} + 1$ is at most $j - 1$. Thus $(M^{\#})_n^{(j)} \le M_n^{(j)} + 1$, and of course $M_n^{(j)} \le (M^{\#})_n^{(j)}$. So

$$|M_n^{(j)} - (M^{\#})_n^{(j)}| \le 1, \qquad j = 1, 2, \dots, n.$$

Also

$$\begin{aligned} ^{(r)}S_n^{\#} &= \sum_{i=1}^{n} (X_i + Z_i(n)) - (M^{\#})_n^{(1)} - \cdots - (M^{\#})_n^{(r)} \\ &= {}^{(r)}S_n + \sum_{i=1}^{n} Z_i(n) + \left[M_n^{(1)} - (M^{\#})_n^{(1)}\right] + \cdots + \left[M_n^{(r)} - (M^{\#})_n^{(r)}\right] \\ &\ge {}^{(r)}S_n - r, \quad \text{a.s.} \end{aligned}$$

Now if X is bounded above w.p.1, then $(M^{\#})_n^{(j)}$ is bounded above, and (2.3) implies $^{(r)}S_n^{\#}/(M^{\#})_n^{(r)} \to_P \infty$. If X is not bounded above, then $(M^{\#})_n^{(j)} \to_P \infty$ and hence, in probability,

$$M_n^{(j)} \sim (M^{\#})_n^{(j)}, \qquad j = 1, 2, \dots, r.$$

Again we see that $^{(r)}S_n^{\#}/(M^{\#})_n^{(r)} \to_P \infty$. Now the $X_{in}^{\#}$ are independent and identically distributed for each n, and the conditional distributions of $X_{in}^{\#}$ given $X_{in}^{\#} \ge 0$ or given $X_{in}^{\#} < 0$ satisfy the conditions of Lemma 3.1 uniformly in n. Therefore, the work up to (3.7) tells us that, for large n, $\mu_n^{\#}(1/n) > 0$ and, as $n \to \infty$,

$$n^{-1/2}\sigma_n^{\#}(1/n) + n^{-1}[L_-^{\#}(1/n) + L_+^{\#}(1/n)] = o\{\mu_n^{\#}(1/n)\}. \tag{3.9}$$

Here, for $n \geq 1$,

$$\mu_n^\#(\varepsilon) = E\{X_{1n}^\# I(L_-^\#(\varepsilon) \leq X_{1n}^\# \leq L_+^\#(\varepsilon))\},$$
$$\sigma_n^{\#2}(\varepsilon) = E\{(X_{1n}^\#)^2 I(L_-^\#(\varepsilon) \leq X_{1n}^\# \leq L_+^\#(\varepsilon))\},$$

and $L_-^\#(\varepsilon)$ and $L_+^\#(\varepsilon)$ are defined by

$$\varepsilon = F_n^\#(-L_-^\#(\varepsilon)) = 1 - F_n^\#(L_+^\#(\varepsilon)),$$

where $F_n^\#$ is the (continuous) distribution function of $X_{1n}^\#$ provided $P(X_{1n}^\# < 0) > 0$, or equivalently $P(X < 0) > 0$. If $P(X < 0) = 0$, $L_-^\#(\varepsilon)$ should be taken equal to 0. Now μ and $\mu_n^\#$, and σ and $\sigma_n^\#$, satisfy the following inequalities for $\varepsilon > 0$:

$$\mu(\varepsilon) \geq \mu_n^\#(\varepsilon) - \varepsilon[L_-^\#(\varepsilon) + L_+^\#(\varepsilon)] - z_n \tag{3.10}$$

and

$$\sigma^2(\varepsilon) \leq 2\sigma_n^{\#2}(\varepsilon) + 2\varepsilon\left[(L_-^\#(\varepsilon))^2 + (L_+^\#(\varepsilon))^2\right] + 4. \tag{3.11}$$

(3.10) is established by writing μ in the form

$$\mu(\varepsilon) = \int_{[-L_-(\varepsilon), L_+(\varepsilon)]} x\,dF(x) = \left\{\int_{[-L_-(\varepsilon), 0]} + \int_{(0, L_+(\varepsilon)]}\right\} x\,dF(x)$$

and integrating by parts to obtain

$$\begin{aligned}\mu(\varepsilon) = {} & -L_+(\varepsilon)[1 - F(L_+(\varepsilon))] \\ & + \int_0^{L_+(\varepsilon)}[1 - F(x)]\,dx + L_-(\varepsilon)F(-L_-(\varepsilon)-) \\ & - \int_{-L_-(\varepsilon)}^0 F(x)\,dx.\end{aligned}$$

The inequalities

$$F_n^\#(x) \leq F(x-) \leq F(x) \leq F_n^\#(x + z_n)$$

and

$$-L_-(\varepsilon) \leq -L_-^\#(\varepsilon) \leq -L_-(\varepsilon) + z_n, \qquad L_+(\varepsilon) \leq L_+^\#(\varepsilon) \leq L_+(\varepsilon) + z_n$$

follow for all x and for $\varepsilon > 0$ from the definitions of the # quantities. Use these together with the inequality $[1 - F(L_+(\varepsilon))] \leq \varepsilon$, to write

$$\begin{aligned}\mu(\varepsilon) &\geq -\varepsilon L_+(\varepsilon) + \int_0^{L_+(\varepsilon)}[1 - F_n^\#(x + z_n)]\,dx - \int_{-L_-(\varepsilon)}^0 F_n^\#(x + z_n)\,dx \\ &= -\varepsilon L_+(\varepsilon) + \int_{z_n}^{L_+(\varepsilon)+z_n}[1 - F_n^\#(x)]\,dx - \int_{-L_-(\varepsilon)+z_n}^{z_n} F_n^\#(x)\,dx \\ &\geq -\varepsilon L_+^\#(\varepsilon) + \int_{z_n}^{L_+^\#(\varepsilon)}[1 - F_n^\#(x)]\,dx - \int_{-L_-^\#(\varepsilon)}^{z_n} F_n^\#(x)\,dx \\ &= -\varepsilon L_+^\#(\varepsilon) + \int_0^{L_+^\#(\varepsilon)}[1 - F_n^\#(x)]\,dx - \int_{-L_-^\#(\varepsilon)}^0 F_n^\#(x)\,dx - z_n.\end{aligned}$$

If we integrate by parts in the last expression we obtain

$$\mu(\varepsilon) \geq \int_{-L_-^\#(\varepsilon)}^{L_+^\#(\varepsilon)} x\, dF_n^\#(x) - \varepsilon L_-^\#(\varepsilon) - \varepsilon L_+^\#(\varepsilon) - z_n,$$

which is (3.10); (3.11) is proved similarly.

Now, by (3.11) with $\varepsilon = n^{-1}$ and the fact that $\liminf \sigma_n^{\#2}(1/n) \geq E(X^+)^2 > 0$, we have

$$\begin{aligned} n^{-1}\sigma^2(1/n) &= O\Big\{n^{-1}\sigma_n^{\#2}(1/n) + 2n^{-2}\Big[\big(L_-^\#(1/n)\big)^2 + \big(L_+^\#(1/n)\big)^2\Big]\Big\} \\ &= o\big\{\mu_n^\#(1/n)\big\}^2, \quad \text{by (3.9)}. \end{aligned}$$

So, by (3.8),

$$\limsup \mu(1/n')/\mu_n^\#(1/n') \leq 0. \tag{3.12}$$

But if we choose $z_n = n^{-1/2} = O\{n^{-1/2}\sigma_n^\#(1/n)\}$, we have by (3.9):

$$z_n/\mu_n^\#(1/n) = O\big\{n^{-1/2}\sigma_n^\#(1/n)/\mu_n^\#(1/n)\big\} = o(1).$$

Thus, by (3.10) with $\varepsilon = n^{-1}$ and (3.9),

$$\begin{aligned} &\mu(1/n)/\mu_n^\#(1/n) \\ &\quad \geq 1 - O\big\{\big[L_-^\#(1/n) + L_+^\#(1/n)\big]/\big[n\mu_n^\#(1/n)\big]\big\} - o(1) \\ &\quad \to 1 \quad \text{as } n \to \infty. \end{aligned} \tag{3.13}$$

(3.12) and (3.13) contradict each other, so (3.8) cannot hold and we have

$$n^{1/2}\frac{\mu(1/n)}{\sigma(1/n)} \to \infty, \quad \text{as } n \to \infty.$$

To prove (2.4) in general from this, take $0 < \varepsilon < 1$ and choose $n = n(\varepsilon)$ as the integer part of $2/\varepsilon$. Then $1/n \leq \varepsilon \leq 2/n$, $L_+(\varepsilon) \leq L_+(1/n)$, $L_-(\varepsilon) \leq L_-(1/n)$ and $\sigma(\varepsilon) \leq \sigma(1/n)$. Thus

$$\begin{aligned} &\left|\mu(\varepsilon) - \mu\left(\frac{1}{n}\right)\right| \\ &\quad = \left|E[XI(-L_-(\varepsilon) \leq X \leq L_+(\varepsilon))] - E\left[XI\left(-L_-\left(\frac{1}{n}\right) \leq X \leq L_+\left(\frac{1}{n}\right)\right)\right]\right| \\ &\quad \leq E\left[XI\left(L_+(\varepsilon) < X \leq L_+\left(\frac{1}{n}\right)\right)\right] + E\left[|X|I\left(-L_-\left(\frac{1}{n}\right) \leq X < -L_-(\varepsilon)\right)\right] \\ &\quad \leq E\left[XI\left(L_+\left(\frac{2}{n}\right) < X \leq L_+\left(\frac{1}{n}\right)\right)\right] + E\left[|X|I\left(-L_-\left(\frac{1}{n}\right) \leq X < -L_-\left(\frac{2}{n}\right)\right)\right] \\ &\quad \leq \left\{\left(\frac{2}{n}\right)E\left[X^2I\left(L_+\left(\frac{2}{n}\right) \leq X \leq L_+\left(\frac{1}{n}\right)\right)\right]\right\}^{1/2} \\ &\qquad + \left\{\left(\frac{2}{n}\right)E\left[X^2I\left(-L_-\left(\frac{1}{n}\right) \leq X \leq -L_-\left(\frac{2}{n}\right)\right)\right]\right\}^{1/2} \\ &\quad \leq 2\left[2n^{-1}\sigma^2\left(\frac{1}{n}\right)\right]^{1/2} \end{aligned} \tag{3.14}$$

by the Cauchy–Schwarz inequality. [Remember that $1 - F(L_+(2/n)) \le 2/n$ and $F(-L_-(2/n)-) \le 2/n$.] Thus $n^{1/2}[\mu(\varepsilon) - \mu(1/n)]/\sigma(1/n)$ is bounded and

$$\frac{n^{1/2}\mu(\varepsilon)}{\sigma(1/n)} \to \infty \quad \text{as } \varepsilon \to 0+.$$

In particular, $\mu(\varepsilon) > 0$ for small $\varepsilon > 0$. But then

$$\frac{2\mu(\varepsilon)}{\varepsilon^{1/2}\sigma(\varepsilon)} \ge \frac{n^{1/2}\mu(\varepsilon)}{\sigma(1/n)} \to \infty \quad \text{as } \varepsilon \to 0+,$$

which proves (2.4) in the general case.

Next we show that (2.4) implies (2.5). This is simply based on Chebyshev's inequality. With $\delta \in (0, \frac{1}{3})$ and $x > 0$, we have

$$\begin{aligned} P&\left\{\left|\sum_{i=1}^n X_i I\left(-L_-\left(\frac{\delta}{n}\right) \le X_i \le L_+\left(\frac{\delta}{n}\right)\right) - n\mu\left(\frac{\delta}{n}\right)\right| > xn\mu\left(\frac{\delta}{n}\right)\right\} \\ &\le \frac{\sigma^2(\delta/n)}{x^2 n\mu^2(\delta/n)} \to 0 \quad \text{as } n \to \infty, \text{ by (2.4)}. \end{aligned}$$

Also

$$\begin{aligned} P&\left\{\left|\sum_{i=1}^n X_i I\left(X_i \notin \left[-L_-\left(\frac{\delta}{n}\right), L_+\left(\frac{\delta}{n}\right)\right]\right)\right| > 0\right\} \\ &\le P\left\{X_i \notin \left[-L_-\left(\frac{\delta}{n}\right), L_+\left(\frac{\delta}{n}\right)\right] \text{ for some } i \le n\right\} \\ &\le n\left[P\left(X > L_+\left(\frac{\delta}{n}\right)\right) + P\left(X < -L_-\left(\frac{\delta}{n}\right)\right)\right] \\ &< 2\delta, \end{aligned}$$

so we have for all $x > 0$,

$$\limsup_{n\to\infty} P\left\{\left|\frac{\sum_{i=1}^n X_i}{n\mu(\delta/n)} - 1\right| > x\right\} \le 2\delta.$$

All we need is to replace $\mu(\delta/n)$ by something independent of δ [$\mu(1/n)$, say] in this. But the same kind of estimates as used for (3.14) give

$$\begin{aligned} &\left|\mu\left(\frac{1}{n}\right) - \mu\left(\frac{\delta}{n}\right)\right| \\ &\quad = \left|E\left[XI\left(-L_-\left(\frac{1}{n}\right) \le X \le L_+\left(\frac{1}{n}\right)\right)\right]\right. \\ &\qquad \left. - E\left[XI\left(-L_-\left(\frac{\delta}{n}\right) \le X \le L_+\left(\frac{\delta}{n}\right)\right)\right]\right| \\ &\quad \le 2\left[n^{-1}\sigma^2\left(\frac{\delta}{n}\right)\right]^{1/2}. \end{aligned} \tag{3.15}$$

By (2.4) we now see that $\mu(1/n) > 0$ and $\mu(\delta/n) \sim \mu(1/n)$ as $n \to \infty$, so we have $S_n/B_n \to_P 1$, where $B_n = n\mu(1/n)$. This in turn implies by the degenerate convergence criterion [Gnedenko and Kolmogorov (1968), page 134] that $nP(|X| > xB_n) \to 0$ for $x > 0$, so $M_n^{(r)}/B_n \to_P 0$, giving ${}^{(r-1)}S_n/B_n \to_P 1$, $r = 1, 2, \dots$. This is (2.5) except for the monotonicity of B_n. That we may replace B_n by the increasing sequence $\max_{k \le n} B_k$ follows as in Rogozin [(1976), page 376].

Next, if (2.5) holds we wish to deduce that $M_n^{(r)}/B_n \to_P 0$. This follows from the inequality

$$8P\left\{|{}^{(r-1)}S_n - B_n| > xB_n\right\} \ge P\left\{|M_n^{(r)}| > 5xB_n\right\}$$

if n is sufficiently large and $0 < 5x < 1$, which is derived as in Lemma 1 of Maller (1982) with continuity, or as in Lemma 3 of Mori (1984). Since, further, $P(M_n^{(r)} \ge 0) \to 1$ as a result of $F(0) < 1$, we have ${}^{(r)}S_n/M_n^{(r)} \to_P \infty$, which is (2.3). This completes the proof of the first three equivalences in Theorem 2.1 when large values are trimmed. [It also shows that the choice of $L_-(\varepsilon)$ and $L_+(\varepsilon)$ is irrelevant.]

We now prove that (2.9) implies (2.3) (with $r = 1$, hence for all r). In fact we prove a little more, namely, that (2.21) implies (2.9) and (2.3), since we will need this in the proof of Theorem 2.2. This is a result of the following lemma.

LEMMA 3.2. *Assume that, for some* $T > 0$,

$$P\left\{{}^{(r-1)}\tilde{S}_n \ge T|X_n^{(r)}|\right\} \to 1. \tag{3.16}$$

Then, for all $j = 1, 2, \dots$,

$$\frac{{}^{(j-1)}\tilde{S}_n}{|X_n^{(j)}|} \to_P \infty.$$

PROOF. Let $L(\varepsilon)$ be such that

$$P\{|X| \ge L(\varepsilon)\} \ge \varepsilon \ge P\{|X| > L(\varepsilon)\}.$$

When $|X| = L(\varepsilon)$ we shall randomize X in a suitable way and count X as a value with $|X| < L(\varepsilon)$ [even though $|X| = L(\varepsilon)$] with a suitable probability such that $P\{|X| < L(\varepsilon)\} = 1 - \varepsilon$. Then, for any $\varepsilon > 0$,

$$P\left\{{}^{(r-1)}\tilde{S}_n \le T|X_n^{(r)}|\right\}$$
$$\ge \binom{n}{r}\left[P\left\{|X| < L\left(\frac{\varepsilon}{n}\right)\right\}\right]^{n-r}\left[P\left\{|X| \ge L\left(\frac{\varepsilon}{n}\right)\right\}\right]^{r} P_{n-r}\left(T, L\left(\frac{\varepsilon}{n}\right)\right),$$

where

$$P_{n-r}(T, L) = P\{S_{n-r}(L) < TL\}$$

and $S_{n-r}(L)$ is the sum of $n - r$ independent copies of a random variable with the conditional distribution of X, given $|X| < L$. Thus, uniformly in $\frac{1}{2} \le \varepsilon \le 2$, (3.16) implies

$$P_{n-r}(T, L(\varepsilon/n)) \to 0, \quad \text{as } n \to \infty,$$

or

$$P_n(T, L(\varepsilon/n)) \to 0, \quad \text{as } n \to \infty.$$

Now consider

$${}^{(j-1)}\tilde{S}_{kn} = \sum_{i=1}^{k} (S_{in} - S_{(i-1)n}) - X_{kn}^{(1)} - \cdots - X_{kn}^{(j-1)}.$$

We have

$$P\left\{ \sum_{i=1}^{k} (S_{in} - S_{(i-1)n}) < kTL(\varepsilon/n) \text{ or } |X_i| \ge L(\varepsilon/n) \text{ for some } i \le kn \right\}$$
$$\le knP\{|X| \ge L(\varepsilon/n)\} + kP\{S_n(L(\varepsilon/n)) < TL(\varepsilon/n)\}$$
$$\le kn\varepsilon/n + k\delta(n, \varepsilon),$$

where $\delta(n, \varepsilon) \to 0$ as $n \to \infty$. But, on the event

$$\{S_{kn} \ge kTL(\varepsilon/n) \quad \text{and } |X_i| \le L(\varepsilon/n) \text{ for all } i \le kn\}$$

we have

$${}^{(j-1)}\tilde{S}_{kn} \ge kTL(\varepsilon/n) - (j-1)L(\varepsilon/n) = (kT - j + 1)L(\varepsilon/n)$$

and $|X_{kn}^{(j)}| \le L(\varepsilon/n)$, whence

$$\frac{{}^{(j-1)}\tilde{S}_{kn}}{|X_{kn}^{(j)}|} \ge kT - j + 1.$$

Consequently, for any $\varepsilon > 0$,

$$\liminf_{n\to\infty} P\left\{ \frac{{}^{(j-1)}\tilde{S}_{kn}}{|X_{kn}^{(j)}|} \ge kT - j + 1 \right\} \ge 1 - k\varepsilon.$$

But then also

$$\liminf_{n\to\infty} P\left\{ \min_{kn \le l < k(n+1)} \frac{{}^{(j-1)}\tilde{S}_l}{|X_l^{(j)}|} \ge kT - j \right\} \ge 1 - (k+1)\varepsilon,$$

because

$$\min_{kn \le l < k(n+1)} {}^{(j-1)}\tilde{S}_l \ge {}^{(j-1)}\tilde{S}_{kn} - \sum_{i=kn+1}^{k(n+1)} |X_i|,$$
$$\max_{kn \le l < k(n+1)} |X_l^{(j)}| \le |X_{kn}^{(j)}|$$

on the event

$$\{|X_i| < |X_{kn}^{(j)}|, \text{ for } kn + 1 \le i < k(n+1)\},$$

and

$$\frac{\sum_{i=kn+1}^{k(n+1)} |X_i|}{|X_{kn}^{(j)}|} \to_P 0.$$

Thus, for all $\varepsilon > 0$,

$$\liminf_{l \to \infty} P\left\{ {}^{(j-1)}\tilde{S}_l \geq (kT - j)|X_l^{(j)}| \right\} \geq 1 - (k+1)\varepsilon,$$

or

$$P\left\{ {}^{(j-1)}\tilde{S}_l \geq (kT - j)|X_l^{(j)}| \right\} \to 1.$$

Since this holds for all k, the lemma follows. □

Taking $j = 1$ in Lemma 3.2, we see that (2.21) implies $S_n/|X_n^{(1)}| \to_P \infty$. Hence $S_n/M_n^{(1)} \to_P \infty$, since clearly $|X_n^{(1)}| \geq M_n^{(1)}$ a.s. Thus (2.3) holds with $r = 1$ and so, by the first part of the proof, $S_n/B_n \to_P 1$, giving ${}^{(r-1)}\tilde{S}_n/B_n \to_P 1$ again by the degenerate convergence criterion. This is (2.10). Finally, ${}^{(r-1)}\tilde{S}_n/B_n \to_P 1$ implies $|X_n^{(r)}|/B_n \to_P 0$ [as in the implication from (2.5) to (2.3)] and hence $|X_n^{(1)}|/B_n \to_P 0$. Thus ${}^{(r)}\tilde{S}_n/|X_n^{(r)}| \to_P \infty$. This shows the equivalence of (2.3)–(2.5) with (2.9) and (2.10), and indeed with (2.21). Similarly, one can show that these are equivalent to (2.22).

Note that we have also shown that (2.5) or (2.9) for any r implies (2.5) for $r = 1$, possibly with a different sequence B_n, B_n^1, say. However, as before, $S_n/B_n^1 \to_P 1$ implies $X_n^{(1)}/B_n^1 \to_P 0$ and hence $X_n^{(j)}/B_n^1 \to_P 0$ for each j. Thus ${}^{(r-1)}S_n/B_n^1 \to_P 1$ and ${}^{(r-1)}\tilde{S}_n^1 \to_P 1$, which shows that B_n can be taken equal to B_n^1 for all n in (2.5) and (2.10).

The remaining equivalences in Theorem 2.1 can be proved either as conditions equivalent to relative stability (in probability) of S_n [see Maller (1978, 1979) for (2.6) and (2.7)] or as purely analytical equivalences among (2.4), (2.6), (2.7) and (2.8). We sketch these proofs here since they may be of interest.

Let (2.4) hold. Then for fixed $0 < \delta < 1$,

$$\int_{[-L_-(\delta\varepsilon),\, L_+(\delta\varepsilon)]} x\, dF(x) \Bigg/ \left\{ \varepsilon \int_{[-L_-(\delta\varepsilon),\, L_+(\delta\varepsilon)]} x^2\, dF(x) \right\}^{1/2} \to \infty, \qquad \varepsilon \to 0.$$

Since

$$\begin{aligned} \int_{[-L_-(\delta\varepsilon),\, L_+(\delta\varepsilon)]} x^2\, dF(x) &\geq L_+^2(\varepsilon) \int_{[L_+(\varepsilon),\, L_+(\delta\varepsilon)]} dF(x) \\ &\quad + L_-^2(\varepsilon) \int_{[-L_-(\delta\varepsilon),\, -L_-(\varepsilon)]} dF(x) \\ &\geq L_+^2(\varepsilon)(1-\delta)\varepsilon + L_-^2(\varepsilon)(1-\delta)\varepsilon, \end{aligned}$$

this means [for any permissible choice of $L_-(\varepsilon)$ or $L_+(\varepsilon)$] that

$$\int_{[-L_-(\delta\varepsilon),\, L_+(\delta\varepsilon)]} x\, dF(x)/\{\varepsilon[L_+(\varepsilon) + L_-(\varepsilon)]\} \to \infty, \qquad \varepsilon \to 0+.$$

Finally, we saw in (3.15) (with n^{-1} replaced by ε) that

$$\mu(\delta\varepsilon) - \mu(\varepsilon) = O\big[\varepsilon^{1/2}\sigma(\delta\varepsilon)\big]$$

as $\varepsilon \to 0$, so that also

$$\text{(3.17)} \quad \int_{[-L_-(\varepsilon),\, L_+(\varepsilon)]} x\, dF(x)/\{\varepsilon[L_+(\varepsilon) + L_-(\varepsilon)]\} \to \infty, \qquad \varepsilon \to 0+.$$

Now, given a large $x > 0$, define $\varepsilon = \varepsilon(x)$ by $\varepsilon = \frac{1}{2}[P(X > x) + P(X < -x)]$. If $P(X > x) \geq P(X < -x)$, then $P(X > x) \geq \varepsilon \geq P(X < -x)$ so we may take $L_+(\varepsilon) \geq x \geq L_-(\varepsilon)$. If $P(X < -x) \geq P(X > x)$, then $P(X > x) \leq \varepsilon \leq P(X < -x)$ so $L_-(\varepsilon) \geq x \geq L_+(\varepsilon)$. When $L_+(\varepsilon) \geq x \geq L_-(\varepsilon)$ we have

$$\begin{aligned}
\mu(\varepsilon) &= \int_{[0,\, L_+(\varepsilon)]} y\, dF(y) - \int_{[-L_-(\varepsilon),\, 0)} |y|\, dF(y) \\
&= \left\{\int_{[0,x]} + \int_{(x,\, L_+(\varepsilon)]}\right\} y\, dF(y) - \left\{\int_{[-x,0]} - \int_{[-x,\, L_-(\varepsilon))}\right\} |y|\, dF(y) \\
&\leq \int_{[0,x]} y\, dF(y) + \int_{[-x,0]} y\, dF(y) + L_+(\varepsilon)[1 - F(x)] + xF(-L_-(\varepsilon)-) \\
&\leq \nu(x) + L_+(\varepsilon)[1 - F(x) + \varepsilon] \leq \nu(x) + 3\varepsilon L_+(\varepsilon),
\end{aligned}$$

so that

$$\text{(3.18)} \qquad \frac{\mu(\varepsilon)}{\varepsilon[L_+(\varepsilon) + L_-(\varepsilon)]} \leq \frac{2\nu(x)}{xP(|X| > x)} + 3.$$

When $L_-(\varepsilon) \geq x \geq L_+(\varepsilon)$ we have

$$\begin{aligned}
\mu(\varepsilon) = &\left\{\int_{[0,x]} - \int_{(L_+(\varepsilon),\, x]}\right\} y\, dF(y) \\
&- \left\{\int_{[-x,0]} + \int_{[-L_-(\varepsilon),\, -x)}\right\} |y|\, dF(y) \leq \nu(x),
\end{aligned}$$

so again (3.18) holds. This shows, via (3.17), that (2.4) implies (2.6).

Now (2.6) implies that $\nu(x) > 0$ for x sufficiently large, and also that ν is slowly varying as $x \to \infty$. To see this, let $x > 0$ and $\lambda > 1$. Then

$$\begin{aligned}
\frac{|\nu(\lambda x) - \nu(x)|}{\nu(x)} &= \frac{\left|\int_{(x,\, \lambda x]} y\, d[F(y) + F(-y)]\right|}{\nu(x)} \\
&\leq \frac{\lambda x P(|X| > x)}{\nu(x)},
\end{aligned}$$

so $\nu(\lambda x) \sim \nu(x)$ as $x \to \infty$, that is, ν is slowly varying. For (2.7) fix $\varepsilon > 0$ and choose x_0 so large that $xP(|X| > x) \le \varepsilon\nu(x)$ for $x \ge x_0$. Then

$$V(x) = -\int_{[0,x]} y^2\, dP(|X| > y) \le 2\int_{[0,x]} yP(|X| > y)\, dy$$

$$\le x_0^2 + 2\varepsilon\int_{x_0}^{x} \nu(y)\, dy \le x_0^2 + 3\varepsilon x\nu(x)$$

by a property of slow variation [Bingham, Goldie and Teugels (1987), Proposition 1.5.8]. If $V(x) \to \infty$ as $x \to \infty$, then (2.7) follows. If $V(x)$ is bounded, then $\lim_{x\to\infty} \nu(x)$ exists. Since $\nu(x) > 0$ for large x, the limit must be positive. If it is strictly positive, then (2.7) again holds. Finally, $\lim_{x\to\infty} \nu(x) = 0$ and the boundedness of $V(x)$ is excluded by (2.6). Indeed $\nu(x) > 0$, $\nu(x) \to 0$ and ν slowly varying imply $\nu(x) > x^{-\varepsilon}$ eventually, for any fixed $\varepsilon > 0$ [Bingham, Goldie and Teugels (1987), Proposition 1.3.6(v)], while mean zero and finite variance for X imply

$$|\nu(x)| = \left| -\int_{|y|>x} y\, dF(y) \right| \le \int_{|y|>x} \frac{y^2\, dF(y)}{x} = o\left(\frac{1}{x}\right).$$

Thus (2.6) always implies (2.7).

Conversely, assume (2.7). Then for any fixed $0 < \varepsilon < 1$ and $\varepsilon x \le y \le x$,

$$|\nu(x) - \nu(y)| = \left| \int_{y<|z|\le x} z\, dF(z) \right| \le \int_{|z|\le x} \frac{z^2\, dF(z)}{y}$$

$$\le \frac{V(x)}{\varepsilon x} = o(\nu(x)), \quad \text{as } x \to \infty.$$

Thus ν is again slowly varying. Also, for fixed $\eta > 0$,

$$P(x < |X| \le 2x) \le \frac{V(2x)}{x^2} = o\left(\frac{\nu(2x)}{x}\right) \le \frac{\eta\nu(x)}{x}, \quad \text{eventually}.$$

Consequently,

$$P(|X| > x) \le \sum_{k=0}^{\infty} P(2^k x < |X| \le 2^{k+1}x) \le \eta \sum_{k=0}^{\infty} \frac{\nu(2^k x)}{2^k x} \le \frac{4\eta\nu(x)}{x},$$

as soon as $\nu(2y) \le 3\nu(y)/2$ for $y \ge x$. Thus (2.7) implies (2.6).

Next, assume (2.6) or (2.7) (and hence both) hold. Then since ν is slowly varying, we obtain by taking limits from the left in (2.6) that also

$$\frac{\nu(x)}{xP(|X| \ge x)} \to \infty, \qquad x \to \infty.$$

Multiplying this by (2.7) yields (2.8).

To complete the chain of implications in Theorem 2.1, we show that (2.8) implies (2.5). To prove this, we wish to choose B_n such that

$$(3.19)\qquad B_n \uparrow \infty, \qquad nP(|X| > B_n) \to 0 \quad \text{and} \quad \frac{V(B_n)}{[n\nu^2(B_n)]} \to 0, \qquad n \to \infty.$$

If we can find such B_n, then

$$P\left\{S_n \neq \sum_{i=1}^{n} X_i I(|X_i| \le B_n)\right\} \le nP(|X| > B_n) \to 0, \qquad n \to \infty,$$

while, by Chebyshev's inequality,

$$\frac{\sum_{i=1}^{n} X_i I(|X_i| \le B_n)}{n\nu(B_n)} \to_P 1.$$

Thus (3.19) will imply (2.5) for $r = 1$ and hence (2.3) and (2.5) for all r. It remains to construct B_n satisfying (3.19). Assume (2.8) holds and define

$$x_n = \inf\{x\colon nP(|X| \ge x) \le 1\}$$

and

$$k_n^2 = \sup_{x \ge x_n} \left\{\frac{V(x)P(|X| \ge x)}{\nu^2(x)}\right\}.$$

Then $x_n \uparrow \infty$ and $k_n \downarrow 0$. Finally, let

$$B_n = \inf\{x\colon nP(|X| > x) \le k_n\}.$$

Then

$$B_n \uparrow \infty \quad \text{and} \quad nP(|X| > B_n) \le k_n \to 0, \qquad n \to \infty.$$

Also $nP(|X| \ge B_n) \ge k_n$ and $B_n \ge x_n$, so that, by the definition of k_n,

$$\frac{V(B_n)P(|X| \ge B_n)}{\nu^2(B_n)} \le k_n^2$$

and

$$\frac{V(B_n)}{[n\nu^2(B_n)]} \le \frac{k_n^2}{k_n} = k_n \to 0, \quad \text{as } n \to \infty,$$

as required.

REMARKS 8. Note that we have also proved that (3.7) and (3.17) are equivalent to the conditions of Theorem 2.1.

REMARK 9. When (2.4) holds, it is not difficult to show that $\mu(\varepsilon) \sim \nu(L(\varepsilon))$, as $\varepsilon \to 0$, where $L(\varepsilon) = [L_-(\varepsilon) + L_+(\varepsilon)]/2$.

4. Proof of Theorem 2.2. We start with assuming the analytic condition (2.13). This implies

$$\frac{A(x)}{xP\{|X| > x\}} \to \infty \quad \text{as } x \to \infty. \tag{4.1}$$

To see this, let

$$f(x) = \left\{ \frac{xH(x)}{A(x)} \right\}^r,$$

where $H(x) = P(|X| > x)$. Then (2.13) says that $\int_{x_0}^{\infty} f(x)\,dx/x$ converges. This implies that

$$\inf_{e^k < x \le e^{k+1}} f(x) \to 0, \qquad k \to \infty. \tag{4.2}$$

Indeed if (4.2) failed there would be a $\delta > 0$ and an infinite set K such that $f(x) \ge \delta$ for all $x \in (e^k, e^{k+1}]$ and $k \in K$, but then

$$\int_{x_0}^{\infty} \frac{f(x)\,dx}{x} \ge \sum_{k \in K} \int_{e^k}^{e^{k+1}} \frac{\delta\,dx}{x} \ge \delta \sum_{k \in K} 1 = \infty,$$

giving a contradiction. Thus (4.2) holds and this implies that for each k we can choose $x_k \in (e^k, e^{k+1}]$ such that $f(x_k) \to 0$, as $k \to \infty$. Equivalently, (4.1) holds along the sequence x_k. To prove (4.1) itself, take a large x, then choose $k = k(x)$ so that $x_k \le x < x_{k+1}$. Then

$$x_k \le x < x_{k+1} \le e^{k+2} \le e^2 x_k.$$

Let $T > 0$ be fixed and assume without loss of generality that k is so large that $A(x_k) \ge T x_k H(x_k)$. Then

$$\begin{aligned} A(x) &= A(x_k) + \int_{x_k}^{x} [1 - F(y) - F(-y)]\,dy \\ &\ge T x_k H(x_k) - (x - x_k) H(x_k) \\ &\ge (T - e^2) x_k H(x_k) \\ &\ge (T - e^2) e^{-2} x H(x). \end{aligned}$$

Thus (4.1) holds.

We next show that $x/A(x)$ is eventually strictly increasing. Note that the differential

$$\begin{aligned} d\left[\frac{x}{A(x)}\right] &= \frac{1}{A(x)}\,dx - \frac{x}{A^2(x)}[1 - F(x) - F(-x)]\,dx \\ &= \frac{1 + o(1)}{A(x)}\,dx \end{aligned} \tag{4.3}$$

as $x \to \infty$, by (4.1). This is positive for large x since $A(x)$ is positive for such x. This strict monotonicity of $A(x)/x$ means that it has no "flat spots"; thus if we define

$$B(x) = \inf\left\{y\colon \frac{y}{A(y)} > x\right\}, \tag{4.4}$$

then, for large $x > 0$, $B(x)$ is positive, continuous and strictly increasing and satisfies

$$\frac{B(x)}{A(B(x))} = x. \tag{4.5}$$

Now (2.13) and (4.3) imply

$$\int_{x_0}^{\infty} [P\{|X| > x\}]^r \, d\left[\frac{x}{A(x)}\right]^r < \infty.$$

This means the integral J_r of Mori (1977) is finite, with Mori's $B(x)$ replaced by $x/A(x)$. The inverse function of this is just our $B(\cdot)$. The conclusion of Theorem 1 of Mori (1977) then gives

$$\frac{{}^{(r-1)}\tilde{S}_n - nA(B(n))}{B(n)} \to 0 \quad \text{a.s.}, \tag{4.6}$$

provided we check Mori's hypotheses:

$$\frac{B(x)}{x^{1/\alpha}} \quad \text{is increasing for some } 0 < \alpha < 2$$

and

$$\frac{B(2x)}{B(x)} \quad \text{is bounded}.$$

These follow from Bingham, Goldie and Teugels [(1987), Theorem 1.5.13] since $A(x)$ and hence also $B(x)/x$ is slowly varying, as follows from (4.1). Thus indeed (4.6) holds. Then since by (4.5) we have $nA(B(n)) = B(n)$, (4.6) implies

$$\frac{{}^{(r-1)}\tilde{S}_n}{B(n)} \to 1 \quad \text{a.s.}$$

This is (2.14), and (2.15) follows from (2.14).

Now assume (2.15). To prove (2.11), we need the following lemma, which provides a simpler proof of results of Kesten [(1972), equation (4.15)] and Maller [(1984), Theorem 3].

LEMMA 4.1. *Let $F(0) < 1$. If C_n and B_n are nonstochastic sequences with $B_n > 0$, $B_n \uparrow \infty$ and $r = 1, 2, \ldots,$ then either of*

$$\limsup_{n\to\infty} \frac{|{}^{(r-1)}\tilde{S}_n - C_n|}{B_n} < \infty \quad a.s. \quad \text{or} \quad \limsup_{n\to\infty} \frac{|{}^{(r-1)}S_n - C_n|}{B_n} < \infty \quad a.s.$$

implies $(S_n - n\nu(B_n))/B_n \to_P 0$ *and hence, for all* $j \geq 0$,

$$\frac{{}^{(j)}\tilde{S}_n - n\nu(B_n)}{B_n} \to_P 0 \quad \text{and} \quad \frac{{}^{(j)}S_n - n\nu(B_n)}{B_n} \to_P 0.$$

PROOF. We need the following relations:

$$|{}^{(r-1)}\tilde{S}_n - {}^{(r-1)}\tilde{S}_{n-1}| = |X_n| \wedge |X_n^{(r)}| \tag{4.7}$$

and

$$|X_n^{(r)}| = \max_{r \leq j \leq n} |{}^{(r-1)}\tilde{S}_j - {}^{(r-1)}\tilde{S}_{j-1}|. \tag{4.8}$$

The first of these is due to Mori [(1976), page 192, equation (4)], and the second follows easily from the first. By the Hewitt–Savage 0–1 law, $\limsup |{}^{(r-1)}\tilde{S}_n - C_n|/B_n$ is a constant a.s., which by the first version of the condition of the lemma is finite, so, for some $c > 0$,

$$P\Big\{\max_{r \leq j \leq n} |{}^{(r-1)}\tilde{S}_j - C_j| > cB_n\Big\} \to 0, \qquad n \to \infty.$$

Thus (taking $C_0 = 0$) we have from (4.7) that

$$|C_n - C_{n-1}| \leq |{}^{(r-1)}\tilde{S}_n - C_n| + |{}^{(r-1)}\tilde{S}_{n-1} - C_{n-1}| + |X_n|. \tag{4.9}$$

Since $X_n/B_n \to_P 0$ this means that

$$\limsup_{n \to \infty} |C_n - C_{n-1}|/B_n \leq 2c,$$

which further implies that

$$\limsup_{n \to \infty} \max_{1 \leq j \leq n} |C_j - C_{j-1}|/B_n \leq 2c.$$

From (4.8) we now obtain

$$P\big\{|X_n^{(r)}| > (4c+1)B_n\big\} \to 0, \qquad n \to \infty,$$

and since the number of X_i with $|X_i| > (4c+1)B_n$ has a binomial distribution with n trials and success probability $P(|X| > (4c+1)B_n)$, we see that also

$$P\big\{|X_n^{(1)}| > (4c+1)B_n\big\} \to 0, \qquad n \to \infty.$$

Thus, for $r = 2, 3, \dots,$

$$P\big\{|X_n^{(1)}| + \cdots + |X_n^{(r-1)}| > (4c+1)(r-1)B_n\big\} \to 0, \qquad n \to \infty.$$

Then since

$$S_n - C_n = {}^{(r-1)}\tilde{S}_n - C_n + X_n^{(1)} + \cdots + X_n^{(r-1)},$$

we obtain

$$P\big\{|S_n - C_n| > (4c+1)rB_n\big\} \to 0. \tag{4.10}$$

Now argue as follows. Given any sequence $n' \to \infty$ of integers, use Helly's theorem to find a subsequence so that $(S_{n'} - C_{n'})/B_{n'} \to_D Z'$, where Z' is a random variable which is proper by (4.10) and infinitely divisible. We must have

$$P(|Z'| > (4c + 1)r) = 0,$$

and a result of Feller [(1971), page 177] tells us that Z' is constant a.s. So we have shown $(S_{n'} - C_{n'})/B_{n'} \to_P Z'$. Symmetrizing in the usual way gives $S_{n'}^S/B_{n'} \to_P 0$, which holds for all subsequences, so $S_n^S/B_n \to_P 0$ and $(S_n - n\nu(B_n))/B_n \to_P 0$, where the choice of $n\nu(B_n)$ as centering constants is justified by Gnedenko and Kolmogorov [(1968), page 124]. The same reference shows that $nP(|X| > \varepsilon B_n) \to 0$ for every $\varepsilon > 0$, whence $|S_n - {}^{(j)}\tilde{S}_n|/B_n \to_P 0$. This proves the lemma when trimming by absolute values.

For trimming by large values, the proof is similar if (4.7) and (4.8) are replaced by

$$^{(r-1)}S_n - {}^{(r-1)}S_{n-1} = X_n \wedge M_n^{(r)}$$

and

$$M_n^{(r)} = \max\left\{^{(r-1)}S_j - {}^{(r-1)}S_{j-1} \colon r \le j \le n\right\}$$

for $r = 1, 2, \dots$ and $n \ge r$. We omit the details of this. □

Returning to the main proof, if (2.15) holds, then, by Lemma 4.1, $(^{(r-1)}\tilde{S}_n - C_n)/B_n \to_P 0$, where $C_n = n\nu(B_n)$. This means $C_n \asymp B_n$, otherwise we could take a subsequence through which $^{(r-1)}\tilde{S}_n/B_n \to 0$ or ∞ a.s., contradicting (2.15). By (4.8) and (2.15) we have

$$\limsup_{n\to\infty} |X_n^{(r)}|/B_n = \limsup_{n\to\infty} \max_{r \le j \le n} |^{(r-1)}\tilde{S}_j - {}^{(r-1)}\tilde{S}_{j-1}|/B_n < \infty \quad \text{a.s.},$$

so Lemma 3 of Mori (1976) shows that

$$\sum_{n\ge 1} n^{r-1}[1 - F(\varepsilon B_n) + F(-\varepsilon B_n)]^r < \infty \tag{4.11}$$

holds for some $\varepsilon > 0$. Also, by Lemma 4.1,

$$(S_n/C_n) - 1 = [(S_n - C_n)/B_n](B_n/C_n) \to_P 0,$$

showing that S_n is relatively stable with norming sequence C_n. The sequence C_n is regularly varying of index 1 [see (1.9) and Bingham, Goldie and Teugels (1987), Theorem 1.5.13], and, since $B_n \asymp C_n$, $B_n \le B_{\lambda n}/2$ for some $\lambda > 1$ and all large n. This means that (4.11) holds for every $\varepsilon > 0$, by, for example, a proof like that of Lemma 1 of Maller (1978). Lemma 3 of Mori (1976) then gives $X_n^{(r)}/B_n \to 0$ a.s., $n \to \infty$, which together with (2.15) implies (2.11).

Trivially (2.11) implies (2.12), and we now show that (2.12) implies (2.13). Define $L_1 = 1 < L_2 < \cdots$ as the successive indices at which $|X_n|$ takes on a maximum, that is,

$$|X_n| \ge \max\{|X_i| \colon i < n\} \quad \text{if and only if } n \text{ is one of the } L_j.$$

Define for $T > 0$ and $\varepsilon > 0$ the following events:

$$\Gamma(n,r,T,\varepsilon) = \left\{0 \leq {}^{(r)}\tilde{S}_n \leq T|X_n^{(r)}| \text{ and } |X_n^{(r+1)}| \leq \varepsilon |X_n^{(r)}|\right\} \tag{4.12}$$

and

$$\Delta(j,r,T,\varepsilon) = \left\{j \text{ is some } L_n, {}^{(r)}\tilde{S}_j \leq T|X_j^{(r)}| \text{ and } |X_j^{(r+1)}| \leq \varepsilon |X_j^{(r)}|\right\}. \tag{4.13}$$

We first show that a necessary condition for (2.12) is that

$$\sum_1^\infty P\{\Gamma(L_n,r,T,\varepsilon)\} < \infty \tag{4.14}$$

for all $T > 0$, $\varepsilon > 0$, for which

$$\liminf_{n\to\infty} {}^{(r)}\tilde{S}_n/|X_n^{(r)}| > T + r\varepsilon \quad \text{a.s.}$$

Such T and ε exist, since by the Hewitt–Savage zero–one law, $\liminf_{n\to\infty} {}^{(r)}\tilde{S}_n/|X_n^{(r)}|$ is a constant a.s.

Assume that the sum in (4.14) diverges for some such T and ε. Let $m > n + r$ and assume that $\Gamma(L_n, r, T, \varepsilon)$ and $\Gamma(L_m, r, T, \varepsilon)$ occur. Then the $r + 1$ largest values among $|X_{L_n+1}|, \dots, |X_{L_m}|$ are just $|X_{L_m}^{(1)}|, \dots, |X_{L_m}^{(r+1)}|$. Hence

$$\begin{aligned}
&\sum_{L_n+1}^{L_m} X_i - r \text{ largest (in absolute value) among } X_{L_n+1}, \dots, X_{L_m} \\
&\quad = {}^{(r)}\tilde{S}_{L_m} - S_{L_n} \\
&\quad \leq T|X_{L_m}^{(r)}| - S_{L_n} \\
&\quad \leq T|X_{L_m}^{(r)}| - {}^{(r)}\tilde{S}_{L_n} + r|X_{L_n}^{(1)}| \\
&\quad \leq T|X_{L_m}^{(r)}| + r|X_{L_m}^{(r+1)}| \\
&\quad \leq (T + r\varepsilon)|X_{L_m}^{(r)}|.
\end{aligned}$$

Hence on $\Gamma(L_n, r, T, \varepsilon) \cap \Gamma(L_m, r, T, \varepsilon)$, also $\Delta((L_m - L_n), r, (T + r\varepsilon), \varepsilon)$ occurs among $X_{L_n+1}, \dots, X_{L_m}$, and

$$\begin{aligned}
&\Big\{\Gamma(L_n,r,T,\varepsilon) \cap \big[\Delta(l,r,(T+r\varepsilon),\varepsilon) \\
&\qquad\qquad \text{occurs among } X_{L_n+1}, X_{L_n+2}, \dots \text{ at most } N \text{ times}\big]\Big\} \\
&\quad \subset \Big\{\Gamma(L_n,r,T,\varepsilon) \cap \big[\Gamma(L_m,r,T,\varepsilon) \\
&\qquad\qquad \text{occurs for at most } (N + r) \text{ values of } m > n\big]\Big\}.
\end{aligned}$$

Consequently,

$$\sum_{n=1}^{\infty} P\{\Gamma(L_n, r, T, \varepsilon)\} P\{\Delta(j, r, (T + r\varepsilon), \varepsilon) \text{ occurs for at most } N \text{ values of } j\}$$

$$\le \sum_{n=1}^{\infty} P\{\Gamma(L_n, r, T, \varepsilon) \cap [\Gamma(L_m, r, T, \varepsilon) \text{ occurs for at most } (N + r) \text{ values of } m > n]\}$$

$$\le N + r + 1.$$

Since we assumed that

$$\sum_1^{\infty} P\{\Gamma(L_n, r, T, \varepsilon)\} = \infty,$$

this shows that

$$P\{\Delta(j, r, (T + r\varepsilon), \varepsilon) \text{ occurs for at most } N \text{ values of } j\} = 0.$$

This proves that w.p.1 $\Delta(j, r, (T + r\varepsilon), \varepsilon)$ occurs for infinitely many j. However, $\Delta(j, r, T + r\varepsilon, \varepsilon)$ implies that ${}^{(r)}\tilde{S}_j \le (T + r\varepsilon)|X_j^{(r)}|$. Thus, for our choice of T and ε, (2.12) shows that $\Delta(j, r, (T + r\varepsilon), \varepsilon)$ occurs a.s. only for finitely many j, and we have arrived at a contradiction. In other words, (2.11) implies (4.14), as claimed.

Now,

$$\sum_1^{\infty} P\{\Gamma(L_n, r, T, \varepsilon)\}$$

$$= E\Big\{\text{number of } n \text{ for which } 0 \le {}^{(r)}\tilde{S}_{L_n} \le T|X_{L_n}^{(r)}| \text{ and } |X_{L_n}^{(r+1)}| \le \varepsilon |X_{L_n}^{(r)}|\Big\}$$

$$= E\Big\{\text{number of } j \text{ for which } j \text{ equals some } L_n,\ 0 \le {}^{(r)}\tilde{S}_j \le T|X_j^{(r)}| \text{ and } |X_j^{(r+1)}| \le \varepsilon |X_j^{(r)}|\Big\}$$

$$= \sum_1^{\infty} P\Big\{\Delta(j, r, T, \varepsilon) \cap \big({}^{(r)}\tilde{S}_j \ge 0\big)\Big\}.$$

Thus we also have

$$\sum_1^{\infty} P\Big\{\Delta(n, r, T, \varepsilon) \cap \big({}^{(r)}\tilde{S}_n \ge 0\big)\Big\} < \infty \tag{4.15}$$

for the same T and ε as in (4.14). Next we observe that, for $r \geq 2$,

$$\begin{aligned}
&P\Big\{\Delta(n,r,T,\varepsilon) \cap \big({}^{(r)}\tilde{S}_n \geq 0\big)\Big\} \\
&\quad = P\Big\{0 \leq {}^{(r)}\tilde{S}_n \leq T|X_n^{(r)}|,\ |X_n^{(1)}| = |X_n|,\ |X_n^{(r+1)}| \leq \varepsilon |X_n^{(r)}|\Big\} \\
&\quad = P\Big\{|X_n^{(r)}| \geq \varepsilon^{-1}|X_n^{(r+1)}|,\ 0 \leq {}^{(r)}\tilde{S}_n \leq T|X_n^{(r)}|,\ |X_n^{(1)}| = |X_n|\Big\} \\
&\quad \geq \frac{1}{r} E\Big\{\#j\colon |X_j| = |X_n^{(r)}| \geq \varepsilon^{-1}|X_n^{(r+1)}|,\ 0 \leq {}^{(r)}\tilde{S}_n \leq T|X_n^{(r)}|,\ |X_n^{(1)}| = |X_n|\Big\} \\
&\quad = \frac{1}{r}\sum_{j=1}^{n-1} \sum\nolimits^{*} P\Big\{0 \leq T^{-1}\sum\nolimits^{\#} X_i \leq |X_j| \leq |X_{i_1}| \leq \cdots \leq |X_{i_{r-2}}| \leq |X_n|, \\
&\qquad\qquad |X_i| \leq \varepsilon |X_j| \text{ for } i \notin \{i_1,\dots,i_{r-2},j,n\}\Big\}.
\end{aligned}$$

Here the notation Σ^* denotes summation over subsets of $(r-2)$ distinct integers in $[1, n-1] \setminus \{j\}$, and the notation $\Sigma^{\#}$ denotes summation over $1 \leq i \leq n-1$, $i \notin \{i_1,\dots,i_{r-2}, j\}$. It follows that, for some constant C_1,

$$\begin{aligned}
&P\Big\{\Delta(n,r,T,\varepsilon) \cap \big({}^{(r)}\tilde{S}_n \geq 0\big)\Big\} \\
&\quad \geq \frac{n-1}{r}\binom{n-2}{r-2}\int_0^\infty [P\{|X| \geq x\}]^{r-1} \\
&\qquad \times P\Big\{\max_{i \leq n-r} |X_i| \leq \varepsilon x \text{ and } 0 \leq S_{n-r} \leq Tx\Big\}\, dh(x) \\
&\quad \geq C_1 n^{r-1}\int_0^\infty [P\{|X| \geq x\}]^{r-1} \\
&\qquad \times P\Big\{\max_{i \leq n-r} |X_i| \leq \varepsilon x \text{ and } 0 \leq S_{n-r} \leq Tx\Big\}\, dh(x),
\end{aligned}$$

where $h(x) = P\{|X| \leq x\}$. The last estimate remains valid for $r = 1$.

At this stage we use the fact that

$$\liminf_{n\to\infty} \frac{{}^{(r)}\tilde{S}_n}{|X_n^{(r)}|} > 0 \quad \text{a.s.}$$

implies, for some $T > 0$,

$$P\Big\{{}^{(r)}\tilde{S}_n \geq T|X_n^{(r)}|\Big\} \to 1.$$

Then by Theorem 2.1 (see also Lemma 3.2), (4.1) must hold and this implies

$$\frac{S_n}{B(n)} \to_P 1 \tag{4.16}$$

with $B(n)$ satisfying (4.5). That this choice of norming constants is permitted was proven by Rogozin (1976), or it can be obtained by an easy Chebyshev estimate, truncating X_i at $B(n)$ and using

$$nP\{|X| > B(n)\} = \frac{B(n)}{A(B(n))}P\{|X| > B(n)\} \to 0.$$

We can choose an x_0 large enough for $A(x) > 0$ when $x \geq x_0$, and, as we saw in the proof of Theorem 2.1, (4.16) also implies $X_n^{(1)}/B(n) \to_P 0$. These, together with (4.16), show that if x_0 is large enough, $x \geq x_0$ and

$$B(n) < \frac{Tx}{2}, \tag{4.17}$$

then

$$\begin{aligned} &P\{|X_i| \leq \varepsilon x \text{ for } i \leq n \text{ and } 0 \leq S_n \leq Tx\} \\ &\quad \geq P\{0 \leq S_n \leq 2B(n)\} - P\left\{|X_n^{(1)}| > \frac{\varepsilon B(n)}{T}\right\} \\ &\quad \geq \frac{1}{2}. \end{aligned}$$

In particular, this holds for all

$$n < \frac{Tx}{4A(x/4)}. \tag{4.18}$$

This can be seen from the definition of $B(n)$, since for such n,

$$\frac{Tx}{2A(Tx/2)} \sim \frac{Tx}{2A(x/4)} > 2n$$

because A is slowly varying under (4.1). It follows that, for some constants $C_i = C_i(\varepsilon) > 0$,

$$\begin{aligned} &\sum_1^\infty P\left\{\Delta(n,r,T,\varepsilon) \cap \left({}^{(r)}\tilde{S}_n \geq 0\right)\right\} \\ &\quad \geq C_2 \int_{[x_0,\infty)} [P\{|X| \geq x\}]^{r-1} \sum_{n < Tx/(4A(x/4))} n^{r-1}\, dh(x) \\ &\quad \geq C_3 \int_{[x_0,\infty)} [P\{|X| \geq x\}]^{r-1} \left[\frac{x}{A(x)}\right]^r dh(x). \end{aligned}$$

In view of (4.15) this implies

$$\int_{[x_0,\infty)} [P\{|X| \geq x\}]^{r-1} \left[\frac{x}{A(x)}\right]^r dh(x) < \infty. \tag{4.19}$$

To connect this with (2.13), we require a formula for the differential $dh^r(x)$, which can be obtained as follows. If $r = 1, 2, \dots,$ let $m_r = \min_{1 \leq i \leq r} Y_i$, where $Y_1, Y_2, \dots, Y_r$ are independent and identically distributed random variables with distribution G. Interpret $\min_{r+1\le i\le r} Y_i$ and $\min_{1\le i\le 0} Y_i$ as ∞. We can calculate, for any x,

$$
\begin{aligned}
P\{m_r > x\} &= \sum_{j=1}^{r} P\{m_r > x, \text{exactly } j \text{ of } Y_i = m_r\}\\
&= \sum_{j=1}^{r} \binom{r}{j} P\Big\{x < Y_1 = \cdots = Y_j < \min_{j+1\le i\le r} Y_i\Big\}\\
&= \sum_{j=1}^{r} \binom{r}{j} \int_{(x,\infty)} P\{x < Y_1 = \cdots = Y_j < y\}\, dP\Big\{\min_{1\le i\le r-j} Y_i \le y\Big\}\\
&= \int_{(x,\infty)} \sum_{j=1}^{r} \binom{r}{j} \Big\{\int_{(x,y)} [\Delta G(z)]^{j-1}\, dG(z)\Big\}\, dP\{m_{r-j} \le y\},
\end{aligned}
\tag{4.20}
$$

where

$$\Delta G(z) = G(z) - G(z-) = P\{Y = z\}.$$

Define also $\overline{G}(x) = 1 - G(x)$, so that

$$P\{m_{r-j} > z\} = \big[\overline{G}(z)\big]^{r-j}.$$

Then, using Fubini's theorem to interchange integrations, we have from (4.20) that

$$
\begin{aligned}
P\{m_r > x\} &= \int_{(x,\infty)} \sum_{j=1}^{r} \binom{r}{j} [\Delta G(z)]^{j-1} \big[\overline{G}(z)\big]^{r-j}\, dG(z)\\
&= \int_{(x,\infty)} g_r(z)\, dG(z), \quad \text{say.}
\end{aligned}
\tag{4.21}
$$

Now, if $\Delta G(z) > 0$,

$$
\begin{aligned}
g_r(z) &= \frac{\sum_{j=0}^{r} \binom{r}{j} [\Delta G(z)]^{j} \big[\overline{G}(z)\big]^{r-j} - \big[\overline{G}(z)\big]^{r}}{\Delta G(z)}\\
&= \frac{\big[\Delta G(z) + \overline{G}(z)\big]^{r} - \big[\overline{G}(z)\big]^{r}}{\Delta G(z)}\\
&= \frac{\big[\overline{G}(z-)\big]^{r} - \big[\overline{G}(z)\big]^{r}}{\Delta G(z)}.
\end{aligned}
$$

Thus

$$g_r(z) = \sum_{j=0}^{r-1} \big[\overline{G}(z-)\big]^{j} \big[\overline{G}(z)\big]^{r-1-j},$$

and, from (4.21), the last expression is still correct if $\Delta G(z) = 0$, since then

$g_r(z) = r[\overline{G}(z)]^{r-1}$. Thus we have shown that

$$P\{m_r > x\} = \left[\overline{G}(x)\right]^r = \int_{(x,\infty)} \sum_{j=0}^{r-1} \left[\overline{G}(z-)\right]^j \left[\overline{G}(z)\right]^{r-1-j} dG(z). \tag{4.22}$$

In other words, a version of the Radon–Nikodym derivative of $[\overline{G}(x)]^r$ is

$$-\frac{d\left[\overline{G}(x)\right]^r}{dG(x)} = \frac{d\left[\overline{G}(x)\right]^r}{d\overline{G}(x)} = \sum_{j=0}^{r-1} \left[\overline{G}(z-)\right]^j \left[\overline{G}(z)\right]^{r-1-j}. \tag{4.23}$$

[See also Goldie and Maller (1992) for a related discussion.]

Returning now to the proof of (2.13), we see that (4.23) with Y_i replaced by $|X_i|$ implies

$$\begin{aligned} -d\left[P\{|X| > x\}\right]^r &= \sum_0^{r-1} P^j\{|X| \geq x\} P^{r-1-j}\{|X| > x\}(-dP\{|X| > x\}) \\ &\leq rP^{r-1}\{|X| \geq x\}(-dP\{|X| > x\}) \\ &= rP^{r-1}\{|X| \geq x\}\, dh(x). \end{aligned} \tag{4.24}$$

Then (4.19), together with (4.24), shows that

$$-\int_{[x_0,\infty)} \left[\frac{x}{A(x)}\right]^r d\left[P\{|X| > x\}\right]^r < \infty;$$

after integrating by parts this gives

$$\int_{[x_0,\infty)} \left[P\{|X| > x\}\right]^r \left[\frac{x}{A(x)}\right]^{r-1} d\left[\frac{x}{A(x)}\right] < \infty. \tag{4.25}$$

[Note that $x/A(x)$ is continuous so that

$$d\left[\frac{x}{A(x)}\right]^r = r\left[\frac{x}{A(x)}\right]^{r-1} d\left[\frac{x}{A(x)}\right]$$

does not need the elaborate justification of (4.24).] Finally, (4.3) shows that (4.25) implies (2.13).

Finally, we show that (2.13) holds with $r = 1$ if and only if

$$0 < EX \leq E|X| < \infty.$$

If (2.13) holds, then since, for $x \geq x_0$,

$$\begin{aligned} 0 < A(x) &= \int_0^x [1 - F(y) - F(-y)]\, dy \leq \int_0^x [1 - F(y) - F(-y-)]\, dy \\ &\leq \int_0^x [1 - F(y) + F(-y-)]\, dy = \int_0^x H(y)\, dy \end{aligned}$$

[where $H(x) = P\{|X| > x\}$], we have

$$\int_{x_0}^{\infty} \frac{H(y)}{\int_0^x H(y)\, dy}\, dx < \infty,$$

and, by the Abel–Dini theorem, this is only possible if

$$E|X| = \int_0^\infty H(y)\,dy < \infty.$$

Now $EX = \lim_{x\to\infty} A(x) \ge 0$, but $EX = 0$ is impossible. This would imply $A(x) \to 0$ as $x \to \infty$, and thus the integral $\int_0^\infty [1 - F(y) - F(-y)]\,dy$ would converge (absolutely, since $E|X| < \infty$) and equal 0. However, then

$$0 < A(x) = -\int_x^\infty [1 - F(y) - F(-y)]\,dy \le \int_x^\infty H(y)\,dy$$

would give

$$\int_{x_0}^T \frac{H(x)}{A(x)}\,dx \ge \int_{x_0}^T \frac{H(x)}{\int_x^\infty H(y)\,dy}\,dx$$

$$= -\log\left\{\frac{\int_T^\infty H(y)\,dy}{\int_{x_0}^\infty H(y)\,dy}\right\} \to \infty, \qquad T \to \infty.$$

This contradicts (2.13) when $r = 1$. Thus $0 < EX \le E|X| < \infty$ as claimed. Conversely, $0 < EX \le E|X| < \infty$ implies $A(x) \to EX > 0$, so

$$\int_{x_0}^\infty \frac{H(x)}{A(x)}\,dx \asymp \int_{x_0}^\infty H(y)\,dy \asymp E|X| < \infty$$

and (2.13) holds. This completes the proof of Theorem 2.2. □

5. Proof of Theorem 2.3. When $X_i \ge 0$ a.s., Theorem 2.3 follows from Theorem 2.2, and for the general case we show that the contribution of the negative tail is negligible with respect to that of the positive, as outlined in Section 2, when (2.18b) holds. To do this, recall that $X_i^+ = \max(X_i, 0)$ and $X_i^- = X_i^+ - X_i$, and let $(X^+)_n^{(n)} \le \cdots \le (X^+)_n^{(1)}$ denote $X_1^+, \dots, X_n^+$, arranged in increasing order, with a convention for breaking ties. Since $F(0) < 1$, $P(M_n^{(r)} \le 0 \text{ i.o.}) = 0$, so

$$P\big(M_n^{(j)} \ne (X^+)_n^{(j)} \text{ i.o.}\big) = 0 \quad \text{for each fixed } j. \tag{5.1}$$

Let (2.16) hold. This implies (2.17), which implies

$$\liminf_{n\to\infty} \frac{\sum_{i=1}^n X_i^+ - \sum_{i=1}^n X_i^- - (X^+)_n^{(1)} - \cdots - (X^+)_n^{(r)}}{(X^+)_n^{(r)}} > 0 \quad \text{a.s.} \tag{5.2}$$

A fortiori,

$$\liminf_{n\to\infty} \frac{\sum_{i=1}^n X_i^+ - (X^+)_n^{(1)} - \cdots - (X^+)_n^{(r)}}{(X^+)_n^{(r)}} > 0 \quad \text{a.s.,} \tag{5.3}$$

from which (2.18a) follows by applying Theorem 2.2 to X_i^+. Also, we see from (5.2) that

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^n X_i^-}{\sum_{i=1}^n X_i^+} \le 1 \quad \text{a.s.} \tag{5.4}$$

From this (2.18b) follows by Erickson [(1973), Lemma 3] or Pruitt [(1981), Lemma 8.1].

To prove that (2.18) implies (2.19), the following lemma is useful.

LEMMA 5.1. *Suppose* $E|X| = \infty$. *For* $r = 1, 2, \dots,$ (2.18b) *is equivalent to*

$$\frac{\sum_{i=1}^n X_i^-}{\sum_{i=1}^n X_i^+ - M_n^{(1)} - \cdots - M_n^{(r-1)}} \to 0 \quad a.s. \tag{5.5}$$

PROOF. For $r = 1$, (5.5) is to be interpreted as

$$\frac{\sum_{i=1}^n X_i^-}{\sum_{i=1}^n X_i^+} \to 0 \quad \text{a.s.};$$

this case was proved in Pruitt [(1981), Lemma 8.1], so we assume it here. Thus, defining

$$\beta(x) = \frac{x}{\int_0^x [1 - F(y)]\, dy} = \frac{1}{\int_0^1 [1 - F(xy)]\, dy},$$

the case $r = 1$ tells us that $E(\beta(X^-)) < \infty$. Also (5.5) for some r implies (5.5) for $r = 1$ and hence (2.18b). Next assume (2.18b) or, equivalently, $E(\beta(X^-)) < \infty$. It is easy to check that β is subadditive, so for any $r \ge 1$,

$$E(\beta(X_1^- + \cdots + X_r^-)) \le rE(\beta(X^-)) < \infty.$$

Now take $r = 2, 3, \dots$ and fix $j \in [0, r-1]$. Define independent and identically distributed random variables Y_k, $k \ge 1$, by

$$Y_k^- = X_{kr+j}^- + X_{kr+j+1}^- + \cdots + X_{(k+1)r+j-1}^-,$$

$$Y_k^+ = \begin{cases} X_{kr+j}^+, & \text{if } X_{kr+j}^- = X_{kr+j+1}^- = \cdots = X_{(k+1)r+j-1}^- = 0, \\ 0, & \text{otherwise}. \end{cases}$$

Then for $x > 0$,

$$P(Y_k^- > x) = P\left\{\sum_{m=0}^{r-1} X_{kr+j+m}^- > x\right\} = P\left\{\sum_{i=1}^{r} X_i^- > x\right\},$$

and

$$\begin{aligned} P(Y_k > x) &= P(Y_k^+ > x) = P(X_{kr+j}^+ > x,\, X_{kr+j+m}^- = 0,\, 0 \le m \le r-1) \\ &= [1 - F(x)]P^{r-1}(X \ge 0). \end{aligned}$$

By the Pruitt–Erickson result, $\Sigma_{k=1}^n Y_k^- / \Sigma_{k=1}^n Y_k^+ \to 0$ a.s. provided

$$\int_0^\infty \left\{ \frac{x}{\int_0^x P(Y > y)\, dy} \right\} |dP(Y^- > x)| < \infty,$$

and this is true if

$$\int_0^\infty \left\{ \frac{x}{\int_0^x [1 - F(y)]\, dy} \right\} \left| dP\left\{ \sum_{i=1}^r X_i^- > x \right\} \right| = \int_0^\infty \beta(x)\, dP\left\{ \sum_{i=1}^r X_i^- \le x \right\}$$
$$= E\left\{ \beta\left(\sum_{i=1}^r X_i^- \right) \right\}$$

is finite. However, we showed at the beginning of the proof that this is so. So indeed we have $\Sigma_{k=1}^n Y_k^- / \Sigma_{k=1}^n Y_k^+ \to 0$ a.s. and this implies, for each $0 \le j \le r - 1$,

$$\sum_{i=1}^n X_i^- = o\left\{ \sum_{ir+j \le n} X_{ir+j}^+ \right\} \quad \text{a.s.} \tag{5.6}$$

Now $M_n^{(1)}, \dots, M_n^{(r-1)}$ can occur at times lying in at most $r - 1$ residue classes, mod r. Thus one sequence X_{ir+j}^+, $i = 1, 2, \dots$, does not contain any of $M_n^{(1)}, \dots, M_n^{(r-1)}$, so (5.5) follows from (5.6). □

Let (2.18) hold and suppose first that $E|X| = \infty$. (2.18a) implies, by (5.1) and Theorem 2.2 applied to X_i^+, that

$$\frac{\Sigma_{i=1}^n X_i^+ - M_n^{(1)} - \cdots - M_n^{(r-1)}}{B_n} \to 1 \quad \text{a.s.}$$

for some increasing B_n. (2.18b) implies that $\Sigma_{i=1}^n X_i^-$ is, by Lemma 5.1, of smaller order than the numerator here, so $\Sigma_{i=1}^n X_i^- / B_n \to 0$ a.s. Consequently, (2.19) holds and this of course implies (2.20).

The proof of (2.16) from (2.20) is the same as that of (2.11) from (2.15).

We next consider the case when $E|X| < \infty$. We now replace (2.18a) and (2.18b) by $E(X^-) < E(X^+) < \infty$. If the revised (2.18a) and (2.18b) hold, then clearly (2.16)–(2.20) hold for any $r \ge 1$. It remains to show that any of (2.16), (2.17), (2.19) or (2.20) for some r imply $E(X^-) < E(X^+)$, knowing that $E|X| < \infty$. Now (2.16) obviously implies (2.17), which implies (5.4). Thus $E(X^-) \le E(X^+) < \infty$. However, if $E(X^-) = E(X^+)$, then $E(X) = 0$ and S_n is recurrent, giving $P(S_n \le 0 \text{ i.o.}) = 1$. Since ${}^{(r)}S_n = S_n - M_n^{(1)} - \cdots - M_n^{(r)} \le S_n$ eventually, a.s., we also have $P({}^{(r)}S_n / M_n^{(r)} \le 0 \text{ i.o.}) = 1$, contradicting (2.17). Thus (2.16) or (2.17) imply $E(X^-) < E(X^+) < \infty$. Next, (2.19) implies (2.20), and when $E|X| < \infty$, (2.20) implies $E(X) > 0$, again because $E(X) = 0$ would make S_n recurrent.

Next we show that any one of (2.16)–(2.20) for $r = 1$ implies $E|X| < \infty$. Again (2.16) implies (2.17), which implies (5.3) with $r = 1$. This implies $EX^+ < \infty$, by Kesten (1971). Alternatively, by Theorem 2.2 applied to X_i^+, (5.3) implies (2.18a) for $r = 1$. But (2.18a) with $r = 1$ is equivalent to $EX^+ < \infty$ by the Abel–Dini theorem [or see Erickson (1973)]. Also, (2.17) implies (5.4) and hence $EX^- \le EX^+ < \infty$. As just stated, (2.18a) for $r = 1$ implies $EX^+ < \infty$ and hence $\lim_{x \to \infty} A_+(x) < \infty$. Then (2.18b) shows again that $EX^- < \infty$. Finally, (2.19) implies (2.20), clearly. If (2.20) holds with $r = 1$, then $S_n/B_n \asymp 1$ a.s. and by Chow and Robbins (1961), Kesten (1971), or Maller (1978), we have $E|X| < \infty$. This completes the proof of Theorem 2.3. □

Acknowledgment. The first author is grateful for hospitality and financial assistance from the Department of Statistics, University of Melbourne and from the University of Western Australia.

REFERENCES

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge Univ. Press.

BREIMAN, L. (1965). On some limit theorems similar to the arc-sine law. *Theory Probab. Appl.* **10** 323–331.

CHOW, Y. S. and ROBBINS, H. (1961). On sums of independent random variables with infinite moments and 'fair games'. *Proc. Nat. Acad. Sci. U.S.A.* **47** 330–335.

DURRETT, R., KESTEN, H. and LAWLER, G. (1991). Making money from fair games. In *Random Walks, Brownian Motion and Interacting Particle Systems* (R. Durrett and H. Kesten, eds.) 255–267. Birkhäuser, Boston.

ERICKSON, K. B. (1973). On the strong law of large numbers when the mean is undefined. *Trans. Amer. Math. Soc.* **185** 371–381.

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

GNEDENKO, B. V. and KOLMOGOROV, A. N. (1968). *Limit Theorems for Sums of Independent Random Variables*, 2nd ed. Addison-Wesley, Reading, Mass.

GOLDIE, C. M. and MALLER, R. A. (1992). Almost sure behaviour of record values and order statistics. Unpublished manuscript.

KESTEN, H. (1971). Sums of random variables with infinite expectation. *Amer. Math. Monthly* **78** 305–308.

KESTEN, H. (1972). Sums of independent random variables—without moment conditions. *Ann. Math. Statist.* **43** 701–732.

KESTEN, H. and LAWLER, G. F. (1992). A necessary condition for making money from fair games. *Ann. Probab.* **20** 855–882.

MALLER, R. A. (1978). Relative stability and the strong law of large numbers. *Z. Wahrsch. Verw. Gebiete* **43** 141–148.

MALLER, R. A. (1979). Relative stability, characteristic functions and stochastic compactness. *J. Austral. Math. Soc. Ser. A* **28** 499–509.

MALLER, R. A. (1982). Asymptotic normality of lightly trimmed means—A converse. *Math. Proc. Cambridge Philos. Soc.* **92** 535–545.

MALLER, R. A. (1984). Relative stability of trimmed sums. *Z. Wahrsch. Verw. Gebiete* **66** 61–80.

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

MORI, T. (1976). The strong law of large numbers when extreme terms are removed from sums. *Z. Wahrsch. Verw. Gebiete* **36** 189–193.

MORI, T. (1977). Stability of sums of iid random variables when extremes are excluded. *Z. Wahrsch. Verw. Gebiete* **40** 159–167.

MORI, T. (1984). On the limit distributions of lightly trimmed sums. *Math. Proc. Cambridge Philos. Soc.* **96** 507–516.

O'Brien, G. L. (1980). A limit theorem for sample maxima and heavy branches in Galton–Watson trees. *J. Appl. Probab.* **17** 539–545.

Pruitt, W. E. (1981). General one-sided laws of the iterated logarithm. *Ann. Probab.* **9** 1–48.

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

Rogozin, B. A. (1976). Relatively stable walks. *Theory Probab. Appl.* **21** 375–379.

Department of Mathematics
Cornell University
Ithaca, New York 14853-7901

Department of Mathematics
University of Western Australia
Nedlands 6009
Western Australia