The Annals of Probability
1989, Vol. 17, No. 3, 1186–1219

ASYMPTOTIC NORMALITY AND SUBSEQUENTIAL LIMITS OF TRIMMED SUMS

By Philip S. Griffin[1] and William E. Pruitt[2]

Syracuse University and University of Minnesota

Let $\{X_i\}$ be i.i.d. and $S_n(s_n, r_n)$ the sum of the first n X_i with the r_n largest and s_n smallest excluded. Assume $r_n \to \infty$, $s_n \to \infty$, $n^{-1}r_n \to 0$, $n^{-1}s_n \to 0$. Necessary and sufficient conditions are obtained for the existence of $\{\delta_n\}, \{\gamma_n\}$ such that $\gamma_n^{-1}(S_n(s_n, r_n) - \delta_n)$ converges weakly to a standard normal. The set of all subsequential limit laws for these sequences is characterized and sufficient conditions are given for X_i to be in the domain of partial attraction of a given law in the class. These conditions are also necessary if a unique factorization result for characteristic functions is true.

1. Introduction. Let $X, X_1, X_2, \dots$ be a sequence of nondegenerate i.i.d. random variables and let X_{nk} denote the kth smallest of $\{X_1, \dots, X_n\}$; thus

$$X_{n1} \le X_{n2} \le \cdots \le X_{nn}.$$

The trimmed sums we will deal with are defined by

$$S_n(s_n, r_n) = \sum_{k=s_n+1}^{n-r_n} X_{nk} = S_n - \sum_{k=1}^{s_n} X_{nk} - \sum_{k=n-r_n+1}^{n} X_{nk},$$

the usual sum but with the s_n smallest and r_n largest summands discarded. The principal results of this article are: (i) a necessary and sufficient condition for asymptotic normality of $S_n(s_n, r_n)$ and (ii) a complete description of the class of all possible subsequential limit laws for $S_n(s_n, r_n)$. We also give sufficient conditions for convergence of a particular subsequence to a given limit which will be necessary as well if a result on unique factorization of characteristic functions can be proved.

In order to state the results we need a little notation. We will assume that

$$r_n \to \infty, \qquad s_n \to \infty, \qquad n^{-1}r_n \to 0, \qquad n^{-1}s_n \to 0. \tag{1.1}$$

For $\alpha, \beta \in \mathbb{R}$ define $\{a_n(\alpha)\}, \{b_n(\beta)\}$ by

$$a_n(\alpha) = \inf\{x: P\{X \le -x\} \le n^{-1}(s_n - \alpha s_n^{1/2})\}, \tag{1.2}$$

$$b_n(\beta) = \inf\{x: P\{X > x\} < n^{-1}(r_n - \beta r_n^{1/2})\}; \tag{1.3}$$

for each fixed α, β these will be defined for n sufficiently large. Let

$$\lambda_n(\alpha, \beta) = E((X \wedge b_n(\beta)) \vee (-a_n(\alpha)))^2,$$

Received July 1986; revised January 1988.

[1]This work was partially supported by NSF Grant DMS-85-01732 and a grant from Syracuse University.

[2]This work was partially supported by NSF Grant DMS-83-01080.

AMS 1980 *subject classification.* 60F05.

Key words and phrases. Asymptotic normality, stochastic compactness, discarding outliers.

where $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$. We can now state the results. The necessary and sufficient condition for asymptotic normality of $S_n(s_n, r_n)$ is that for all $\alpha, \beta \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{\lambda_n(\alpha,\beta)}{\lambda_n(0,0)} = 1. \tag{1.4}$$

This condition holds for all $\{r_n\}, \{s_n\}$ satisfying (1.1) provided the tails of the distribution of X do not have long flat stretches where their relative decay is slower than any power. Even so, the question of whether it fails can be quite delicate. An example is given at the end of Section 4 which shows that it is possible to have $P\{X > x\} \sim P\{Y > x\}$ and $P\{X \le -x\} \sim P\{Y \le -x\}$ as $x \to \infty$ and yet have asymptotic normality for the trimmed sum for the X's but not for the Y's.

The class of subsequential limits is large even though asymptotic normality is so prevalent. The members of this class are of the form

$$\tau N_1 + f(N_2) - g(N_3), \tag{1.5}$$

where N_1, N_2, N_3 are independent $N(0,1)$, $\tau \ge 0$ and f and g are arbitrary nondecreasing convex functions. We will also show that for any given $\{r_n\}, \{s_n\}$ there is a universal law for X such that all limits in (1.5) arise by taking different subsequences of $S_n(s_n, r_n)$. It is natural to ask whether all of these limit laws are infinitely divisible. This question was answered for us (negatively) by Fred Steutel. For example, if one takes $f(x) = (x^+)^\rho$ for $\rho \in (1,2)$, then it follows easily from Steutel [16] that $f(N_2)$ cannot be infinitely divisible.

For an example, consider the family of distributions with slowly varying tails given by

$$P\{X < -x\} = P\{X > x\} = \tfrac{1}{2}(\log x)^{-\rho}, \qquad x \ge e,$$

where $\rho > 0$. Then $S_n(s_n, r_n)$ is asymptotically normal iff

$$\lim_{n\to\infty} r_n^{-1} n^{2/(2+\rho)} = 0 \quad \text{and} \quad \lim_{n\to\infty} s_n^{-1} n^{2/(2+\rho)} = 0.$$

If one takes $r_n \wedge s_n$ proportional to $n^{2/(2+\rho)}$, then one obtains subsequential limits as in (1.5) with $\tau = 0$ and f, g or both exponential. Finally, if

$$\liminf_{n\to\infty} (r_n \wedge s_n) n^{-2/(2+\rho)} = 0,$$

then it is impossible to normalize $S_n(s_n, r_n)$ to even make it stochastically compact, that is, tight with no degenerate subsequential limits. A more complete description of this example is given in Section 6.

It is natural to ask whether the class of limit laws along the entire sequence is smaller than the class of subsequential limit laws. The answer depends on the phrasing of the question. For a given limit law W of the form (1.5) there exist sequences $\{r_n\}, \{s_n\}$, and a distribution for X such that the normalized trimmed sum converges to W. On the other hand, if $\{r_n\}$ and $\{s_n\}$ are given satisfying (1.1), then the class of limit laws for the normalized trimmed sums may be smaller than (1.5). It would be interesting to know what the class of limit laws for the entire sequence is under the natural restriction that $\{r_n\}, \{s_n\}$ are nondecreasing.

Stigler [17] found the limit law for $S_n(s_n, r_n)$ when $r_n = [pn]$, $s_n = [qn]$ with $p > 0$, $q > 0$, $p + q < 1$. Generally it is normal but flat stretches in the distribution function of X can lead to nonnormal limits. Maller [12] and Mori [14] showed that when r is fixed and the r terms largest in absolute value are trimmed, the normalized trimmed sum cannot have a normal limit unless X is in the domain of attraction of the normal. For X in the domain of attraction of a stable law, the exact limit distribution is obtained by Arov and Bobrov [1] for this problem and by Csörgő, Csörgő, Horváth and Mason [3] if the r largest and s smallest terms are trimmed where r and s are fixed.

Csörgő, Horváth and Mason [5] have shown that $S_n(s_n, r_n)$ is asymptotically normal when $r_n = s_n$, (1.1) holds and X is in the domain of attraction of a stable law. This also follows easily from (1.4) as we will show in Section 4.

In [9] we have considered the central limit problem for the trimmed sum obtained by deleting the r_n summands which are largest in absolute value. We denote this sum by ${}^{(r_n)}S_n$. This seems to be a more difficult problem and we have necessary and sufficient conditions for asymptotic normality only in the case of a symmetric distribution for X. For a continuous distribution (the condition is more complicated in general), one defines $c_n(\alpha)$ by

$$P\{|X| > c_n(\alpha)\} = n^{-1}\left(r_n - \alpha r_n^{1/2}\right) \tag{1.6}$$

and the condition becomes

$$EX^2 1\{|X| \le c_n(\alpha)\} \sim EX^2 1\{|X| \le c_n(0)\} \quad \text{for all } \alpha \in \mathbb{R}. \tag{1.7}$$

If we take a symmetric distribution for X, then if $S_n(r_n, r_n)$ is asymptotically normal so is ${}^{(2r_n)}S_n$. An interesting feature of this is that the asymptotic variance is always smaller for ${}^{(2r_n)}S_n$ than for $S_n(r_n, r_n)$—even much smaller when the tail of the X distribution is slowly varying. This is proved in Section 6. On the other hand, the converse does not hold: ${}^{(r_n)}S_n$ is asymptotically normal for any symmetric distribution for X for which S_n can be normalized to be stochastically compact whereas asymptotic normality of $S_n(s_n, r_n)$ may fail within this class even when $s_n = r_n$. An example is given in Section 6. For asymmetric distributions, asymptotic normality may hold for $S_n(r_n, r_n)$ while failing for ${}^{(2r_n)}S_n$. An example is given in [9]. The class of subsequential limits analogous to (1.5) for ${}^{(r_n)}S_n$ is $h(N_1)N_2 + \mu$, where $\mu \in \mathbb{R}$ and h is nonnegative, nondecreasing with h^2 convex.

Some of the results in this article have been obtained independently by Csörgő, Haeusler and Mason [4]; in particular, they have an equivalent form of (1.4). Our approaches to these problems are very different. They use a Brownian bridge approximation to the empirical distribution function, while our methods are more classical relying on the theorems of Berry-Esseen and Liapounov.

Our interest in trimmed sums was stimulated by the work of Mori [13] who showed that the domain of applicability of the strong law of large numbers is increased by trimming a fixed number of terms. There are many recent articles on the law of the iterated logarithm for trimmed sums which we have not listed.

In other recent work, Kuelbs and Ledoux [11] and Hahn and Kuelbs [10] obtain asymptotic normality for sums which are trimmed only when the summands exceed a level which is deterministic but varies with n. They show that, under their assumptions, the main effect stems from the deterministic level rather than from the trimming. This enables them to obtain asymptotic normality in general. If the levels are chosen so that both the trimming and the deterministic cutoff points play a significant role, this mixed or conditional trimming will presumably lead to an even wider class of subsequential limit laws than the one we obtain here. Weiner [18] studies censored sums where the r_n summands largest in absolute value are reduced to a deterministic level when they exceed it. Griffin [8] considers Winsorized means where the s_n smallest summands are each replaced by X_{n,s_n} and the r_n largest by $X_{n,n-r_n+1}$. Weiner [18] also has some remarks comparing the various methods.

We now give a brief description of the contents of the remaining sections. Some preliminaries are in Section 2 including necessary and sufficient conditions to be able to normalize $S_n(s_n, r_n)$ so as to make this sequence stochastically compact. The solution to the general limit problem is in Section 3. The criteria for asymptotic normality along with some related conditions are in Section 4. The construction of the universal law is in Section 5. Examples and some comparisons with the trimmed sums ${}^{(r_n)}S_n$ are in Section 6. A few comments about the statistical implications of these results are in Section 7.

2. Preliminaries and stochastic compactness. The key to our proofs is that by conditioning on X_{n,s_n} and $X_{n,n-r_n+1}$, the distribution of $S_n(s_n, r_n)$ is seen to be a mixture of distributions of sums of i.i.d. random variables whose distributions are very close to that of X, appropriately truncated [see (3.12)]. A discontinuous distribution for X complicates this slightly even though it does not cause any problem in defining the trimmed sums. In order to overcome this complication we introduce an i.i.d. uniform $(0,1)$ sequence $Y, Y_1, Y_2, \ldots$ and then define $X_i = F^{-1}(Y_i)$, where F^{-1} is the right-continuous inverse of the distribution function F of X, that is,

$$F^{-1}(y) = \inf\{x: F(x) > y\}, \qquad 0 < y < 1.$$

Then $\{X_i\}$ is an i.i.d. sequence with distribution F. This leads to a unique ordering of the X's defined by $X_{ni} = F^{-1}(Y_{ni})$, where $Y_{n1} < Y_{n2} < \cdots < Y_{nn}$ are the ordered values of $\{Y_1, \ldots, Y_n\}$. Thus $X_{n1} \le X_{n2} \le \cdots \le X_{nn}$. Of course, it suffices to prove the various results for the $\{X_i\}$ defined in this way.

Another approximation to the trimmed sum will also be used—see (2.11)—to obtain information about the location and dispersion of the distribution of $S_n(s_n, r_n)$ in Lemma 5. The earlier lemmas will describe the limiting behavior of the order statistics and sums of truncated variables. The assumption (1.1) will always be made. It will also be convenient to assume that

$$P\{X > 0\} > 0 \quad \text{and} \quad P\{X < 0\} > 0.$$

This is harmless since the translation $X \to X + c$ does not affect the ordering of the X_i's and so only adds $(n - r_n - s_n)c$ to $S_n(s_n, r_n)$. The centering can be adjusted to compensate for this. This assumption does not change the various conditions for convergence.

Define $v_n(\alpha) = n^{-1}(s_n - \alpha s_n^{1/2})$, $u_n(\beta) = 1 - n^{-1}(r_n - \beta r_n^{1/2})$ and note that if α, β are fixed, then by (1.1) $0 < v_n(\alpha) < u_n(\beta) < 1$ for large n and $v_n(\alpha) \to 0$, $u_n(\beta) \to 1$. Furthermore, with $a_n(\alpha)$, $b_n(\beta)$ as in (1.2), (1.3) we have $-a_n(\alpha) = F^{-1}(v_n(\alpha))$, $b_n(\beta) = F^{-1}(u_n(\beta))$ and

$$(2.1) \qquad \begin{aligned} P\{X < -a_n(\alpha)\} &\le n^{-1}(s_n - \alpha s_n^{1/2}) \le P\{X \le -a_n(\alpha)\}, \\ P\{X > b_n(\beta)\} &\le n^{-1}(r_n - \beta r_n^{1/2}) \le P\{X \ge b_n(\beta)\}. \end{aligned}$$

Also, the above assumption that there is mass on both sides of the origin means that when α, β are fixed $a_n(\alpha)$, $b_n(\beta)$ are positive for large n. We will sometimes assume without comment that n is large enough that this is true. The following notation will be used for $\alpha, \beta \in \mathbb{R}$:

$$(2.2) \qquad \begin{aligned} (\sigma_n^+(\beta))^2 = \operatorname{Var}(X^+ \wedge b_n(\beta)), &\qquad (\sigma_n^-(\alpha))^2 = \operatorname{Var}(X^- \wedge a_n(\alpha)), \\ \sigma_n^2(\alpha, \beta) &= \operatorname{Var}((X \wedge b_n(\beta)) \vee (-a_n(\alpha))) \\ &= \operatorname{Var}((X^+ \wedge b_n(\beta)) - (X^- \wedge a_n(\alpha))) \\ &= (\sigma_n^+(\beta))^2 + (\sigma_n^-(\alpha))^2 \\ &\quad + 2E(X^+ \wedge b_n(\beta))E(X^- \wedge a_n(\alpha)), \end{aligned}$$

$$(2.3) \qquad \begin{aligned} \lambda_n^+(\beta) = E(X^+ \wedge b_n(\beta))^2, &\qquad \lambda_n^-(\alpha) = E(X^- \wedge a_n(\alpha))^2, \\ \lambda_n(\alpha, \beta) = E((X \wedge b_n(\beta)) \vee &(-a_n(\alpha)))^2 = \lambda_n^+(\beta) + \lambda_n^-(\alpha). \end{aligned}$$

We will make use of the monotonicity of $\lambda_n^+(\cdot)$, $\lambda_n^-(\cdot)$ without further mention. Note that by (2.1)

$$(2.4) \qquad \begin{aligned} n\lambda_n^+(\beta) &\ge b_n^2(\beta) nP\{X \ge b_n(\beta)\} \\ &\ge b_n^2(\beta)(r_n - \beta r_n^{1/2}) \sim b_n^2(\beta) r_n, \\ n\lambda_n^-(\alpha) &\ge a_n^2(\alpha) nP\{X \le -a_n(\alpha)\} \\ &\ge a_n^2(\alpha)(s_n - \alpha s_n^{1/2}) \sim a_n^2(\alpha) s_n. \end{aligned}$$

We will use the fact that

$$(2.5) \qquad \begin{aligned} &E(X^+)^2 = \infty \quad \text{implies} \\ &(E(X^+ \wedge m))^2 = o\big(E(X^+ \wedge m)^2\big) \quad \text{as } m \to \infty; \end{aligned}$$

see, for example, page 11 of [15]. Thus

$$(2.6) \qquad E(X^+)^2 = \infty \quad \text{implies} \quad (\sigma_n^+(\beta))^2 \sim \lambda_n^+(\beta).$$

If, on the other hand, $E(X^+)^2 < \infty$, then $(\sigma_n^+(\beta))^2 \to \operatorname{Var}(X^+)$ and $\lambda_n^+(\beta) \to E(X^+)^2$ so that

$$(2.7) \qquad (\sigma_n^+(\beta))^2 \approx \lambda_n^+(\beta)$$

in any case. ($c_n \approx d_n$ will be used to mean that $c_n d_n^{-1}$ is bounded above and below by positive, finite constants for large n.) Furthermore, using (2.2), (2.5), (2.7) and (2.6), we see that $EX^2 = \infty$ implies

$$\sigma_n^2(\alpha, \beta) \sim (\sigma_n^+(\beta))^2 + (\sigma_n^-(\alpha))^2 \sim \lambda_n^+(\beta) + \lambda_n^-(\alpha) = \lambda_n(\alpha, \beta). \tag{2.8}$$

This is valid even if $E(X^+)^2$ or $E(X^-)^2$ is finite.

We will now prove some lemmas. The first three are asymptotic normality results for triangular arrays which follow easily from Liapounov's condition (see, e.g., page 312 of [2]). Define for $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} T_n^+(\beta) &= \sum_{i=1}^n X_i^+ \wedge b_n(\beta), \qquad T_n^-(\alpha) = \sum_{i=1}^n X_i^- \wedge a_n(\alpha), \\ T_n(\alpha, \beta) &= \sum_{i=1}^n (X_i \wedge b_n(\beta)) \vee (-a_n(\alpha)) = T_n^+(\beta) - T_n^-(\alpha). \end{aligned} \tag{2.9}$$

LEMMA 1. *For all $\alpha, \beta \in \mathbb{R}$, (T_n^{-*}, T_n^{+*}) is asymptotically normal with mean 0 and covariance matrix A, where*

$$T_n^{+*} = (T_n^+(\beta) - ET_n^+(\beta))(n^{1/2}\sigma_n^+(\beta))^{-1},$$

$$T_n^{-*} = (T_n^-(\alpha) - ET_n^-(\alpha))(n^{1/2}\sigma_n^-(\alpha))^{-1},$$

and $A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ *with* $\rho = 0$ *if* $EX^2 = \infty$ *or*

$$\rho = -\frac{EX^+EX^-}{(\mathrm{Var}(X^+)\mathrm{Var}(X^-))^{1/2}} \quad \text{if } EX^2 < \infty.$$

PROOF. It suffices, by using characteristic functions, to show that $uT_n^{-*} + vT_n^{+*}$ is asymptotically normal with mean 0 and variance $u^2 + v^2 + 2uv\rho$. Observe that this is the sum of i.i.d. random variables. If $EX^2 = \infty$, the variance converges to $u^2 + v^2$ by using (2.5) as in (2.8) while if $EX^2 < \infty$, the covariance converges to ρ. It remains to check Liapounov's condition. We have by (2.7) and (2.4),

$$\begin{aligned} &nE\left|u(X^- \wedge a_n(\alpha))(n^{1/2}\sigma_n^-(\alpha))^{-1} + v(X^+ \wedge b_n(\beta))(n^{1/2}\sigma_n^+(\beta))^{-1}\right|^3 \\ &\quad \le n^{-1/2}|u|^3 a_n(\alpha)\lambda_n^-(\alpha)(\sigma_n^-(\alpha))^{-3} + n^{-1/2}|v|^3 b_n(\beta)\lambda_n^+(\beta)(\sigma_n^+(\beta))^{-3} \\ &\quad \approx n^{-1/2}|u|^3 a_n(\alpha)(\lambda_n^-(\alpha))^{-1/2} + n^{-1/2}|v|^3 b_n(\beta)(\lambda_n^+(\beta))^{-1/2} \\ &\quad = O(s_n^{-1/2} + r_n^{-1/2}) \to 0. \end{aligned} \tag{2.10}$$

□

Define for $\alpha, \beta \in \mathbb{R}$,

$$U_n^+(\beta) = \sum_{i=1}^n 1\{Y_i > u_n(\beta)\}, \qquad U_n^-(\alpha) = \sum_{i=1}^n 1\{Y_i < v_n(\alpha)\},$$

$$U_n^{+*}(\beta) = r_n^{-1/2}(U_n^+(\beta) - r_n + \beta r_n^{1/2}),$$

$$U_n^{-*}(\alpha) = s_n^{-1/2}(U_n^-(\alpha) - s_n + \alpha s_n^{1/2}).$$

The dependence on α, β will often be suppressed.

LEMMA 2. *For all $\alpha, \beta \in \mathbb{R}$, the pair (U_n^{-*}, U_n^{+*}) is asymptotically normal with mean 0 and identity covariance.*

PROOF. Note that

$$\mathrm{Var}(U_n^-(\alpha)) = nv_n(\alpha)(1 - v_n(\alpha)) \sim s_n.$$

Next

$$\begin{aligned}\mathrm{Cov}(1\{Y < v_n(\alpha)\}, 1\{Y > u_n(\beta)\}) &= -v_n(\alpha)(1 - u_n(\beta)) \\ &\sim -n^{-2}s_n r_n = o(n^{-1}s_n^{1/2}r_n^{1/2})\end{aligned}$$

and so $\mathrm{Var}(uU_n^{-*} + vU_n^{+*}) \to u^2 + v^2$. Liapounov is easy to check. □

LEMMA 3. *For all $\alpha, \beta \in \mathbb{R}$, the pair (U_n^{-*}, T_n^{+*}) is asymptotically normal with mean 0 and identity covariance. This also holds for (T_n^{-*}, U_n^{+*}).*

PROOF. Note that by (2.7)

$$\begin{aligned}\mathrm{Cov}(1\{Y < v_n(\alpha)\}, X^+ \wedge b_n(\beta)) &= -v_n(\alpha)E(X^+ \wedge b_n(\beta)) \\ &\sim -n^{-1}s_n E(X^+ \wedge b_n(\beta)) \\ &= O\big(n^{-1}s_n(\lambda_n^+(\beta))^{1/2}\big) \\ &= O\Big(n^{-1}(n^{-1}s_n)^{1/2}s_n^{1/2}(n^{1/2}\sigma_n^+(\beta))\Big) \\ &= o\big(n^{-1}s_n^{1/2}(n^{1/2}\sigma_n^+(\beta))\big).\end{aligned}$$

This means that $\mathrm{Var}(uU_n^{-*} + vT_n^{+*}) \to u^2 + v^2$. To check Liapounov,

$$\begin{aligned}&nE\Big|u1\{Y < v_n(\alpha)\}s_n^{-1/2} + v(X^+ \wedge b_n(\beta))\big(n^{1/2}\sigma_n^+(\beta)\big)^{-1}\Big|^3 \\ &\quad \le n|u|^3 v_n(\alpha)s_n^{-3/2} + n^{-1/2}|v|^3 b_n(\beta)\lambda_n^+(\beta)(\sigma_n^+(\beta))^{-3} \\ &\quad = O\big(s_n^{-1/2} + r_n^{-1/2}\big) \to 0\end{aligned}$$

as in (2.10). □

Next, we introduce for $\alpha < \alpha_1$, $\beta < \beta_1$,

$$N_n^+(\beta, \beta_1) = \sum_{i=1}^{n} 1\{u_n(\beta) < Y_i \le u_n(\beta_1)\},$$

$$N_n^-(\alpha, \alpha_1) = \sum_{i=1}^{n} 1\{v_n(\alpha_1) \le Y_i < v_n(\alpha)\}.$$

LEMMA 4. *For all $\alpha < \alpha_1$, $\beta < \beta_1$, $r_n^{-1/2}N_n^+(\beta, \beta_1) \to \beta_1 - \beta$ and $s_n^{-1/2}N_n^-(\alpha, \alpha_1) \to \alpha_1 - \alpha$ in probability.*

PROOF. This follows immediately from Chebyshev:

$$EN_n^+(\beta, \beta_1) = (\beta_1 - \beta) r_n^{1/2}, \qquad \operatorname{Var}(N_n^+(\beta, \beta_1)) \sim (\beta_1 - \beta) r_n^{1/2}. \qquad \square$$

Now we are almost ready to do some estimation of the trimmed sums. Observe that

$$S_n(s_n, r_n) - s_n a_n(\alpha) + r_n b_n(\beta) = T_n(\alpha, \beta) - V_n^-(\alpha) + V_n^+(\beta), \tag{2.11}$$

where

$$V_n^+(\beta) = \sum_{i=1}^{n-r_n} (X_{ni} - b_n(\beta)) 1\{b_n(\beta) < X_{ni}\} + \sum_{i=n-r_n+1}^{n} (b_n(\beta) - X_{ni}) 1\{X_{ni} < b_n(\beta)\},$$

$$V_n^-(\alpha) = \sum_{i=1}^{s_n} (X_{ni} + a_n(\alpha)) 1\{-a_n(\alpha) < X_{ni}\} + \sum_{i=s_n+1}^{n} (-a_n(\alpha) - X_{ni}) 1\{X_{ni} < -a_n(\alpha)\}.$$

To see the equality, think of $r_n b_n(\beta)$ as representing the r_n missing large terms and $s_n(-a_n(\alpha))$ the s_n missing small terms. Also note that

$$V_n^+(\beta) \ge 0, \qquad V_n^-(\alpha) \ge 0. \tag{2.12}$$

We will use this decomposition to obtain the following bounds on the location and spread of the distribution of $S_n(s_n, r_n)$.

LEMMA 5. *Let*

$$\delta_n(\alpha, \beta) = ET_n(\alpha, \beta) + s_n a_n(\alpha) - r_n b_n(\beta). \tag{2.13}$$

For every $\varepsilon > 0$, there exist ξ, M such that for all n sufficiently large

$$P\left\{ |S_n(s_n, r_n) - \delta_n(0,0)| \ge M(n\lambda_n(\xi, \xi))^{1/2} \right\} < \varepsilon. \tag{2.14}$$

Furthermore, if α, β satisfy

$$4(1 - \Phi(\alpha)) < 1, \qquad 4(1 - \Phi(\beta)) < 1, \tag{2.15}$$

where Φ is the $N(0,1)$ distribution function, and M is given, then there exists $\eta = \eta(\alpha, \beta, M) > 0$ such that for all n sufficiently large

$$P\left\{ S_n(s_n, r_n) - \delta_n(\alpha, \beta) \ge Mn^{1/2}\sigma_n^+(\beta) - 3\alpha a_n(\alpha) s_n^{1/2} \right\} \ge \eta, \tag{2.16}$$

$$P\left\{ S_n(s_n, r_n) - \delta_n(\alpha, \beta) \le -Mn^{1/2}\sigma_n^-(\alpha) + 3\beta b_n(\beta) r_n^{1/2} \right\} \ge \eta. \tag{2.17}$$

PROOF. Define

$$B_n = B_n(\xi) = \left\{ v_n(\xi) \le Y_{n, s_n} < v_n(-\xi) \right\}, \qquad C_n = \left\{ |U_n^-(0) - s_n| \le \xi s_n^{1/2} \right\}.$$

Then by Lemma 2, for $\xi > 0$, $P(C_n) \to 1 - 2\Phi(-\xi)$. Next, note that for all ζ,

$$\{Y_{n,s_n} < v_n(\zeta)\} = \{U_n^-(\zeta) \geq s_n\} = \{U_n^{-*}(\zeta) \geq \zeta\}, \tag{2.18}$$

so that by Lemma 2

$$P\{Y_{n,s_n} < v_n(\zeta)\} \to \Phi(-\zeta)$$

and so $P(B_n(\xi)) \to 1 - 2\Phi(-\xi)$. Thus for any $\varepsilon > 0$, by choosing ξ large enough, we have $P(B_nC_n) > 1 - \varepsilon/4$ for large n. But

$$\begin{aligned} B_n &\subset \{-a_n(\xi) \leq X_{n,s_n} \leq -a_n(-\xi)\} \\ &\subset \{V_n^-(0) \leq a_n(0)(s_n - U_n^-(0))^+ + a_n(\xi)(U_n^-(0) - s_n)^+\} \end{aligned}$$

so that

$$B_nC_n \subset \{V_n^-(0) \leq \xi s_n^{1/2} a_n(\xi)\}.$$

Thus we have by (2.12), (2.4) and (2.3)

$$P\{0 \leq V_n^-(0) \leq 2\xi(n\lambda_n(\xi,\xi))^{1/2}\} \geq 1 - \varepsilon/4. \tag{2.19}$$

By Lemma 1

$$P\{|T_n^+(0) - ET_n^+(0)| \geq \xi(n^{1/2}\sigma_n^+(0))\} \to 2\Phi(-\xi)$$

so that by (2.7) and (2.3)

$$P\{|T_n^+(0) - ET_n^+(0)| \geq \xi(n\lambda_n(0,0))^{1/2}\} < \varepsilon/4$$

for appropriate ξ and large n. Since the same methods work for $V_n^+(0)$ and $T_n^-(0)$, using these estimates in (2.11) yields (2.14). To prove (2.16), we note first that it suffices to prove it for large M. We introduce

$$\begin{aligned} A_n &= \{T_n^-(\alpha) \leq ET_n^-(\alpha),\ T_n^+(\beta) \geq ET_n^+(\beta) + Mn^{1/2}\sigma_n^+(\beta)\}, \\ D_n &= \{N_n^-(-\alpha,\alpha) < 3\alpha s_n^{1/2}\}. \end{aligned}$$

Now

$$A_n = \{T_n^{-*}(\alpha) \leq 0,\ T_n^{+*}(\beta) \geq M\}$$

so that by Lemma 1

$$\begin{aligned} \lim P(A_n) &= P\{W_1 \leq 0,\ W_2 \geq M\} \\ &= 1 - \Phi(M) - P\{W_1 > 0,\ W_2 \geq M\}, \end{aligned}$$

where $W = (W_1, W_2)$ is $N(0, A)$. A little algebra shows that since $\rho \leq 0$ the density for W is no larger than the standard normal density throughout the region $\{w\colon w_1 > 0,\ w_2 \geq M\}$ provided that $M^2 \geq -\rho^{-2}(1-\rho^2)\log(1-\rho^2)$. Thus we have

$$\liminf P(A_n) \geq \tfrac{1}{2}(1 - \Phi(M)).$$

By Lemma 3, with $B_n = B_n(\alpha)$,

$$\begin{aligned} P(A_nB_n^c) &\leq P\{T_n^{+*}(\beta) \geq M,\ U_n^{-*}(\alpha) \geq \alpha\} \\ &\quad + P\{T_n^{+*}(\beta) \geq M,\ U_n^{-*}(-\alpha) < -\alpha\} \\ &\to 2(1 - \Phi(M))(1 - \Phi(\alpha)). \end{aligned}$$

Since $P(D_n^c) \to 0$ by Lemma 4 this leads to

$$\liminf P(A_n B_n D_n) \geq \tfrac{1}{2}(1 - \Phi(M)) - 2(1 - \Phi(M))(1 - \Phi(\alpha)) > 0$$

by (2.15). Next, observe that

$$B_n(\alpha) \subset \{V_n^-(\alpha) \leq a_n(\alpha) N_n^-(-\alpha, \alpha)\}$$

and so by (2.11)

$$\begin{aligned} A_n B_n D_n \subset \big\{S_n(s_n, r_n) \geq ET_n^+(\beta) + Mn^{1/2}\sigma_n^+(\beta) - ET_n^-(\alpha) \\ -3\alpha s_n^{1/2} a_n(\alpha) + s_n a_n(\alpha) - r_n b_n(\beta)\big\} \\ \subset \big\{S_n(s_n, r_n) \geq \delta_n(\alpha, \beta) + Mn^{1/2}\sigma_n^+(\beta) - 3\alpha s_n^{1/2} a_n(\alpha)\big\}. \end{aligned}$$

This proves (2.16) and the proof of (2.17) is the same. □

Now we can obtain necessary and sufficient conditions for tightness, nondegenerate subsequential limits and stochastic compactness of the normalized trimmed sums.

PROPOSITION 1. *Assume* (1.1) *and let* $\{n_i\}$ *be any subsequence. There exists a sequence* $\{\delta_n\}$ *such that*

$$\gamma_{n_i}^{-1}\big(S_{n_i}(s_{n_i}, r_{n_i}) - \delta_{n_i}\big) \text{ is tight} \tag{2.20}$$

iff for all $\alpha, \beta \in \mathbb{R}$,

$$\limsup_{i \to \infty} \gamma_{n_i}^{-2} n_i \lambda_{n_i}(\alpha, \beta) < \infty. \tag{2.21}$$

There exists a sequence $\{\delta_n\}$ *such that all subsequential weak limits of*

$$\gamma_{n_i}^{-1}\big(S_{n_i}(s_{n_i}, r_{n_i}) - \delta_{n_i}\big) \text{ are nondegenerate} \tag{2.22}$$

iff there exist $\alpha, \beta \in \mathbb{R}$ *such that*

$$\liminf_{i \to \infty} \gamma_{n_i}^{-2} n_i \lambda_{n_i}(\alpha, \beta) > 0. \tag{2.23}$$

There exist sequences $\{\delta_n\}, \{\gamma_n\}$ *such that*

$$\gamma_{n_i}^{-1}\big(S_{n_i}(s_{n_i}, r_{n_i}) - \delta_{n_i}\big) \text{ is stochastically compact} \tag{2.24}$$

iff there exist $\alpha_0, \beta_0 \in \mathbb{R}$ *such that for all* $\alpha, \beta \in \mathbb{R}$,

$$\limsup_{i \to \infty} \frac{\lambda_{n_i}(\alpha, \beta)}{\lambda_{n_i}(\alpha_0, \beta_0)} < \infty. \tag{2.25}$$

In this case, one may take $\gamma_n = (n\lambda_n(\alpha_0, \beta_0))^{1/2}$ *and in* (2.20) *and* (2.24) *one may take* $\delta_n = \delta_n(0,0)$ *as defined in* (2.13). [*This sequence may not work in* (2.22), *however.*]

REMARK. To make (2.22) have content, we consider a subsequence along which there is convergence in probability to $+\infty$ (or to $-\infty$) to have a degenerate limit. This prevents one from taking δ_n too small (or large). If there is mass escaping to both $\pm\infty$, we consider the subsequential limit to be nondegenerate.

PROOF. The sufficiency of (2.21) follows immediately from Lemma 5. For the necessity, first observe that by (2.7) we can find a constant C such that for fixed α, β satisfying (2.15)

$$(\lambda_n(\alpha,\beta))^{1/2} \le C(\sigma_n^-(\alpha) + \sigma_n^+(\beta)), \tag{2.26}$$

for all sufficiently large n. Then take $M > C(1 + 6(\alpha \vee \beta))$ and apply (2.16), (2.17) and the tightness to obtain C_1 such that (for sufficiently large n in the subsequence)

$$\delta_n(\alpha,\beta) + Mn^{1/2}\sigma_n^+(\beta) - 3\alpha a_n(\alpha)s_n^{1/2} \le \delta_n + C_1\gamma_n,$$
$$\delta_n(\alpha,\beta) - Mn^{1/2}\sigma_n^-(\alpha) + 3\beta b_n(\beta)r_n^{1/2} \ge \delta_n - C_1\gamma_n.$$

Subtracting yields

$$\begin{aligned} 2C_1\gamma_n &\ge Mn^{1/2}(\sigma_n^+(\beta) + \sigma_n^-(\alpha)) - 3\alpha a_n(\alpha)s_n^{1/2} - 3\beta b_n(\beta)r_n^{1/2} \\ &\ge (n\lambda_n(\alpha,\beta))^{1/2} \end{aligned} \tag{2.27}$$

by (2.26), the choice of M, (2.4) and (2.3). Thus (2.21) holds for α, β satisfying (2.15). It is valid for smaller α, β by monotonicity.

For the necessity of (2.23) we will prove the contrapositive. Thus we assume for all $\alpha, \beta \in \mathbb{R}$ (along a subsequence)

$$\liminf_n \gamma_n^{-2} n\lambda_n(\alpha,\beta) = 0.$$

Then by the diagonal procedure we can find a subsequence such that

$$\lim_n \gamma_n^{-2} n\lambda_n(\alpha,\beta) = 0 \tag{2.28}$$

along this subsequence for all $\alpha, \beta \in \mathbb{R}$. Now

$$P\{|T_n(0,0) - ET_n(0,0)| \ge \varepsilon\gamma_n\} \le \varepsilon^{-2}\gamma_n^{-2} n\lambda_n(0,0) \to 0 \tag{2.29}$$

and by (2.19) and (2.28) we have with probability near 1,

$$0 \le V_n^-(0) \le 2\xi(n\lambda_n(\xi,\xi))^{1/2} = o(\gamma_n). \tag{2.30}$$

A similar argument applies to $V_n^+(0)$. Thus for any $\{\delta_n'\}$, with $\delta_n = \delta_n(0,0)$, we have

$$\gamma_n^{-1}(S_n(s_n, r_n) - \delta_n') = \gamma_n^{-1}(\delta_n - \delta_n') + \gamma_n^{-1}(S_n(s_n, r_n) - \delta_n)$$

and the second term on the right approaches 0 in probability by (2.11), (2.29) and (2.30). Taking a further subsequence to make the first term on the right converge (possibly to $+\infty$ or $-\infty$), we see that (2.22) fails. It remains to prove the sufficiency of (2.23). If α, β satisfy (2.15) and (2.23) and $M > C(1 + 6(\alpha \vee \beta))$, then (2.16) and (2.17) imply that, for the subsequence, $S_n(s_n, r_n)$ has mass bounded below on both sides of an interval of width

$$Mn^{1/2}(\sigma_n^+(\beta) + \sigma_n^-(\alpha)) - 3\alpha a_n(\alpha)s_n^{1/2} - 3\beta b_n(\beta)r_n^{1/2} \geq (n\lambda_n(\alpha,\beta))^{1/2} \geq c\gamma_n$$

by (2.27) and (2.23). Then take δ_n to be the left endpoint of the interval.

Finally, the equivalence of (2.24) and (2.25) follows easily from the first two parts except that we must check that in this case $\delta_n(0,0)$ will work in (2.22). We may increase α_0, β_0 to satisfy (2.15) and (2.25) still holds. Then a straightforward computation shows that

$$\delta_n(\alpha_0,\beta_0) - Mn^{1/2}\sigma_n^-(\alpha_0) + 3\beta_0 b_n(\beta_0)r_n^{1/2} - \delta_n(0,0) = O\big((n\lambda_n(\alpha_0,\beta_0))^{1/2}\big).$$

This means that the change in centering is of order γ_n in this case and thus the interval cannot slide off to $\pm\infty$. The problem in general with (2.22) is that γ_n may be much smaller than $(n\lambda_n(\alpha,\beta))^{1/2}$. □

3. The general limit theorem. We are now almost ready to state and prove the main limit theorems. But first we need a little more notation. Recall that

$$v_n(\alpha) = n^{-1}(s_n - \alpha s_n^{1/2}), \qquad u_n(\beta) = 1 - n^{-1}(r_n - \beta r_n^{1/2})$$

and assume that n is large enough that $0 < v_n(\alpha) < u_n(\beta) < 1$. Then define a random variable $X(n,\alpha,\beta) = F^{-1}(Y)1\{v_n(\alpha) < Y < u_n(\beta)\}$. The distribution function is given by

$$P\{X(n,\alpha,\beta) \leq x\} = \begin{cases} 0, & \text{if } x < -a_n(\alpha), \\ F(x) - v_n(\alpha), & \text{if } -a_n(\alpha) \leq x < 0, \\ F(x) + 1 - u_n(\beta), & \text{if } 0 \leq x < b_n(\beta), \\ 1, & \text{if } x \geq b_n(\beta). \end{cases}$$

If there is no atom at $-a_n(\alpha)$ or $b_n(\beta)$, then this is just the distribution of $X1\{-a_n(\alpha) < X < b_n(\beta)\}$. But an atom at either endpoint must be split appropriately. Now define

$$\mu_n = nEX(n,0,0), \qquad \tau_n = \{n\operatorname{Var}(X(n,0,0))\}^{1/2}. \tag{3.1}$$

Except when $EX^2 < \infty$, $\operatorname{Var}(X(n,0,0))$ may be replaced by $E(X(n,0,0))^2$ in the definition of τ_n as they are asymptotically equivalent; this makes some computations easier. Next define

$$f_n(\beta) = nEX(n,0,\beta) - nEX(n,0,0), \tag{3.2}$$

$$g_n(\alpha) = nEX(n,0,0) - nEX(n,\alpha,0). \tag{3.3}$$

LEMMA 6. *f_n, g_n are convex, nondecreasing, $f_n(0) = g_n(0) = 0$, and if $\beta_1 < \beta_2 < \beta_3$, $\alpha_1 < \alpha_2 < \alpha_3$,*

$$\frac{f_n(\beta_2) - f_n(\beta_1)}{\beta_2 - \beta_1} \leq b_n(\beta_2)r_n^{1/2} \leq \frac{f_n(\beta_3) - f_n(\beta_2)}{\beta_3 - \beta_2}, \tag{3.4}$$

$$\frac{g_n(\alpha_2) - g_n(\alpha_1)}{\alpha_2 - \alpha_1} \leq a_n(\alpha_2)s_n^{1/2} \leq \frac{g_n(\alpha_3) - g_n(\alpha_2)}{\alpha_3 - \alpha_2}. \tag{3.5}$$

Furthermore, for all α, β,

$$\beta b_n(0) r_n^{1/2} \le f_n(\beta) \le \beta b_n(\beta) r_n^{1/2}, \tag{3.6}$$

$$\alpha a_n(0) s_n^{1/2} \le g_n(\alpha) \le \alpha a_n(\alpha) s_n^{1/2}. \tag{3.7}$$

PROOF. Note that

$$\begin{aligned} f_n(\beta_2) - f_n(\beta_1) &= nEX(n,0,\beta_2) - nEX(n,0,\beta_1) \\ &= nEF^{-1}(Y) 1\{u_n(\beta_1) \le Y < u_n(\beta_2)\} \\ &\le n b_n(\beta_2)(u_n(\beta_2) - u_n(\beta_1)) \\ &= b_n(\beta_2)(\beta_2 - \beta_1) r_n^{1/2}. \end{aligned}$$

This gives the first inequality in (3.4). The other is similar as is (3.5). These imply the convexity and nondecreasing properties of f_n, g_n. Finally, (3.6) follows by using $\beta_1 = \beta < \beta_2 = 0$, and $\beta_1 = 0 < \beta_2 = \beta$, $\beta_2 = \beta < \beta_3 = 0$ and $\beta_2 = 0 < \beta_3 = \beta$ in (3.4). (3.7) is similar. □

We are now ready to state and prove the main limit theorems.

THEOREM 1. *Assume that* $\{r_n\}, \{s_n\}$ *satisfy* (1.1). *Then the class of subsequential limits of the normalized trimmed sums* $\{\gamma_n^{-1}(S_n(s_n, r_n) - \delta_n)\}$ *consists of all laws of the form*

$$Z = \tau N_1 + f(N_2) - g(N_3) + \mu, \tag{3.8}$$

where N_1, N_2, N_3 *are independent* $N(0,1)$, $\tau \ge 0$, $\mu \in \mathbb{R}$ *and* f, g *are convex, nondecreasing functions which vanish at* 0. *Sufficient conditions for*

$$\gamma_{n_i}^{-1}\left(S_{n_i}(s_{n_i}, r_{n_i}) - \delta_{n_i}\right) \to Z \tag{3.9}$$

are

$$\begin{aligned} &\gamma_{n_i}^{-1}\tau_{n_i} \to \tau, \qquad \gamma_{n_i}^{-1} f_{n_i}(\beta) \to f(\beta), \\ &\gamma_{n_i}^{-1} g_{n_i}(\alpha) \to g(\alpha), \qquad \gamma_{n_i}^{-1}\left(\mu_{n_i} - \delta_{n_i}\right) \to \mu \end{aligned} \tag{3.10}$$

for all $\alpha, \beta \in \mathbb{R}$, *where* τ_n, f_n, g_n *and* μ_n *are defined in* (3.1)–(3.3). *Furthermore, the conditions are necessary whenever* (τ, f, g, μ) *is uniquely determined by Z.*

REMARK. We have inserted μ and made f and g vanish at 0 so that when Z is not normal there is a chance for uniqueness in (3.8). There is no uniqueness when Z is normal since if $f(x) = cx$, $g(x) = dx$, then the distribution of Z does not change so long as $\tau^2 + c^2 + d^2$ remains fixed. We will obtain the necessary and sufficient condition in case Z is normal in the next section. Since the uniqueness question arises so naturally we feel there is a good chance it has an affirmative answer.

In the proof of Theorem 1 we will need to use an i.i.d. sequence whose distribution is a conditional one closely related to that of $X(n, \alpha, \beta)$. Define

$$P\{\hat{X}(n, \alpha, \beta) \leq x\} = P\{F^{-1}(Y) \leq x | v_n(\alpha) < Y < u_n(\beta)\},$$

and let $P_n(\alpha, \beta) = u_n(\beta) - v_n(\alpha)$. Therefore the distribution of $\hat{X}(n, \alpha, \beta)$ is obtained from that of $X(n, \alpha, \beta)$ by removing some mass from 0 and renormalizing. Since the atom at 0 does not contribute to the moments of $X(n, \alpha, \beta)$,

$$P_n(\alpha, \beta) E\hat{X}(n, \alpha, \beta) = EX(n, \alpha, \beta),$$

$$P_n(\alpha, \beta) E(\hat{X}(n, \alpha, \beta))^2 = E(X(n, \alpha, \beta))^2.$$

Under condition (3.10) we are able to obtain more useful information. These are straightforward but somewhat tedious estimates. The point of them is that they allowed the condition (3.10) to be stated in terms of moments of the $X(n, \alpha, \beta)$ instead of the more complicated $\hat{X}(n, \alpha, \beta)$.

LEMMA 7. *Assume* (1.1) *and* (3.10) *and let* $m = n - r_n - s_n$. *Then*

$$m_i E\hat{X}(n_i, \alpha, \beta) = \mu_{n_i} + f_{n_i}(\beta) - g_{n_i}(\alpha) + o(\gamma_{n_i}),$$

$$b_{n_i}(\beta) r_{n_i}^{1/2} = O(\gamma_{n_i}), \qquad a_{n_i}(\alpha) s_{n_i}^{1/2} = O(\gamma_{n_i}),$$

$$m_i \operatorname{Var}(\hat{X}(n_i, \alpha, \beta)) = \tau_{n_i}^2 + o(\gamma_{n_i}^2),$$

with the error terms uniform for α, β *in compacts.*

PROOF. First observe that

$$|EX(n,0,0)| \leq \{E(X(n,0,0))^2\}^{1/2} \approx \tau_n n^{-1/2}. \tag{3.11}$$

Then

$$\begin{aligned} mE\hat{X}(n, \alpha, \beta) - mE\hat{X}(n,0,0) &= \frac{mEX(n, \alpha, \beta)}{P_n(\alpha, \beta)} - \frac{mEX(n,0,0)}{P_n(0,0)} \\ &= \frac{mEX(n,0,0)(P_n(0,0) - P_n(\alpha, \beta))}{P_n(\alpha, \beta) P_n(0,0)} \\ &\quad + \frac{m}{n} \frac{f_n(\beta) - g_n(\alpha)}{P_n(\alpha, \beta)} \\ &= O(\tau_n n^{-1/2}(r_n^{1/2} + s_n^{1/2})) \\ &\quad + (f_n(\beta) - g_n(\alpha))(1 + o(1)) \\ &= f_n(\beta) - g_n(\alpha) + o(\tau_n + f_n(\beta) + g_n(\alpha)). \end{aligned}$$

(All error terms will be uniform for α, β in compacts.) Since $mE\hat{X}(n,0,0) = nP_n(0,0)E\hat{X}(n,0,0) = \mu_n$, the first bound then follows from the fact that convex functions converge uniformly on compacts. The next two bounds are immediate from (3.4), (3.5) and (3.10). If $\alpha, \beta > 0$, then

$$\begin{aligned} E(X(n,\alpha,\beta))^2 &= E(X(n,0,0))^2 + E(F^{-1}(Y))^2 1\{v_n(\alpha) < Y \le v_n(0)\} \\ &\quad + E(F^{-1}(Y))^2 1\{u_n(0) \le Y < u_n(\beta)\} \\ &= E(X(n,0,0))^2 + O(a_n^2(\alpha)n^{-1}s_n^{1/2}) + O(b_n^2(\beta)n^{-1}r_n^{1/2}). \end{aligned}$$

A similar estimate holds if α or β is negative but with $a_n(\alpha)$, $b_n(\beta)$ replaced by $a_n(0)$, $b_n(0)$, respectively. Thus by (3.10)

$$\begin{aligned} m_i E(\hat{X}(n_i,\alpha,\beta))^2 &= n_i E(X(n_i,0,0))^2(1+o(1)) + o(\gamma_{n_i}^2) \\ &= n_i E(X(n_i,0,0))^2 + o(\gamma_{n_i}^2) \end{aligned}$$

and $m_i E\hat{X}(n_i,\alpha,\beta) = \mu_{n_i} + O(\gamma_{n_i})$ so that by (3.11) and (3.10)

$$\begin{aligned} m_i(E\hat{X}(n_i,\alpha,\beta))^2 &= m_i^{-1}\mu_{n_i}^2 + O(m_i^{-1}\mu_{n_i}\gamma_{n_i}) + O(m_i^{-1}\gamma_{n_i}^2) \\ &= n_i^{-1}\mu_{n_i}^2 + O(n_i^{-2}(r_{n_i}+s_{n_i})\mu_{n_i}^2) \\ &\quad + O(\tau_{n_i} n_i^{-1/2}\gamma_{n_i}) + o(\gamma_{n_i}^2) \\ &= n_i^{-1}\mu_{n_i}^2 + O(\tau_{n_i}^2 n_i^{-1}(r_{n_i}+s_{n_i})) + o(\gamma_{n_i}^2) \\ &= n_i^{-1}\mu_{n_i}^2 + o(\gamma_{n_i}^2). \end{aligned}$$

Subtracting these expressions gives the final result of the lemma. □

PROOF OF THEOREM 1. There will be three parts to the proof: (i) the sufficiency of (3.10); (ii) the proof that all possible limits are included in (3.8) and the necessity of (3.10) when (τ, f, g, μ) are determined by Z; and (iii) construction of a universal law for which all the laws in (3.8) arise as subsequential limits. This construction will be given in Theorem 5 in Section 5.

(i) If $Y_{n,s_n} = v_n(\alpha)$ and $Y_{n,n-r_n+1} = u_n(\beta)$ are given, then, conditionally, the summands X_k with $v_n(\alpha) < Y_k < u_n(\beta)$, are i.i.d. and distributed as $\hat{X}(n,\alpha,\beta)$. Thus we let

$$W_m(n,\alpha,\beta) = \sum_{k=1}^{m} \hat{X}_k(n,\alpha,\beta),$$

where $\{\hat{X}_k(n,\alpha,\beta)\}$ are i.i.d. and $m = n - r_n - s_n$. Then

$$\begin{aligned} &P\{\gamma_n^{-1}(S_n(s_n,r_n) - \delta_n) \le x\} \\ &\qquad = \int\int P\{\gamma_n^{-1}(W_m(n,\alpha,\beta) - \delta_n) \le x\}\nu_n(d\alpha,d\beta), \end{aligned} \tag{3.12}$$

where ν_n is the measure associated with

$$H_n(\alpha,\beta) = P\{Y_{n,s_n} \ge v_n(\alpha), Y_{n,n-r_n+1} \le u_n(\beta)\}.$$

Recalling (2.18) and Lemma 2, we have

$$H_n(\alpha,\beta) = P\{U_n^{-*}(\alpha) < \alpha, U_n^{+*}(\beta) < \beta\} \to \Phi(\alpha)\Phi(\beta). \tag{3.13}$$

For the integrand we write

$$\begin{aligned} W_m(n,\alpha,\beta) - \delta_n &= \left(W_m(n,\alpha,\beta) - EW_m(n,\alpha,\beta)\right) \\ &\quad + \left(EW_m(n,\alpha,\beta) - \delta_n\right) \end{aligned} \tag{3.14}$$

and we will determine the behavior of each of the terms on the right side. If $\tau > 0$, then τ_{n_i} and γ_{n_i} are comparable so that

$$m_i \operatorname{Var}\left(\hat{X}(n_i,\alpha,\beta)\right) \sim \tau_{n_i}^2 \sim \tau^2\gamma_{n_i}^2, \tag{3.15}$$

uniformly for α, β in compacts by Lemma 7. Next, consider

$$\begin{aligned} &m_i\gamma_{n_i}^{-3} E\left|\hat{X}(n_i,\alpha,\beta) - E\hat{X}(n_i,\alpha,\beta)\right|^3 \\ &\quad\le 2\gamma_{n_i}^{-3} m_i\left(b_{n_i}(\beta) + a_{n_i}(\alpha)\right)\operatorname{Var}\left(\hat{X}(n_i,\alpha,\beta)\right) \\ &\quad\sim 2\tau^2\gamma_{n_i}^{-1}\left(b_{n_i}(\beta) + a_{n_i}(\alpha)\right) \to 0 \end{aligned}$$

by (3.15) and Lemma 7. This convergence is also uniform on compacts. By the Berry–Esseen theorem ([6], page 542) we have

$$P\left\{\gamma_{n_i}^{-1}\left(W_{m_i}(n_i,\alpha,\beta) - EW_{m_i}(n_i,\alpha,\beta)\right) \le x\right\} \to \Phi(\tau^{-1}x), \tag{3.16}$$

uniformly on compacts. By Lemma 7

$$\gamma_{n_i}^{-1}\left(EW_{m_i}(n_i,\alpha,\beta) - \delta_{n_i}\right) \to \mu + f(\beta) - g(\alpha).$$

Recalling (3.14) and (3.16),

$$P\left\{\gamma_{n_i}^{-1}\left(W_{m_i}(n_i,\alpha,\beta) - \delta_{n_i}\right) \le x\right\} \to \Phi\left(\tau^{-1}(x - f(\beta) + g(\alpha) - \mu)\right),$$

uniformly on compacts. Then by (3.12) and (3.13)

$$\begin{aligned} &P\left\{\gamma_{n_i}^{-1}\left(S_{n_i}(s_{n_i}, r_{n_i}) - \delta_{n_i}\right) \le x\right\} \\ &\quad\to \int\!\!\int \Phi\left(\tau^{-1}(x - f(\beta) + g(\alpha) - \mu\right)\varphi(\alpha)\varphi(\beta)\,d\alpha\,d\beta \\ &\quad= P\{\tau N_1 + f(N_2) - g(N_3) + \mu \le x\}, \end{aligned}$$

where φ is the $N(0,1)$ density function. If $\tau = 0$, then by Lemma 7

$$\begin{aligned} &P\left\{\gamma_{n_i}^{-1}\left|W_{m_i}(n_i,\alpha,\beta) - EW_{m_i}(n_i,\alpha,\beta)\right| \ge \varepsilon\right\} \\ &\quad\le \varepsilon^{-2}\gamma_{n_i}^{-2} m_i \operatorname{Var}\left(\hat{X}(n_i,\alpha,\beta)\right) \\ &\quad= \varepsilon^{-2}\gamma_{n_i}^{-2}\tau_{n_i}^2 + o(1) = o(1), \end{aligned}$$

uniformly on compacts. Thus, in this case for any bounded continuous h,

$$\begin{aligned} &Eh\left(\gamma_{n_i}^{-1}\left(S_{n_i}(s_{n_i}, r_{n_i}) - \delta_{n_i}\right)\right) \\ &\quad\to \int\int h(f(\beta) - g(\alpha) + \mu)\varphi(\alpha)\varphi(\beta)\,d\alpha\,d\beta \\ &\quad= Eh(\tau N_1 + f(N_2) - g(N_3) + \mu), \end{aligned}$$

which completes the proof of sufficiency.

(ii) We assume (3.9) where Z is any random variable. By (2.21) and (2.4)

$$\tau_{n_i}^2 \le n_i \lambda_{n_i}(0,0) = O\left(\gamma_{n_i}^2\right), \qquad b_{n_i}^2(\beta) r_{n_i} + a_{n_i}^2(\alpha) s_{n_i} = O\left(\gamma_{n_i}^2\right)$$

and so by Lemma 6, $f_{n_i}(\beta) = O(\gamma_{n_i})$, $g_{n_i}(\alpha) = O(\gamma_{n_i})$. Next $\mu_n = \delta_n(0,0)$ as defined in (2.13) and so $\delta_{n_i} - \mu_{n_i} = O(\gamma_{n_i})$ by Proposition 1. Thus by taking a further subsequence we may assume that (3.10) holds along this new subsequence for some τ, f, g and μ. Furthermore, τ will be nonnegative and f, g convex, nondecreasing and vanishing at 0 by Lemma 6. Thus Z has the representation (3.8) by part (i) of the proof. Also if (τ, f, g, μ) are determined by Z, then (3.10) must hold for the original subsequence by the usual subsequence argument. □

We conclude this section with the domain of partial attraction conditions for Z which do not depend on knowing $\{\gamma_n\}, \{\delta_n\}$.

THEOREM 2. *Assume that* $\{r_n\}, \{s_n\}$ *satisfy* (1.1) *and let* Z *be a nondegenerate random variable given by* (3.8). *Sufficient conditions for the existence of* $\{\gamma_n\}, \{\delta_n\}$ *such that* (3.9) *holds are that for some* α_0, β_0 *such that* $g'(\alpha_0)$, $f'(\beta_0)$ *exist and* $\tau^2 + (g'(\alpha_0))^2 + (f'(\beta_0))^2 \ne 0$,

$$\left(n_i^{1/2}\sigma_{n_i}(\alpha_0, \beta_0)\right)^{-1}\tau_{n_i} \to \tau\left(\tau^2 + (g'(\alpha_0))^2 + (f'(\beta_0))^2\right)^{-1/2}, \tag{3.17}$$

$$\left(n_i^{1/2}\sigma_{n_i}(\alpha_0, \beta_0)\right)^{-1} f_{n_i}(\beta) \to f(\beta)\left(\tau^2 + (g'(\alpha_0))^2 + (f'(\beta_0))^2\right)^{-1/2}, \tag{3.18}$$

$$\left(n_i^{1/2}\sigma_{n_i}(\alpha_0, \beta_0)\right)^{-1} g_{n_i}(\alpha) \to g(\alpha)\left(\tau^2 + (g'(\alpha_0))^2 + (f'(\beta_0))^2\right)^{-1/2}, \tag{3.19}$$

where τ_n, f_n, g_n *are defined in* (3.1)–(3.3). *Furthermore, the conditions are necessary* [*even for all* α_0, β_0, *where* g', f' *exist and* $\tau^2 + (g'(\alpha_0))^2 + (f'(\beta_0))^2 \ne 0$] *whenever* (τ, f, g, μ) *is determined by* Z. *In this case one may take*

$$\begin{aligned} \gamma_{n_i} &= n_i^{1/2}\sigma_{n_i}(\alpha_0, \beta_0)\left(\tau^2 + (g'(\alpha_0))^2 + (f'(\beta_0))^2\right)^{-1/2}, \\ \delta_{n_i} &= \mu_{n_i} - \mu\gamma_{n_i}, \end{aligned} \tag{3.20}$$

where μ_n *is defined in* (3.1), $\sigma_n(\alpha, \beta)$ *in* (2.2) *and* α_0, β_0 *are such that* $g'(\alpha_0)$, $f'(\beta_0)$ *exist and* $\tau^2 + (g'(\alpha_0))^2 + (f'(\beta_0))^2 \ne 0$.

PROOF. Assume (3.17)–(3.19). Then (3.9) follows immediately from Theorem 1 if we take $\gamma_{n_i}, \delta_{n_i}$ as in (3.20). For the converse, suppose that (3.9) holds and (τ, f, g, μ) is determined by Z. Then (3.10) holds for some γ_{n_i} by Theorem 1. By

(3.4), for $\beta_1 < \beta_2 < \beta_3$,

$$\frac{f(\beta_2) - f(\beta_1)}{\beta_2 - \beta_1} \le \liminf \gamma_{n_i}^{-1} b_{n_i}(\beta_2) r_{n_i}^{1/2}$$

$$\le \limsup \gamma_{n_i}^{-1} b_{n_i}(\beta_2) r_{n_i}^{1/2} \le \frac{f(\beta_3) - f(\beta_2)}{\beta_3 - \beta_2}.$$

Letting $\beta_1 \uparrow \beta_2$, $\beta_3 \downarrow \beta_2$ gives

$$\gamma_{n_i}^{-1} b_{n_i}(\beta) r_{n_i}^{1/2} \to f'(\beta) \tag{3.21}$$

for all β for which $f'(\beta)$ exists. Similarly

$$\gamma_{n_i}^{-1} a_{n_i}(\alpha) s_{n_i}^{1/2} \to g'(\alpha) \tag{3.22}$$

when $g'(\alpha)$ exists. Next

$$E(X(n,\alpha,\beta))^2 = E(X(n,0,0))^2 + O\big(a_n^2(\alpha \vee 0) n^{-1} s_n^{1/2}\big) + O\big(b_n^2(\beta \vee 0) n^{-1} r_n^{1/2}\big)$$

so that by (3.21) and (3.22)

$$n_i E(X(n_i,\alpha,\beta))^2 = n_i E(X(n_i,0,0))^2 + o\big(\gamma_{n_i}^2\big).$$

Furthermore,

$$nE\big((X \wedge b_n(\beta)) \vee (-a_n(\alpha))\big) = \mu_n + f_n(\beta) - g_n(\alpha) + b_n(\beta)\big(r_n - \beta r_n^{1/2}\big) - a_n(\alpha)\big(s_n - \alpha s_n^{1/2}\big)$$

so that by (3.10), (3.21) and (3.22)

$$n_i E\Big(\big(X \wedge b_{n_i}(\beta)\big) \vee \big(-a_{n_i}(\alpha)\big)\Big) = \mu_{n_i} + O(\gamma_{n_i}) + O\Big(\gamma_{n_i}\big(r_{n_i}^{1/2} + s_{n_i}^{1/2}\big)\Big) = \mu_{n_i} + o\big(n_i^{1/2}\gamma_{n_i}\big).$$

Thus by (3.11) and (3.10)

$$n_i\Big(E\Big(\big(X \wedge b_{n_i}(\beta)\big) \vee \big(-a_{n_i}(\alpha)\big)\Big)\Big)^2 = n_i^{-1}\mu_{n_i}^2 + o\big(n_i^{-1/2}\mu_{n_i}\gamma_{n_i}\big) + o\big(\gamma_{n_i}^2\big) = n_i^{-1}\mu_{n_i}^2 + o\big(\gamma_{n_i}^2\big).$$

Putting all the pieces together, we have for all $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} n_i \sigma_{n_i}^2(\alpha,\beta) &= n_i E(X(n_i,\alpha,\beta))^2 + b_{n_i}^2(\beta)\big(r_{n_i} - \beta r_{n_i}^{1/2}\big) + a_{n_i}^2(\alpha)\big(s_{n_i} - \alpha s_{n_i}^{1/2}\big) \\ &\quad - n_i\Big(E\Big(\big(X \wedge b_{n_i}(\beta)\big) \vee \big(-a_{n_i}(\alpha)\big)\Big)\Big)^2 \\ &= n_i E(X(n_i,0,0))^2 + b_{n_i}^2(\beta) r_{n_i} + a_{n_i}^2(\alpha) s_{n_i} - n_i^{-1}\mu_{n_i}^2 + o\big(\gamma_{n_i}^2\big) \\ &= \tau_{n_i}^2 + b_{n_i}^2(\beta) r_{n_i} + a_{n_i}^2(\alpha) s_{n_i} + o\big(\gamma_{n_i}^2\big) \\ &\sim \big\{\tau^2 + (f'(\beta))^2 + (g'(\alpha))^2\big\}\gamma_{n_i}^2. \end{aligned}$$

Now take α_0, β_0 so that $\tau^2 + (f'(\beta_0))^2 + (g'(\alpha_0))^2 \neq 0$. ($\alpha_0$, β_0 must exist since Z is nondegenerate.) Then (3.17)–(3.19) follow from (3.10). □

4. Asymptotic normality. We are now ready to prove the theorem giving necessary and sufficient conditions for asymptotic normality of the trimmed sums. Recall that

$$\lambda_n(\alpha, \beta) = E\big((X \wedge b_n(\beta)) \vee (-a_n(\alpha))\big)^2.$$

THEOREM 3. *Assume* (1.1). *There exist sequences* $\{\delta_n\}, \{\gamma_n\}$ *such that*

$$\gamma_n^{-1}(S_n(s_n, r_n) - \delta_n) \to N(0,1) \tag{4.1}$$

iff for all $\alpha, \beta \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{\lambda_n(\alpha, \beta)}{\lambda_n(0,0)} = 1. \tag{4.2}$$

In this case, one may take $\gamma_n = \{n\sigma_n^2(0,0)\}^{1/2}$ *and* $\delta_n = \mu_n$ *as in* (3.1). *Furthermore,* (4.1) *holds for a subsequence iff* (4.2) *holds for that subsequence.*

PROOF. First suppose that $EX^2 = \infty$, (4.2) holds for a subsequence, and let $\gamma_n = \{n\sigma_n^2(0,0)\}^{1/2} \sim \{n\lambda_n(0,0)\}^{1/2}$. Since

$$\begin{aligned} \lambda_n(\alpha, \beta) &= E(X(n, \alpha, \beta))^2 + b_n^2(\beta)(r_n - \beta r_n^{1/2})n^{-1} \\ &\quad + a_n^2(\alpha)(s_n - \alpha s_n^{1/2})n^{-1} \\ &\sim E(X(n, \alpha, \beta))^2 + b_n^2(\beta) r_n n^{-1} + a_n^2(\alpha) s_n n^{-1}, \end{aligned} \tag{4.3}$$

we may take a further subsequence so that

$$\gamma_{n_i}^{-2} n_i E(X(n_i, \alpha, \beta))^2 \to \tau^2, \qquad \gamma_{n_i}^{-1} b_{n_i}(\beta) r_{n_i}^{1/2} \to b, \qquad \gamma_{n_i}^{-1} a_{n_i}(\alpha) s_{n_i}^{1/2} \to a.$$

The limits are independent of α and β since all three sequences are nondecreasing in α and β and by (4.3), $\tau^2 + b^2 + a^2 = 1$. By (3.6) and (3.7)

$$\gamma_{n_i}^{-1} f_{n_i}(\beta) \to b\beta, \qquad \gamma_{n_i}^{-1} g_{n_i}(\alpha) \to a\alpha.$$

Now since $EX^2 = \infty$,

$$\gamma_{n_i}^{-1}\tau_{n_i} \sim \gamma_{n_i}^{-1}\big\{n_i E(X(n_i, 0, 0))^2\big\}^{1/2} \to \tau.$$

Taking $\delta_n = \mu_n$, we have (3.10) so that by Theorem 1, (3.9) holds with

$$Z = \tau N_1 + bN_2 - aN_3,$$

which is $N(0,1)$ since $\tau^2 + b^2 + a^2 = 1$. Since the limit is independent of the original subsequence chosen this also applies to the entire sequence. If $EX^2 < \infty$, let $\sigma^2 = \operatorname{Var} X$. Then with $\gamma_n = \{n\sigma_n^2(0,0)\}^{1/2}$ we have $\gamma_n \sim \sigma n^{1/2}$, $\tau_n \sim \sigma n^{1/2}$ so that $\tau = 1$ in (3.10). Next

$$n^{-1} b_n^2(\beta) r_n \sim b_n^2(\beta) n^{-1}(r_n - \beta r_n^{1/2}) \le b_n^2(\beta) P\{X \ge b_n(\beta)\} \to 0$$

unless $b_n(\beta)$ is bounded in which case $n^{-1}b_n^2(\beta)r_n \to 0$ by (1.1). Thus we have $f \equiv 0$ in this case by Lemma 6 and a similar argument applies to g so (4.1) holds in this case also.

Now suppose that (4.1) holds for a subsequence. We may assume $EX^2 = \infty$ since (4.2) holds when $EX^2 < \infty$. As in the proof of Theorem 1, part (ii), there is a further subsequence such that (3.10) holds and then

$$Z = \tau N_1 + f(N_2) - g(N_3) + \mu,$$

where Z is $N(0,1)$. By the Cramér–Lévy theorem ([6], page 525), $f(N_2)$ and $g(N_3)$ must be normal. This forces f and g to be linear (recall that f, g are monotone and vanish at 0) and so $\mu = 0$. If $f(x) = bx$, $g(x) = ax$, then by (3.21) and (3.22)

$$\gamma_{n_i}^{-1} b_{n_i}(\beta) r_{n_i}^{1/2} \to b, \qquad \gamma_{n_i}^{-1} a_{n_i}(\alpha) s_{n_i}^{1/2} \to a.$$

Furthermore,

$$E(X(n,\alpha,\beta))^2 = E(X(n,0,0))^2 + O\big(a_n^2(\alpha \vee 0) s_n^{1/2} n^{-1}\big) \\ + O\big(b_n^2(\beta \vee 0) r_n^{1/2} n^{-1}\big)$$

so by (4.3), for all $\alpha, \beta \in \mathbb{R}$,

$$\gamma_{n_i}^{-2} n_i \lambda_{n_i}(\alpha, \beta) \to \tau^2 + b^2 + a^2 = 1$$

since this is the variance of Z. Thus we have (4.2) along the subsequence. This is enough since we could have started with an arbitrary subsequence of the given subsequence. □

Now we give some sufficient conditions for asymptotic normality of $S_n(s_n, r_n)$ which are quite general and even easier to check than (4.2).

THEOREM 4. *Assume* (1.1). *Each of the following conditions is sufficient for asymptotic normality of* $S_n(s_n, r_n)$:

(a) $a_n(\alpha) \sim a_n(0)$, $b_n(\beta) \sim b_n(0)$ *for all* $\alpha, \beta \in \mathbb{R}$;

(b) $x^\varepsilon P\{X > x\}$ *and* $x^\varepsilon P\{X < -x\}$ *are eventually nonincreasing for some* $\varepsilon > 0$;

(c) $x^\varepsilon P\{X > x\}$ *and* $x^\varepsilon P\{X < -x\}$ *are slowly varying for some* $\varepsilon > 0$.

PROOF. By (2.4), for $\alpha > 0$,

$$\begin{aligned} \lambda_n^-(\alpha) - \lambda_n^-(-\alpha) &\le \big(a_n^2(\alpha) - a_n^2(-\alpha)\big) P\{X < -a_n(-\alpha)\} \\ &= O(\lambda_n^-(\alpha)) a_n^{-2}(\alpha) \big(a_n^2(\alpha) - a_n^2(-\alpha)\big) \\ &= o\big(\lambda_n^-(\alpha)\big) \end{aligned}$$

under (a). Thus $\lambda_n^-(\alpha) \sim \lambda_n^-(-\alpha)$ which implies $\lambda_n^-(\alpha) \sim \lambda_n^-(0)$; similarly $\lambda_n^+(\beta) \sim \lambda_n^+(0)$. This implies (4.2). Under (b), we have for $\beta > 0$, if $b_n(-\beta) < b_n(\beta)$, then

$$\lim_{x \uparrow b_n(\beta)} x^\varepsilon P\{X > x\} \le b_n^\varepsilon(-\beta) P\{X > b_n(-\beta)\}$$

and so

$$b_n^\varepsilon(\beta)n^{-1}(r_n - \beta r_n^{1/2}) \le b_n^\varepsilon(-\beta)n^{-1}(r_n + \beta r_n^{1/2}).$$

Thus $b_n(\beta) \sim b_n(-\beta)$; a similar argument holds for $a_n(\alpha)$. Now use (a). If $x^\varepsilon P\{X > x\}$ is slowly varying, then for any $\delta > 0$,

$$((1+\delta)b_n(0))^\varepsilon P\{X > (1+\delta)b_n(0)\} \sim b_n^\varepsilon(0)P\{X > b_n(0)\} \le b_n^\varepsilon(0)n^{-1}r_n$$

so that for $1 < \rho < (1+\delta)^\varepsilon$,

$$P\{X > (1+\delta)b_n(0)\} \le \rho(1+\delta)^{-\varepsilon}n^{-1}r_n < n^{-1}(r_n - \beta r_n^{1/2}).$$

Thus, for large n, $b_n(\beta) \le (1+\delta)b_n(0)$; similar arguments complete the proof. □

Part (c) shows that $S_n(s_n, r_n)$ is always asymptotically normal when X is in the domain of attraction of a stable law (even when $r_n \ne s_n$) *except* for the extremely asymmetric case when either the right or left tail of F may not vary regularly. In the extreme case when $P\{X > x\}$ is dominating, if $r_n = s_n$, then it is easy to see that $\lambda_n^-(\alpha) = o(\lambda_n^+(\beta))$ and so $\lambda_n(\alpha, \beta) \sim \lambda_n^+(\beta) \sim \lambda_n^+(0) \sim \lambda_n(0,0)$ since $P\{X > x\}$ is regularly varying. But if s_n is much smaller than r_n, asymptotic normality may fail. Here is an example. Let $P\{X > x\} = x^{-1} \wedge 2^{-1}$,

$$P\{X < -x\} = \begin{cases} k^{-1}2^{-k}, & 2^k \le x < 2^{k+1},\ k = 1,2,\dots, \\ 2^{-1}, & x < 2. \end{cases}$$

Then $nP\{X > nx\} \to x^{-1}$ and $nP\{X < -nx\} \to 0$ so X is in the domain of attraction of an asymmetric Cauchy. Let $s_n = [n^{1/2}]$, $r_n = [n^\delta]$ where $\delta > \frac{1}{2}$. By Lemma 2.1 of [15]

$$\lambda_n^+(\beta) = \int_0^{b_n(\beta)} 2yP\{X > y\}\,dy \sim 2b_n(\beta) \sim 2nr_n^{-1} \tag{4.4}$$

since $b_n(\beta) = n(r_n - \beta r_n^{1/2})^{-1}$. Consider $n_k = k^2 2^{2k}$; then $n_k^{-1}s_{n_k} = k^{-1}2^{-k}$ and so $a_{n_k}(\alpha) = 2^{k+1}$ if $\alpha > 0$ and $a_{n_k}(\alpha) = 2^k$ if $\alpha \le 0$ when k is large. Thus, for $\alpha > 0$,

$$\begin{aligned} \lambda_{n_k}^-(\alpha) &= \sum_{j=2}^{k} 2^{2j}\big((j-1)^{-1}2^{-(j-1)} - j^{-1}2^{-j}\big) + 2^{2k+2}k^{-1}2^{-k} \sim 6k^{-1}2^k, \\ \lambda_{n_k}^-(-\alpha) &= \sum_{j=2}^{k} 2^{2j}\big((j-1)^{-1}2^{-(j-1)} - j^{-1}2^{-j}\big) + 2^{2k}k^{-1}2^{-k} \sim 3k^{-1}2^k. \end{aligned} \tag{4.5}$$

Since

$$\lambda_{n_k}^+(\beta) \sim 2n_k^{1-\delta} \sim 2k^{2(1-\delta)}2^{2k(1-\delta)} = o\big(\lambda_{n_k}^-(\alpha)\big),$$

we have $\lambda_{n_k}(\alpha, \beta) \sim \lambda_{n_k}^-(\alpha)$ and (4.2) fails by (4.5). This example could be made continuous by spreading the mass at -2^k uniformly over $[-2^k - 1, -2^k]$.

Next we will show that condition (a) of Theorem 4 is also necessary for asymptotic normality of $S_n(s_n, r_n)$ under rather special conditions.

PROPOSITION 2. *Assume* (1.1), $r_n = s_n$ *and suppose that* X *is symmetric and not in the domain of partial attraction of the normal. Then asymptotic normality of* $S_n(s_n, r_n)$ *implies* $a_n(\alpha) \sim a_n(0)$, $b_n(\beta) \sim b_n(0)$ *for all* $\alpha, \beta \in \mathbb{R}$.

REMARK. The conclusion holds for all nondecreasing $\{r_n\}, \{s_n\}$, even if they are not equal. But the proof is somewhat harder and we will omit it.

PROOF. Since $P\{X > x\} = P\{X < -x\}$ and $r_n = s_n$ we have $a_n(\alpha) = b_n(\alpha)$ for all $\alpha \in \mathbb{R}$. Then $\lambda_n(\alpha, \alpha) = 2\lambda_n^+(\alpha)$ so by (4.2) $\lambda_n^+(\alpha) \sim \lambda_n^+(0)$. For $\alpha > 0$,

$$\lambda_n^+(\alpha) - \lambda_n^+(-\alpha) \geq \left(b_n^2(\alpha) - b_n^2(-\alpha)\right)P\{X \geq b_n(\alpha)\}. \tag{4.6}$$

By the symmetry and the condition that X is not in the domain of partial attraction of the normal (see page 190 of [7]) we have

$$b_n^2(0)n^{-1}r_n \approx \lambda_n^+(0),$$

which in conjunction with (4.6) gives

$$\frac{\lambda_n^+(\alpha) - \lambda_n^+(-\alpha)}{\lambda_n^+(0)} \geq \frac{b_n^2(\alpha) - b_n^2(-\alpha)}{b_n^2(0)}c$$

for some $c > 0$ and so $b_n(\alpha) \sim b_n(0)$. □

The assumption of symmetry in the proposition is critical. Even if one assumes $P\{X > x\} \sim P\{X < -x\}$ the conclusion may fail. Here is an example. Take $r_n = s_n = [n^{3/4}]$ and let $P\{X > x\} = (\log x)^{-1}$, $x > e^2$,

$$P\{X < -x\} = \begin{cases} (k+2)^{-2}, & e^{k^2} \leq x < e^{(k+1)^2},\ k = 1,2,\dots, \\ 2^{-1}, & -e^2 \leq x < e. \end{cases}$$

Note that $P\{X > x\} \sim P\{X < -x\}$. Since both tails are slowly varying, we have

$$\begin{aligned} \lambda_n^+(\beta) &\sim b_n^2(\beta)P\{X > b_n(\beta)\} \sim b_n^2(\beta)n^{-1/4}, \\ \lambda_n^-(\alpha) &\sim a_n^2(\alpha)P\{X < -a_n(\alpha)\} \sim a_n^2(\alpha)n^{-1/4}. \end{aligned} \tag{4.7}$$

Next, observe that

$$\begin{aligned} \log b_n(\beta) &= nr_n^{-1}\left(1 - \beta r_n^{-1/2}\right)^{-1} \\ &= nr_n^{-1} + O\left(nr_n^{-3/2}\right) = nr_n^{-1} + o(1) \end{aligned}$$

so that $b_n(\beta) \sim \exp(nr_n^{-1})$. Suppose that for some k and α we have $a_n(\alpha) = \exp(k^2)$. Then

$$\begin{aligned} n^{-1}\left(s_n - \alpha s_n^{1/2}\right) &\leq P\{X \leq -a_n(\alpha)\} \\ &= P\left\{X < -\exp\left((k-1)^2\right)\right\} = (k+1)^{-2}. \end{aligned}$$

Since the left side is $\sim n^{-1/4}$, this gives $k^8 = O(n)$. Then

$$\begin{aligned} nr_n^{-1} = ns_n^{-1} &= n\left(s_n - \alpha s_n^{1/2}\right)^{-1}\left(1 - \alpha s_n^{-1/2}\right) \geq (k+1)^2\left(1 + O(n^{-3/8})\right) \\ &= (k+1)^2\left(1 + O(k^{-3})\right) \end{aligned}$$

and so

$$b_n(\beta) \sim \exp\left(nr_n^{-1}\right) \geq \exp\left((k+1)^2\left(1+O(k^{-3})\right)\right) \sim \exp\left((k+1)^2\right).$$

Thus $a_n(\alpha) = o(b_n(\beta))$ so by (4.7), $\lambda_n^-(\alpha) = o(\lambda_n^+(\beta))$ and so (4.2) holds. But we can clearly choose a subsequence n_k so that for $\alpha > 0$,

$$a_{n_k}(-\alpha) = \exp\left((k-1)^2\right), \qquad a_{n_k}(\alpha) = \exp(k^2)$$

so that $a_n(\alpha) \sim a_n(0)$ must fail. As before, this can be modified easily to make F continuous. Furthermore, by alternating the flat stretches between the two tails one can make both $a_n(\alpha) \sim a_n(0)$ and $b_n(\beta) \sim b_n(0)$ fail.

We now perturb this example slightly by making the distribution symmetric with both tails like the negative tail above and keep $\{r_n\}$, $\{s_n\}$ as before. Then, as we have seen, $a_n(\alpha) \sim a_n(0)$ fails and so $S_n(s_n, r_n)$ is not asymptotically normal by Proposition 2. If, on the other hand, we take both tails like the positive tail above, we do get asymptotic normality (see Example 1 in Section 6). This is the example mentioned in the Introduction with two distributions with asymptotic tails where the trimmed sum is asymptotically normal for one but not the other.

5. A universal law. The purpose of this section is the following construction.

THEOREM 5. *Given* $\{r_n\}, \{s_n\}$ *satisfying* (1.1), *there exists a distribution for* X *and* $\{\gamma_n\}, \{\delta_n\}$ *such that for every* Z *satisfying* (3.8) *there is a subsequence* $\{n_i\}$ *such that* (3.9) *holds.*

PROOF. We first need to define some sequences. Let $n_1 = 1$ and for $i \geq 1$, let

$$n_{i+1} = \min\left\{n\colon n^{-1}r_n \leq \left(i^5 n_i\right)^{-1} r_{n_i} \text{ and } n^{-1}s_n \leq \left(i^5 n_i\right)^{-1} s_{n_i}\right\},$$

$$\xi_i = n_i^{-1} r_{n_i}, \qquad \eta_i = n_i^{-1} r_{n_i}^{1/2}, \qquad \rho_i = n_i^{-1} s_{n_i}, \qquad \zeta_i = n_i^{-1} s_{n_i}^{1/2}.$$

Next we choose $x_1 = 1$ and then $x_i \uparrow \infty$, $y_i \uparrow \infty$ to satisfy

$$\text{(5.1)} \quad x_i r_{n_i}^{1/2} = y_i s_{n_i}^{1/2}, \qquad x_{i+1} \geq \xi_{i+1}^{-1} x_i (i+1)^2, \qquad y_{i+1} \geq \rho_{i+1}^{-1} y_i (i+1)^2.$$

Now we take a sequence $(\bar{\tau}_i, \bar{f}_i, \bar{g}_i, \bar{\mu}_i)$ which is dense (in the sense of uniform convergence on compacts for f, g) in the set of all (τ, f, g, μ) and suppose that each member of the sequence is repeated infinitely often. We may assume that $\bar{f}_i, \bar{g}_i$ are strictly convex and continuously differentiable. By repeating the sequence more often if necessary, we may also assume that $\bar{\tau}_i \leq i$ and choose $w_i \uparrow \infty$, $t_i \uparrow \infty$ but slowly enough so that

$$w_i\left(1 + \left(\bar{f}_i'(w_i)\right)^2\right) \leq r_{n_i}^{1/4}, \qquad \bar{f}_i'(w_i) \leq i,$$

$$t_i\left(1 + \left(\bar{g}_i'(t_i)\right)^2\right) \leq s_{n_i}^{1/4}, \qquad \bar{g}_i'(t_i) \leq i.$$

Now we can define the distribution function F of X. We let h_{i+} denote the inverse function of the monotone function $\bar{f}_i'$ and h_{i-} the inverse function of $\bar{g}_i'$.

Then define

$$F(x) = 1 - \xi_i + \eta_i h_{i+}(xx_i^{-1}) \quad \text{for } (\bar{f}_i'(-w_i) \vee i^{-1})x_i \le x \le \bar{f}_i'(w_i)x_i, \tag{5.2}$$
$$F \text{ has an atom of } \bar{\tau}_i^2 i^2 \xi_i \text{ at } i^{-1}x_i,$$
$$F \text{ has no other mass on } [i^{-1}\xi_i x_i, \bar{f}_i'(w_i)x_i];$$

$$F(x) = \rho_i - \zeta_i h_{i-}(-xy_i^{-1}) \quad \text{for } -\bar{g}_i'(t_i)y_i \le x \le -(\bar{g}_i'(-t_i) \vee i^{-1})y_i, \tag{5.3}$$
$$F \text{ has no other mass on } [-\bar{g}_i'(t_i)y_i, -i^{-1}\rho_i y_i].$$

We do this for each i to build the distribution function F. Any remaining mass can be filled in arbitrarily. We must check that the intervals do not overlap and that the pieces of F preserve the monotonicity. For the former, we have

$$\bar{f}_i'(w_i)x_i \le ix_i \le (i+1)^{-1}\xi_{i+1}x_{i+1}$$

by (5.1). For the latter, we need

$$F(\bar{f}_i'(w_i)x_i) \le F((i+1)^{-1}\xi_{i+1}x_{i+1}).$$

But

$$\begin{aligned} F(\bar{f}_i'(w_i)x_i) &= 1 - \xi_i + \eta_i w_i = 1 - \xi_i + o(\xi_i) \\ &\le 1 - i^5\xi_{i+1}(1 + o(1)) \\ &\le 1 - \xi_{i+1} - \eta_{i+1}w_{i+1} - \bar{\tau}_{i+1}^2(i+1)^2\xi_{i+1} \\ &\le F((i+1)^{-1}\xi_{i+1}x_{i+1}) \end{aligned}$$

for i sufficiently large. (The definition of F on any fixed interval is irrelevant.) The conditions for the left tail may be checked in a similar way. Define $\gamma_{n_i} = n_i x_i \eta_i = n_i y_i \zeta_i$. Next we will verify that if we fix j and restrict ourselves to a subsequence where only $(\bar{\tau}_j, \bar{f}_j, \bar{g}_j, \bar{\mu}_j)$ occurs, then (3.10) holds with limits $\bar{\tau}_j$, $\bar{f}_j(\beta)$, $\bar{g}_j(\alpha)$ and $\bar{\mu}_j$, respectively. Recalling the definition of $X(n, \alpha, \beta)$ from Section 3, we have

$$EX(n, \alpha, \beta) = \int_{v_n(\alpha)}^{u_n(\beta)} F^{-1}(u)\, du,$$

and then by (3.2)

$$f_n(\beta) = n\int_{u_n(0)}^{u_n(\beta)} F^{-1}(u)\, du.$$

Now $u_{n_i}(\beta) = 1 - \xi_i + \beta\eta_i$ and for each fixed β this will be in the range of F as defined in (5.2) for large i. Computing the inverse yields

$$F^{-1}(u) = x_i \bar{f}_j'(\eta_i^{-1}(u - 1 + \xi_i))$$

and so

$$f_{n_i}(\beta) = n_i \int_{u_{n_i}(0)}^{u_{n_i}(\beta)} x_i \bar{f}_j'\big(\eta_i^{-1}(u - 1 + \xi_i)\big)\, du$$

$$= n_i x_i \eta_i \int_0^{\beta} \bar{f}_j'(v)\, dv$$

$$= \gamma_{n_i} \bar{f}_j(\beta).$$

g_{n_i} is similar. Now we must estimate τ_{n_i}. First note that $b_{n_i}(0)$ will be in the interval where F is defined by (5.2) and $-a_{n_i}(0)$ in the interval where F is defined by (5.3). The main contribution to τ will come from the atom. We have

$$E(X(n_i,0,0))^2 = (i^{-1}x_i)^2 \bar{\tau}_j^2 i^2 \xi_i + O\big(w_i \eta_i (\bar{f}_i'(w_i) x_i)^2\big)$$
$$+ O(i^{-2}\xi_i^2 x_i^2) + O\big(t_i \zeta_i (\bar{g}_i'(t_i) y_i)^2\big) + O(i^{-2}\rho_i^2 y_i^2),$$

where the errors come from the interval in (5.2), $(0, i^{-1}\xi_i x_i)$, the interval in (5.3) and $(-i^{-1}\rho_i y_i, 0)$, respectively. Recalling the conditions on w_i, t_i, we have

$$n_i E(X(n_i,0,0))^2 = \bar{\tau}_j^2 n_i x_i^2 \xi_i + o(n_i x_i^2 \xi_i) + o(n_i y_i^2 \rho_i).$$

Since $n_i \xi_i = r_{n_i} = (n_i \eta_i)^2$, $n_i \rho_i = s_{n_i} = (n_i \zeta_i)^2$ and $n_i \zeta_i y_i = s_{n_i}^{1/2} y_i = r_{n_i}^{1/2} x_i = n_i \eta_i x_i$, this yields

$$n_i E(X(n_i,0,0))^2 = \bar{\tau}_j^2 \gamma_{n_i}^2 + o(\gamma_{n_i}^2).$$

In estimating $n_i(EX(n_i,0,0))^2$, the squares of terms corresponding to the four error terms above will all be no larger than the error terms above so we only need to consider the square of the term coming from the atom,

$$n_i\big(i^{-1}x_i \bar{\tau}_j^2 i^2 \xi_i\big)^2 = \bar{\tau}_j^4 i^2 \xi_i n_i x_i^2 \xi_i = \bar{\tau}_j^4 i^2 \xi_i \gamma_{n_i}^2,$$

and by the defining conditions $\bar{\tau}_j^4 i^2 \xi_i \le i^6((i-1)!)^{-5} = o(1)$ so this term does not contribute. Therefore

$$\tau_{n_i}^2 = \bar{\tau}_j^2 \gamma_{n_i}^2 + o(\gamma_{n_i}^2).$$

The condition on μ_{n_i} in (3.10) is arranged by choice of δ_{n_i}: set $\delta_{n_i} = \mu_{n_i} - \bar{\mu}_i \gamma_{n_i}$. Therefore we have that any distribution for Z corresponding to some $(\bar{\tau}_i, \bar{f}_i, \bar{g}_i, \bar{\mu}_i)$ is a subsequential limit by Theorem 1. For a general Z given by (τ, f, g, μ) take a subsequence from the dense family such that

$$\bar{\tau}_{m_i} \to \tau, \qquad \bar{f}_{m_i} \to f, \qquad \bar{g}_{m_i} \to g, \qquad \bar{\mu}_{m_i} \to \mu.$$

As we have seen we can find n_{ik} such that $\gamma_{n_{ik}}^{-1} \tau_{n_{ik}} \to \bar{\tau}_{m_i}$, and so on. By taking k large enough here, we find n_i such that $|\gamma_{n_i}^{-1}\tau_{n_i} - \bar{\tau}_{m_i}| \le i^{-1}$ and so $\gamma_{n_i}^{-1}\tau_{n_i} \to \tau$. Since the f's and g's are converging uniformly on compacts the same method works there also. Thus Z is a subsequential limit by Theorem 1. □

6. Examples. We start with the example described in the Introduction.

EXAMPLE 1. For $\rho \in (0, \infty)$ let

$$P\{X < -x\} = P\{X > x\} = \tfrac{1}{2}(\log x)^{-\rho}, \qquad x \geq e.$$

We will prove that $S_n(s_n, r_n)$ is asymptotically normal iff

$$\lim_{n\to\infty} n^{2/(2+\rho)} r_n^{-1} = 0 \quad \text{and} \quad \lim_{n\to\infty} n^{2/(2+\rho)} s_n^{-1} = 0. \tag{6.1}$$

(This condition also applies to any subsequence.) There exist $\{\delta_n\}, \{\gamma_n\}$ such that $\gamma_n^{-1}(S_n(s_n, r_n) - \delta_n)$ is stochastically compact iff

$$\limsup_{n\to\infty} n^{2/(2+\rho)} r_n^{-1} < \infty \quad \text{and} \quad \limsup_{n\to\infty} n^{2/(2+\rho)} s_n^{-1} < \infty. \tag{6.2}$$

Finally, we will show that in order to have a nonnormal subsequential limit, one must have

$$\big(b_{n_i}(0)\big)^{-1} a_{n_i}(0) \to \xi, \qquad n_i^{-2/(2+\rho)} r_{n_i} \to c, \qquad n_i^{-2/(2+\rho)} s_{n_i} \to d, \tag{6.3}$$

where $0 \leq \xi \leq \infty$, and

$$\begin{aligned} &0 < c = d < \infty, && \text{if } 0 < \xi < \infty,\\ &0 < c < \infty,\ c \leq d \leq \infty, && \text{if } \xi = 0,\\ &0 < d < \infty,\ d \leq c \leq \infty, && \text{if } \xi = \infty. \end{aligned}$$

If $\xi = 0$, one actually only needs $\liminf n_i^{-2/(2+\rho)} s_{n_i} \geq c$, and, if $\xi = \infty$, $\liminf n_i^{-2/(2+\rho)} r_{n_i} \geq d$. The subsequential limit is then

$$\begin{aligned} h_c(N_2) - \xi h_c(N_3), &\quad \text{if } \xi < \infty,\\ -h_d(N_3), &\quad \text{if } \xi = \infty, \end{aligned} \tag{6.4}$$

where N_2, N_3 are independent $N(0,1)$ and

$$h_c(x) = \exp\big\{x\rho^{-1}(2c)^{-1/\rho} c^{-1/2}\big\}. \tag{6.5}$$

(Of course, one may also have scale and translation changes of these.)

For any distribution with slowly varying tails (see, e.g., Lemma 2.5 of [15])

$$n\sigma_n^2(\alpha, \beta) \sim n\lambda_n(\alpha, \beta) \sim r_n b_n^2(\beta) + s_n a_n^2(\alpha). \tag{6.6}$$

Next we have

$$\begin{aligned} b_n(\beta) &= \exp\left\{\left(\frac{n}{2r_n(1 - \beta r_n^{-1/2})}\right)^{1/\rho}\right\},\\ a_n(\alpha) &= \exp\left\{\left(\frac{n}{2s_n(1 - \alpha s_n^{-1/2})}\right)^{1/\rho}\right\}. \end{aligned} \tag{6.7}$$

Since

$$\big(1 - \beta r_n^{-1/2}\big)^{-1/\rho} = 1 + \beta\rho^{-1} r_n^{-1/2} + O\big(r_n^{-1}\big), \tag{6.8}$$

we see that

$$b_n(\beta) \sim b_n(0) \quad \text{iff} \quad \lim_{n\to\infty} n^{2/(2+\rho)} r_n^{-1} = 0. \tag{6.9}$$

Thus the condition (6.1) implies $\lambda_n(\alpha, \beta) \sim \lambda_n(0,0)$ and hence (by Theorem 3) asymptotic normality of $S_n(s_n, r_n)$. For the converse, suppose $\lambda_n(\alpha, \beta) \sim \lambda_n(0,0)$ and restrict attention to those n for which $r_n \le s_n$. For large n this will mean $r_n b_n^2(0) \ge s_n a_n^2(0)$ and then by (6.6)

$$1 = \lim_{n \to \infty} \frac{r_n b_n^2(\beta) + s_n a_n^2(0)}{r_n b_n^2(0) + s_n a_n^2(0)},$$

which implies $b_n(\beta) \sim b_n(0)$ and then (6.1) by (6.9). The argument is similar for those n where s_n is smaller. This proves the first assertion. A similar argument using (2.25) shows that there exist $\{\delta_n\}, \{\gamma_n\}$ such that $\gamma_n^{-1}\{S_n(s_n, r_n) - \delta_n\}$ is stochastically compact iff (6.2) holds. Thus it remains to examine the subsequential limits using Theorem 2 when $r_n \wedge s_n$ is comparable to $n^{2/(2+\rho)}$. We consider a subsequence for which (6.3) holds. We do not allow c or d to be 0 due to (6.2) and we do not allow them both to be ∞ due to (6.1). Note that $0 < \xi < \infty$ implies that $r_{n_i} \sim s_{n_i}$ and so $c = d$. Recalling (6.7) and (6.8), we have $b_{n_i}(\beta) \sim b_{n_i}(0) h_c(\beta)$ with $h_c(\beta)$ as in (6.5). Then

$$\begin{aligned} f_{n_i}(\beta) &= n_i \int_{b_{n_i}(0)}^{b_{n_i}(\beta)} \frac{\rho}{2} \frac{x}{(\log x)^{\rho+1} x}\, dx \\ &\sim \frac{\rho}{2} n_i \left(n_i^{-1} 2 r_{n_i}\right)^{(\rho+1)/\rho} \left(b_{n_i}(\beta) - b_{n_i}(0)\right) \qquad (6.10) \\ &\sim \rho 2^{1/\rho} c^{(\rho+2)/2\rho} r_{n_i}^{1/2} b_{n_i}(0) \left(h_c(\beta) - 1\right). \end{aligned}$$

If $0 < \xi < \infty$, since $r_{n_i} \sim s_{n_i}$ we have by (6.6)

$$n_i \sigma_{n_i}^2(0,0) \sim r_{n_i} b_{n_i}^2(0)(1 + \xi^2). \qquad (6.11)$$

This is also valid if $\xi = 0$. To see this, let $H(x) = x^{-\rho/2} e^x$. H is ultimately increasing and

$$H\left(x - \log(2\varepsilon^{-1})\right) \sim \varepsilon H(x)/2 \quad \text{as } x \to \infty.$$

Since $\xi = 0$ we have

$$\left(n_i / 2 s_{n_i}\right)^{1/\rho} - \left(n_i / 2 r_{n_i}\right)^{1/\rho} < -\log(2\varepsilon^{-1})$$

for large i and then

$$\begin{aligned} n_i^{-1/2} \left(2 s_{n_i}\right)^{1/2} a_{n_i}(0) &= H\left(\left(n_i / 2 s_{n_i}\right)^{1/\rho}\right) \\ &< H\left(\left(n_i / 2 r_{n_i}\right)^{1/\rho} - \log(2\varepsilon^{-1})\right) \\ &< \varepsilon H\left(\left(n_i / 2 r_{n_i}\right)^{1/\rho}\right) \\ &= \varepsilon n_i^{-1/2} \left(2 r_{n_i}\right)^{1/2} b_{n_i}(0). \end{aligned}$$

This proves that for $\xi = 0$, $s_{n_i} a_{n_i}^2(0) = o(r_{n_i} b_{n_i}^2(0))$ and hence (6.11) holds. Assume for now that $\xi < \infty$. Then by (6.10) and (6.11)

$$\left(n_i \sigma_{n_i}^2(0,0)\right)^{-1/2} f_{n_i}(\beta) \to \rho 2^{1/\rho} c^{(\rho+2)/2\rho} (1 + \xi^2)^{-1/2} \left(h_c(\beta) - 1\right). \qquad (6.12)$$

A similar argument shows that

$$(6.13)\quad \left(n_i\sigma_{n_i}^2(0,0)\right)^{-1/2} g_{n_i}(\alpha) \to \rho 2^{1/\rho} d^{(\rho+2)/2\rho}\xi(1+\xi^2)^{-1/2}(h_d(\alpha)-1).$$

(The coefficient on the right is to be 0 if $d=\infty$.) If $\xi>0$, then $c=d$ so we let

$$(6.14)\qquad f(\beta)=h_c(\beta)-1,\qquad g(\alpha)=\xi(h_c(\alpha)-1).$$

Then

$$f'(0)=\xi^{-1}g'(0)=\rho^{-1}(2c)^{-1/\rho}c^{-1/2}$$

and so

$$(6.15)\qquad \left((g'(0))^2+(f'(0))^2\right)^{1/2}=(1+\xi^2)^{1/2}\rho^{-1}(2c)^{-1/\rho}c^{-1/2}.$$

Recalling (6.12)–(6.14), we see that (3.18) and (3.19) hold with $\tau=0$. Then (3.17) follows (with $\tau=0$) by using Lemma 2.5 of [15] again. For the norming sequence, we have by (3.20), (6.11) and (6.15)

$$(6.16)\qquad \begin{aligned}\gamma_{n_i} &\sim r_{n_i}^{1/2} b_{n_i}(0)\rho(2c)^{1/\rho}c^{1/2}\\ &\sim \rho(2c)^{1/\rho} c n_i^{1/(2+\rho)}\exp\left\{\left(n_i/2r_{n_i}\right)^{1/\rho}\right\},\end{aligned}$$

and for the centering sequence

$$\begin{aligned}\mu_n &= n\int_e^{b_n(0)}\frac{\rho}{2}\frac{dx}{(\log x)^{\rho+1}} - n\int_{-a_n(0)}^{-e}\frac{\rho}{2}\frac{dx}{(\log|x|)^{\rho+1}}\\ &= \frac{\rho}{2}n\frac{b_n(0)}{(\log b_n(0))^{\rho+1}} + O\left(\frac{nb_n(0)}{(\log b_n(0))^{\rho+2}}\right)\\ &\quad -\frac{\rho}{2}n\frac{a_n(0)}{(\log a_n(0))^{\rho+1}} + O\left(\frac{na_n(0)}{(\log a_n(0))^{\rho+2}}\right),\end{aligned}$$

which leads to

$$\begin{aligned}\mu_{n_i} &\sim \rho 2^{1/\rho}c^{(\rho+2)/2\rho} r_{n_i}^{1/2} b_{n_i}(0) - \rho 2^{1/\rho} d^{(\rho+2)/2\rho} s_{n_i}^{1/2} a_{n_i}(0)\\ &\sim \rho 2^{1/\rho} c^{(\rho+2)/2\rho} r_{n_i}^{1/2} b_{n_i}(0)(1-\xi) \sim \gamma_{n_i}(1-\xi).\end{aligned}$$

(This is valid even if $d=\infty$.) By (3.20) if we take $\delta_{n_i}=0$, then $\mu=1-\xi$. Thus we obtain (6.4) for the case $\xi<\infty$ with $\delta_{n_i}=0$ and γ_{n_i} as in (6.16). The case $\xi=\infty$ is clearly dual to $\xi=0$ and we obtain the final result with

$$\gamma_{n_i}=\rho(2d)^{1/\rho} d n_i^{1/(2+\rho)}\exp\left\{\left(n_i/2s_{n_i}\right)^{1/\rho}\right\}.$$

Next we give the example mentioned in the Introduction where asymptotic normality fails for $S_n(r_n, r_n)$ but holds for ${}^{(2r_n)}S_n$.

EXAMPLE 2. Let X have a symmetric distribution with density 2^{-k-1} on $[2^k-1, 2^k]$, $k=1,2,\dots$, and no mass elsewhere. Assume, for simplicity, $r_n\le s_n$ for all n but this is unimportant. We will show that $S_n(s_n, r_n)$ cannot be asymptotically normal for any nondecreasing sequence $\{r_n\}$. [It can be for a general sequence $\{r_n\}$ satisfying (1.1).] Define

$$n_k = \max\{n: n^{-1}r_n > 2^{-k}\}.$$

Since $P\{X > 2^{k-1}\} = 2^{-k}$ we have $b_{n_k}(0) < 2^{k-1}$ and then $s_n \geq r_n$ yields $a_{n_k}(0) < 2^{k-1}$ also. Thus

$$\lambda_{n_k}(0,0) \leq E(|X| \wedge 2^{k-1})^2 \leq 2\sum_{j=1}^{k-1} 2^{2j}2^{-j-1} + 2 \cdot 2^{2k-2}2^{-k} \tag{6.17}$$
$$\leq 3 \cdot 2^{k-1}.$$

Next $(n_k + 1)^{-1}r_{n_k+1} \leq 2^{-k} < n_k^{-1}r_{n_k}$ implies $r_{n_k+1} \leq r_{n_k}$ and equality holds since we assumed $\{r_n\}$ is nondecreasing. Thus

$$\frac{r_{n_k}}{n_k} = \frac{r_{n_k+1}}{n_k+1}\frac{n_k+1}{n_k} \leq \frac{1}{2^k} + \frac{r_{n_k}}{n_k(n_k+1)}$$
$$= \frac{1}{2^k} + o\left(n_k^{-1}r_{n_k}^{1/2}\right)$$

and so

$$\frac{r_{n_k} - r_{n_k}^{1/2}}{n_k} \leq \frac{1}{2^k} - \frac{r_{n_k}^{1/2}}{n_k}(1 + o(1)) < \frac{1}{2^k}$$

for large k. Since $P\{X > 2^k - 1\} = 2^{-k}$ we must have $b_{n_k}(1) > 2^k - 1$. Therefore

$$\lambda_{n_k}(0,1) - \lambda_{n_k}(0,0) \geq \left(b_{n_k}^2(1) - b_{n_k}^2(0)\right)P\{X \geq b_{n_k}(1)\}$$
$$\geq \left((2^k-1)^2 - 2^{2k-2}\right)P\{X \geq b_{n_k}(1)\}$$
$$\sim 3 \cdot 2^{k-2}$$

and, recalling (6.17), we have

$$\liminf_k \left(\lambda_{n_k}(0,0)\right)^{-1}\lambda_{n_k}(0,1) - 1 \geq 2^{-1}$$

so that asymptotic normality must fail by Theorem 3. One will obtain asymptotic normality for this example if one has $\{r_n\}$, $\{s_n\}$ sequences where the values of $n^{-1}r_n$ and $n^{-1}s_n$ stay relatively far away from the sequence 2^{-k}. This cannot happen with monotone sequences. For the problem concerning the trimmed sum ${}^{(r_n)}S_n$, where the trimming is in terms of the largest in absolute value, the condition (1.7) is not sensitive to the sudden change from $b_n(0)$ to $b_n(1)$. The reader may check that (1.7) is valid for this example for any sequence $\{r_n\}$ satisfying (1.1); alternatively, it is an immediate consequence of Corollary 3.12 of [9].

We conclude with the proof of the statement in the Introduction that if X has a symmetric distribution and $S_n(r_n, r_n)$ is asymptotically normal, then so is ${}^{(2r_n)}S_n$. [We assume that X has a continuous distribution so we may use (1.7) but this is not important.] For (1.6) we have

$$2P\{X > c_n(\alpha)\} = n^{-1}\left(2r_n - \alpha(2r_n)^{1/2}\right)$$
$$= 2n^{-1}\left(r_n - \alpha 2^{-1/2} r_n^{1/2}\right)$$

so that $c_n(\alpha) = b_n(\alpha 2^{-1/2}) = a_n(\alpha 2^{-1/2})$. For $\alpha > 0$, consider

$$EX^2 1\{|X| \le c_n(\alpha)\} - EX^2 1\{|X| \le c_n(0)\} \le c_n^2(\alpha) P\{c_n(0) < |X| \le c_n(\alpha)\}$$
$$= \alpha n^{-1}(2r_n)^{1/2} c_n^2(\alpha).$$

Thus for (1.7) we need to prove that

$$n^{-1} r_n^{1/2} c_n^2(\alpha) = o\left(EX^2 1\{|X| \le c_n(\alpha)\}\right). \tag{6.18}$$

Consider a subsequence for which

$$\lim_{n\to\infty} c_n^2(\alpha) P\{X > c_n(\alpha)\}\left(\lambda_n^+(\alpha 2^{-1/2})\right)^{-1} = c \quad (\le 1). \tag{6.19}$$

If $c = 0$, then (6.18) holds along the subsequence since

$$n^{-1} r_n^{1/2} c_n^2(\alpha) = o\left(c_n^2(\alpha) P\{X > c_n(\alpha)\}\right)$$
$$= o\left(\lambda_n^+(\alpha 2^{-1/2})\right)$$
$$= o\left(EX^2 1\{|X| \le c_n(\alpha)\}\right).$$

Thus we may suppose $c > 0$. Then for $\alpha > 0$, by (6.19)

$$\lambda_n^+(\alpha 2^{-1/2}) - \lambda_n^+(-\alpha 2^{-1/2}) \ge P\{X > c_n(\alpha)\}\left(c_n^2(\alpha) - c_n^2(-\alpha)\right)$$
$$\approx \lambda_n^+(\alpha 2^{-1/2})\left(c_n(\alpha)\right)^{-2}\left(c_n^2(\alpha) - c_n^2(-\alpha)\right).$$

Since (1.4) implies $\lambda_n^+(\alpha) \sim \lambda_n^+(0)$ for all α in the symmetric case, we have $c_n(\alpha) \sim c_n(-\alpha)$ along the subsequence. But then, for any $\xi > 0$,

$$EX^2 1\{|X| \le c_n(\alpha)\} \ge EX^2 1\{c_n(-\xi) < |X| \le c_n(\alpha)\}$$
$$\ge c_n^2(-\xi) P\{c_n(-\xi) < |X| \le c_n(\alpha)\}$$
$$= c_n^2(-\xi)(\xi + \alpha) n^{-1}(2r_n)^{1/2}$$

and so for large n,

$$n^{-1} r_n^{1/2} c_n^2(\alpha) \le (\xi + \alpha)^{-1} c_n^2(\alpha)\left(c_n(-\xi)\right)^{-2} EX^2 1\{|X| \le c_n(\alpha)\}$$
$$\sim (\xi + \alpha)^{-1} EX^2 1\{|X| \le c_n(\alpha)\},$$

which again implies (6.18) since we are free to choose ξ large. This proves the result. The asymptotic variance for ${}^{(2r_n)}S_n$ is $nEX^2 1\{|X| \le c_n(0)\}$ which is smaller than $n\lambda_n(0, 0)$ as indicated and even of smaller order with slowly varying tails.

7. Statistical implications.

In these remarks we will assume that $E|X| < \infty$ and that the statistician is interested in estimating EX. Perhaps the first subject to consider is consistency. This may fail if heavy trimming is used (r_n, s_n proportional to n) as then the trimmed mean will converge to a conditional population mean. This is not a problem when intermediate trimming is used [i.e., (1.1) is satisfied] and here one may prove

$$(n - r_n - s_n)^{-1} S_n(s_n, r_n) \to EX \quad \text{a.s.}$$

The next question is concerned with the implications of Theorem 3 for interval estimation. The existence of a limiting normal distribution under such mild conditions appears to be quite useful. But one must examine more carefully the centering and norming sequences $\{\delta_n\}, \{\gamma_n\}$. We start with $\{\gamma_n\}$. It is, of course, undesirable that this depends heavily on the underlying distribution. However, this defect is easily overcome by replacing γ_n by its sample version. We now prove this fact. Define $M_n = (n - r_n - s_n)^{-1} S_n(s_n, r_n)$ and V_n by

$$V_n^2 = \sum_{k=s_n+1}^{n-r_n} (X_{nk} - M_n)^2 + s_n(X_{n,s_n})^2 + r_n(X_{n,n-r_n+1})^2.$$

Then we have the following theorem.

Theorem 6. *Assume* (1.1) *and* (4.2) *and let* $\gamma_n = \{n\sigma_n^2(0,0)\}^{1/2}$. *Then*

$$\frac{V_n}{\gamma_n} \xrightarrow{P} 1 \tag{7.1}$$

and there exists $\{\delta_n\}$ *such that*

$$V_n^{-1}\{S_n(s_n, r_n) - \delta_n\} \to N(0,1). \tag{7.2}$$

Remark. One may replace X_{n,s_n} by $X_{n,s_n} - M_n$ and make a similar change for $X_{n,n-r_n+1}$ in the definition of V_n without changing the results. We chose the simpler although perhaps less aesthetic version.

Proof. It is enough to prove (7.1) since the rest follows from Theorem 3. We condition on $Y_{n,s_n} = v_n(\alpha)$ and $Y_{n,n-r_n+1} = u_n(\beta)$ as in the proof of Theorem 1 and use $E_{\alpha\beta}$ and $P_{\alpha\beta}$ to denote conditional expectations and probabilities. Then

$$E_{\alpha\beta}V_n^2 = (m-1)\operatorname{Var}(\hat{X}(n,\alpha,\beta)) + s_n a_n^2(\alpha) + r_n b_n^2(\beta) \sim n\sigma_n^2(\alpha,\beta)$$

and a straightforward computation (or recalling the standard formula for the variance of the sample variance; see, e.g., Wilks [19], page 199) and (2.4) lead to

$$\begin{aligned} E_{\alpha\beta}\left(V_n^2 - E_{\alpha\beta}V_n^2\right)^2 &\le (m-1)E\left(\hat{X}(n,\alpha,\beta) - E\hat{X}(n,\alpha,\beta)\right)^4 \\ &\le n(b_n(\beta) + a_n(\alpha))^2 \operatorname{Var}(\hat{X}(n,\alpha,\beta)) \\ &= O\left(n(r_n^{-1} + s_n^{-1}) n\lambda_n(\alpha,\beta)\lambda_n(\alpha,\beta)\right) \\ &= o\left(n^2\lambda_n^2(\alpha,\beta)\right). \end{aligned}$$

Using Chebsyhev's inequality, we have

$$P_{\alpha\beta}\{|V_n^2 - E_{\alpha\beta}V_n^2| \geq \varepsilon\gamma_n^2\} \leq \varepsilon^{-2}\gamma_n^{-4}E_{\alpha\beta}\left(V_n^2 - E_{\alpha\beta}V_n^2\right)^2 = o\left(\frac{\lambda_n^2(\alpha,\beta)}{\sigma_n^4(0,0)}\right)$$

and this goes to 0 by (4.2) and (2.7). Using (4.2) again, we see that $\gamma_n^{-2}V_n^2 \to 1$ in probability under $P_{\alpha\beta}$. By using the monotonicity of $a_n(\alpha)$, $b_n(\beta)$, we see that this convergence is uniform for α, β in compacts. Finally, as in (3.12),

$$P\{|\gamma_n^{-2}V_n^2 - 1| \geq \varepsilon\} = \int\int P_{\alpha\beta}\{|\gamma_n^{-2}V_n^2 - 1| \geq \varepsilon\}\nu_n(d\alpha, d\beta)$$

and so this approaches 0 as well. □

It remains to examine $\{\delta_n\}$, where $\delta_n = \mu_n$ as given in (3.1). In the context of interval estimation, one would *like* to replace (7.2) by

$$V_n^{-1}\{S_n(s_n, r_n) - (n - r_n - s_n)EX\} \to N(0,1).$$

But this is typically false. The question is whether

$$\gamma_n^{-1}\{\delta_n - (n - r_n - s_n)EX\} \to 0 \tag{7.3}$$

and this turns out to be quite delicate—even when $EX^2 < \infty$! We give some examples to show what can happen. If the distribution is symmetric about its unknown mean and one takes $r_n = s_n$, then $\delta_n = (n - r_n - s_n)EX$ as desired. But even in this simple situation taking $r_n \neq s_n$ may cause (7.3) to fail. As an example, if the distribution is double exponential, centered at $\mu = EX$, then

$$\delta_n - (n - r_n - s_n)\mu = s_n \log\frac{ne}{2s_n} - r_n \log\frac{ne}{2r_n}.$$

Thus (7.3) holds whenever $r_n, s_n = o(n^{1/2}/\log n)$ but for larger r_n even taking $s_n \sim r_n$ is not enough. Unless r_n is large enough that $\log(n/r_n)$ is much smaller than $\log n$, one needs $s_n - r_n = o(n^{1/2}/\log n)$. Of course, one might feel that taking $r_n = s_n$ is a safe thing to do. But what if we consider the same example with the positive tail having twice as much mass:

$$F(x) = \begin{cases} \frac{1}{3}e^{x-\mu}, & x \leq \mu, \\ 1 - \frac{2}{3}e^{-(x-\mu)}, & x > \mu. \end{cases}$$

Now should one take $s_n = 2r_n$? In this case $EX = \mu + \frac{1}{3}$ and

$$\delta_n - (n - r_n - s_n)\left(\mu + \tfrac{1}{3}\right) = s_n \log\frac{ne^{4/3}}{3s_n} - r_n \log\frac{2ne^{2/3}}{3r_n}$$

so one still needs $s_n \sim r_n$ when r_n is comparable to $n^{1/2}/\log n$ or larger. But more is required whenever r_n is larger than $n^{1/2}/\log n$. For example, if $r_n = [n^{1/2}]$, then one must take

$$s_n = n^{1/2} + \left(\log 4 - \tfrac{4}{3}\right)n^{1/2}/\log n + o\left(n^{1/2}/\log n\right)$$

in order for (7.3) to hold. As a final example, consider a stable law of index

$\rho \in (1, 2)$ with mean μ and asymmetry parameter

$$\tau = \lim_{x \to \infty} \frac{P\{X > x\}}{P\{X < -x\}}.$$

Now r_n and s_n must be chosen carefully even when they are relatively small. If $r_n = o(n^{2/3})$, then one must take

$$s_n = \tau^{1/(\rho-1)} r_n + o\left(r_n^{1/2}\right). \tag{7.4}$$

Thus, unlike the exponential example, if one puts twice as much mass in the positive tail, then it is no longer correct to take $s_n \sim r_n$. In fact, the ratio of s_n to r_n must depend on the index of the stable law. If r_n is as large as $n^{2/3}$, then (7.4) is not good enough. If $r_n = o(n^{4/5})$, then one must take

$$s_n = \tau^{1/(\rho-1)} r_n + K n^{-1} r_n^2 + o\left(r_n^{1/2}\right),$$

where K depends on the coefficients of the second terms in the expansions of the stable density near $\pm\infty$ as well as the leading terms. The situation becomes worse as r_n gets larger; if $r_n = o(n^{2k/(2k+1)})$, then one must take

$$s_n = \tau^{1/(\rho-1)} r_n \left(1 + \sum_{j=1}^{k-1} K_j \left(n^{-1} r_n\right)^j\right) + o\left(r_n^{1/2}\right),$$

for appropriate K_j.

We conclude this section with a positive result followed by some general remarks on how to proceed. If the distribution for X and $\{r_n\}$ are given, then it is always possible to find $\{s_n\}$ so that (7.3) is true. But this requires detailed knowledge of the distribution. If one is willing to assume that the distribution is symmetric about the unknown mean, then one should take $s_n = r_n$. The exponential example suggests that there is also some advantage in taking r_n so that it grows rather slowly as this provides some protection against asymmetric tails if they are thin enough. This is true if the variance is finite but one must make r_n grow more slowly as the tails get thicker to preserve this advantage. The stable examples suggest that making r_n grow slowly does not help if the variance is infinite. In general, the examples seem to convey the message that it would be difficult to devise a statistical procedure that would tell the statistician how to choose the sequences $\{r_n\}, \{s_n\}$ so that (7.3) holds.

Acknowledgments. We take this opportunity to thank the referees for their careful reading of the manuscript and several comments and suggestions which improved the exposition.

REFERENCES

[1] Arov, D. Z. and Bobrov, A. A. (1960). The extreme terms of a sample and their role in the sum of independent variables. *Theory Probab. Appl.* **5** 377–396.

[2] Billingsley, P. (1979). *Probability and Measure*. Wiley, New York.

[3] Csörgő, M., Csörgő, S., Horváth, L. and Mason, D. M. (1986). Normal and stable convergence of integral functions of the empirical distribution function. *Ann. Probab.* **14** 86–118.

[4] Csörgő, S., Haeusler, E. and Mason, D. M. (1988). The asymptotic distribution of trimmed sums. *Ann. Probab.* **16** 672–699.

[5] Csörgő, S., Horváth, L. and Mason, D. M. (1986). What portion of the sample makes a partial sum asymptotically stable or normal? *Probab. Theory Related Fields* **72** 1–16.

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

[7] Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading, Mass.

[8] Griffin, P. S. (1988). Asymptotic normality of Winsorized means. *Stochastic Process. Appl.* **29** 107–127.

[9] Griffin, P. S. and Pruitt, W. E. (1987). The central limit problem for trimmed sums. *Math. Proc. Cambridge Philos. Soc.* **102** 329–349.

[10] Hahn, M. and Kuelbs, J. (1989). Universal asymptotic normality for conditionally trimmed sums. *Statist. Probab. Lett.* **7** 9–15.

[11] Kuelbs, J. and Ledoux, M. (1987). Extreme values and a Gaussian central limit theorem. *Probab. Theory Related Fields* **74** 341–355.

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

[13] Mori, T. (1976). The strong law of large numbers when extreme values are excluded from sums. *Z. Wahrsch. verw. Gebiete* **36** 189–194.

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

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

[16] Steutel, F. W. (1974). On the tails of infinitely divisible distributions. *Z. Wahrsch. verw. Gebiete* **28** 273–276.

[17] Stigler, S. M. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1** 472–477.

[18] Weiner, D. C. (1986). Centrality, scale and asymptotic normality for sums of independent random variables. Preprint.

[19] Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York.

Department of Mathematics
Syracuse University
Syracuse, New York 13244-1150

School of Mathematics
University of Minnesota
Minneapolis, Minnesota 55455