The Annals of Probability
1978, Vol. 6, No. 2, 294–315

A WEAK INVARIANCE PRINCIPLE WITH APPLICATIONS TO DOMAINS OF ATTRACTION

BY GORDON SIMONS[1] AND WILLIAM STOUT[2]

University of North Carolina and University of Illinois

An elementary probabilistic argument is given which establishes a "weak invariance principle" which in turn implies the sufficiency of the classical assumptions associated with the weak convergence of normed sums to stable laws. The argument, which uses quantile functions (the inverses of distribution functions), exploits the fact that two random variables $X = F^{-1}(U)$ and $Y = G^{-1}(U)$ are, in a useful sense, close together when F and G are, in a certain sense, close together. Here U denotes a uniform variable on (0, 1). By-products of the research are two alternative characterizations for a random variable being in the domain of partial attraction to a normal law and some results concerning the study of domains of partial attraction.

1. Introduction. This paper presents a new method for establishing convergence in law of normed sums of independent identically distributed (i.i.d.) random variables. A distinctive feature of this method is that no use of transforms is made (except the distributional form of stable laws is taken for granted). The method is probabilistic. It depends completely on the establishment of an appropriate invariance principle. The invariance principle introduced here is analogous to the almost sure invariance principles appearing in the literature (cf., Strassen (1964, 1965), Csörgö and Révész (1975), Komlós, Major and Tusnády (1975) and Philipp and Stout (1975)) except almost sure convergence is replaced by convergence in probability or, what is equivalent in this setting, convergence in law. For this reason, the terminology "weak invariance principle" will be used. The weak invariance principle encapsulated in Theorem 2 below is based upon very elementary notions of probability. This contrasts sharply with the typical almost sure invariance principle, which depends on the existence and properties of Brownian motion and frequently on some form of Skorokhod embedding. Our weak invariance principle is designed to demonstrate the sufficiency of the classical assumptions associated with the weak convergence of normed sums to stable laws. Perhaps more sophisticated weak invariance principles can be found which would apply more widely within the scope of the central limit problem.

Received December 22, 1975; revised September 9, 1976.

[1] The work of this author was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-75-2796 and the National Science Foundation under Grant No. MCS-75-07556.

[2] The work of this author was supported by the National Science Foundation.

AMS 1970 *subject classifications.* Primary 60F05, 60G50.

Key words and phrases. Invariance principle, domain of attraction, domain of partial attraction, stable distribution, central limit problem, domain of normal attraction, slowly varying function, quantile function.

Our main result, Theorem 2, is concerned with random variables which are within the domain of partial attraction of some infinitely divisible law (implied by assumption A2 below) but outside the domain of partial attraction of a normal law (i.e., satisfying assumption A1). The latter restriction is essential to our approach. It implies that our theory does not concern itself with the convergence of normed sums to the normal law. But it does apply to all other stable laws. In this regard, it should be pointed out that Root and Rubin (1973) have already established the normal convergence criterion by probabilistic methods. The effect of our work is to bring probabilistic methods to another area of the general central limit problem.[3]

By-products of our research are two alternative characterizations (see A1$''$ and A1$'''$) of the class of distributions which are within the domain of partial attraction of a normal law. We believe these are new. They may be of independent interest. We also obtain some results about domains of partial attraction.

In order to introduce our method, unencumbered by technical details, we present a special case in Section 2. Sections 3 and 4 present the method in fuller generality. Section 5 discusses the assumptions being made and their significance in more detail than what seems to be appropriate for Section 3.

2. Domain of normal attraction of the Cauchy law. The following theorem describes a very specialized weak invariance principle but, at the same time, with a minimum of computations, yields an interesting convergence in law result.

Let X be a Cauchy random variable (centered at zero and) scaled so that

$$\Pr(|X| > x) \sim \frac{1}{x} \qquad \text{as} \quad x \to \infty ,$$

and Y be a symmetric random variable satisfying

$$\Pr(|Y| > x) \sim \frac{1}{x} \qquad \text{as} \quad x \to \infty .$$

THEOREM 1. *On some probability space, there exist two sequences of* i.i.d. *random variables* $X_1, X_2, \cdots$ *and* $Y_1, Y_2, \cdots$ *whose members are distributed as* X *and* Y, *respectively, and which satisfy*

$$\sum_{i=1}^n Y_i = \sum_{i=1}^n X_i + o_p(n) \qquad as \quad n \to \infty . \tag{1}$$ [4]

Since $n^{-1} \sum_{i=1}^n X_i$ is distributed as X, the following corollary is immediate.

COROLLARY 1. *If* $Y_1, Y_2, \cdots$ *are* i.i.d. *and distributed as* Y, *then*

$$n^{-1} \sum_{i=1}^n Y_i \to_L X .$$

[3] After the acceptance of this paper, the authors became aware of a relevant preprint by P. Major, entitled An improvement of Strassen's invariance principle. It gives a probabilistic proof of the central limit theorem in the case of a normal limit using quantile functions.

[4] The notation $Z_n = o_p(c_n)$ means $Z_n/c_n \to 0$ in probability.

PROOF OF THEOREM 1. Let $U_1, U_2, \cdots$ be i.i.d. uniform variables on (0, 1), and let F and G denote the (right continuous) distribution functions of X and Y respectively. Set

$$X_i = F^{-1}(U_i), \qquad Y_i = G^{-1}(U_i), \quad i \geqq 1 .$$

Here F^{-1} is defined formally by

$$F^{-1}(u) = \min\{x : F(x) \geqq u\}, \quad u \in (0, 1) . \tag{2}$$

Since

$$F^{-1}(u) \leqq x \qquad \text{iff} \quad u \leqq F(x) , \tag{3}$$

the distribution function of X_1 is F. Let

$$Z_i = Y_i - X_i \quad (i \geqq 1) \qquad \text{and} \qquad Z = Z_1 .$$

Equation (1) is identical to $n^{-1} \sum_{i=1}^n Z_i \to_p 0$ as $n \to \infty$, which, by the classical degenerate convergence theorem, is equivalent (since Z is symmetric) to

$$n \Pr(|Z| > n) \to 0 \qquad \text{as} \quad n \to \infty ; \tag{4}$$

$$n^{-1} \operatorname{Var} ZI(|Z| \leqq n) \to 0 \qquad \text{as} \quad n \to \infty . \tag{5}$$

Since $\operatorname{Var} ZI(|Z| \leqq n) \leqq 2 \int_0^n v \Pr(|Z| > v)\, dv$ (see (10) below), condition (4) implies condition (5).

It is almost immediate that

$$F^{-1}(u) = \frac{1 + o(1)}{2(1 - u)}, \qquad G^{-1}(u) = \frac{1 + o(1)}{2(1 - u)} \quad \text{as} \quad u \to 1 ,$$

and

$$F^{-1}(u) = -\frac{1 + o(1)}{2u}, \qquad G^{-1}(u) = -\frac{1 + o(1)}{2u} \quad \text{as} \quad u \to 0 .$$

Hence,

$$\begin{aligned} G^{-1}(u) - F^{-1}(u) &= o((1 - u)^{-1}) \quad \text{as} \quad u \to 1 , \qquad \text{and} \\ &= o(u^{-1}) \qquad\quad \text{as} \quad u \to 0 . \end{aligned}$$

In as much as $Z = G^{-1}(U_1) - F^{-1}(U_1)$, condition (4) follows immediately: for any fixed $\varepsilon > 0$ and sufficiently large n,

$$\Pr(|Z| > n) \leqq \Pr\left(\frac{\varepsilon}{1 - U_1} > n\right) + \Pr\left(\frac{\varepsilon}{U_1} > n\right) = 2\varepsilon/n . \qquad \square$$

It is a simple matter to replace X by any nonnormal stable (random) variable in Theorem 1. That is, the classical conditions for a random variable Y to be in the domain of *normal* attraction of a (nonnormal) stable variable can easily be shown to be sufficient using a theorem similar to Theorem 1. Rather than do this, we establish an even more general weak invariance principle in Section 3. It is applicable to the study of domains of (not necessarily normal) attraction of stable laws.

If one were to try to prove the classical central limit theorem using the above approach, one would have to establish

$$\sum_{i=1}^{n} (Z_i - EZ)/n^{\frac{1}{2}} \to_p 0 \quad \text{as} \quad n \to \infty .$$

But this would require Var Z to be zero, which is of course too strong. It is from considerations such as this that we were led to work with random variables which are outside the domain of partial attraction of a normal random variable. This is our starting point in Section 3.

3. The weak invariance principle. The weak invariance principle described in Theorem 2 below basically says that, under appropriate assumptions, it is possible on some probability space to define two normalized sums of i.i.d. random variables

$$a(n)^{-1} \sum_{i=1}^{n} X_i \quad \text{and} \quad b(n)^{-1} \sum_{i=1}^{n} Y_i ,$$

with specified distributions for the X's and Y's, so that the difference

$$S_n = b(n)^{-1} \sum_{i=1}^{n} Y_i - a(n)^{-1} \sum_{i=1}^{n} X_i \tag{6}$$

is virtually nonstochastic. Specifically, it is shown that

$$S_n = h(n) + o_p(1) \quad \text{as} \quad n \to \infty$$

for appropriate constants $h(n)$. We shall begin with a discussion of assumptions.

Let X be an unbounded random variable. A positive nondecreasing[5] function $a(\cdot)$ on R^+ (the positive reals) is said to be a *norming function* for X if

$$\limsup_{x\to\infty} x \Pr(|X| > a(x)) < \infty , \tag{7}$$

$$\liminf_{x\to\infty} x \Pr(|X| > a(x)) > 0 , \tag{8}$$

and

$$a(x+1)/a(x) \to 1 \quad \text{as} \quad x \to \infty . \tag{9}$$

Quite clearly, (7) implies that $a(x) \to \infty$ as $x \to \infty$. In Theorem 2, we assume that $a(\cdot)$ is a norming function for X, that $b(\cdot)$ is a norming function for Y, and that $a(\cdot)$ and $b(\cdot)$ are related by the following two conditions:

C1. $a(\cdot)/b(\cdot)$ is a slowly varying function.

C2. For each $B \in (0, 1)$ as $x \to \infty$,

$$x \Pr(\pm Y > B^{-1}b(x)) - o(1) \leqq x \Pr(\pm X > a(x)) \leqq x \Pr(\pm Y > Bb(x)) + o(1)$$

and

$$x \Pr(\pm X > B^{-1}a(x)) - o(1) \leqq x \Pr(\pm Y > b(x)) \leqq x \Pr(\pm X > Ba(x)) + o(1) .$$

[5] The assumption that $a(\cdot)$ is nondecreasing does not appear in our original manuscript (Simons and Stout (1975)). It is a mild restriction, as we use it, and it simplifies some of our discussions.

Here, in C2 (and below in A3 and A3′), the interpretation is that the statements in C2 (A3 and A3′) are true with all pluses preceding the random variables and with all minuses preceding the random variables. It is further assumed that X and Y satisfy the following three assumptions which are described for X:

A1. $$\limsup_{x\to\infty} \frac{EX^2 I(|X| \leqq x)}{x^2 \Pr(|X| > x)} < \infty .$$

A2. $$\sup_{x\geqq 1} \frac{\Pr(|X| > Ax)}{\Pr(|X| > x)} \to 0 \quad \text{as} \quad A \to \infty .$$

A3. For no positive $B \neq 1$ does the vector

$$\left(\frac{\Pr(\pm X > x)}{\Pr(|X| > x)}, \frac{\Pr(\pm X > x)}{\Pr(\pm X > Bx)}\right)$$

have a limit point $(c, 1)$ with $0 < c \leqq 1$ as $x \to \infty$.

Assumption A1 simply means that the random variable X is *not* in the domain of partial attraction of a normal variable. (See Paul Lévy (1954), page 113, for a proof.) Integration by parts yields

(10) $$EX^2 I(|X| \leqq x) = 2 \int_0^x v \Pr(|X| > v)\, dv - x^2 \Pr(|X| > x) .$$

Consequently, A1 is equivalent to

A1′. $$\limsup_{x\to\infty} \frac{\int_0^x v \Pr(|X| > v)\, dv}{x^2 \Pr(|X| > x)} < \infty .$$

In Section 5, it is shown that A1 is also equivalent to each of:

A1″. *For some* $\varepsilon > 0$, $x^{-\varepsilon} \int_0^x v \Pr(|X| > v)\, dv$ *is an increasing function of* x *on* R^+.

A1‴. *For some* $\varepsilon > 0$ *and some* $M > 1$

$$x^{2-\varepsilon} \Pr(|X| > x) \leqq M y^{2-\varepsilon} \Pr(|X| > y) , \qquad y > x > 0 .$$

It is useful to know that assumption A1 implies

(11) For each $A > 1$, $\inf_{x>0} \Pr(|X| > Ax)/\Pr(|X| > x) > 0$.[6]

This is immediate from A1‴. Another way of stating (11) is to say that the function $\Pr(|X| > x)$ is of dominated variation (see Seneta (1976), page 99), i.e.,

$$\liminf_{x\to\infty} \Pr(|X| > Ax)/\Pr(|X| > x) > 0 \qquad \text{for some} \quad A > 1 .$$

Thus A1 implies $\Pr(|X| > x)$ varies dominatedly. But it says somewhat more;

[6] The fact that A1 implies (11) is implicit in the work of C. Heyde (1969). He derives from A1 an inequality like that of A1‴ but with $\varepsilon = 0$. It is weaker than A1‴, but it immediately implies (11). In an earlier paper, Heyde (1967) assumes both A1 and a weakened version of (11). It is now known that the latter is redundant.

e.g., the function $\Pr(|X| > x) = x^{-2}$, $x \geqq 1$, varies dominatedly, but such an X does not satisfy A1.[7]

Assumption A2 implies that the norming fuction $a(\cdot)$ is of dominated variation (see Seneta (1976), page 99), i.e.,

$$\limsup_{x\to\infty} a(Ax)/a(x) < \infty \qquad \text{for some} \quad A > 1\,.$$

This can be stated equivalently as

$$\sup_{x\geqq 1} a(Ax)/a(x) < \infty \qquad \text{for each} \quad A > 1\,. \tag{12}$$

To be precise, the identity

$$\frac{\Pr(|X| > \{a(Ax)/a(x)\}a(x))}{\Pr(|X| > a(x))} = \frac{\{Ax\Pr(|X| > a(Ax))\}}{A\{x\Pr(|X| > a(x))\}}$$

immediately shows that (7), (8) and A2 imply (12).

The statement of A3 is made somewhat cumbersome in order to make it applicable in a wide variety of situations involving unbalanced tails. For instance when the first component of the vector in A3 converges to zero, no conclusion is required concerning the asymptotic behavior of the second component, and it would be seriously restrictive to require one. (The ratio might even assume the form "$\frac{0}{0}$" for large x.) Although somewhat stronger than A3, assumption A3′ below is often true in applications and captures much of the spirit of A3:

A3′. *Given any* $B \in (0, 1)$, *the ratio* $\Pr(\pm X > Bx)/\Pr(\pm X > x)$ *is (defined and) bounded away from* 1 *for large* x.

Note that both A2 and A3 imply that, in certain senses, $\Pr(|X| > y)/\Pr(|X| > x)$ is small for $y > x$.

Before continuing, one should also observe that, because of the identity

$$\frac{r}{s} = \frac{\{r\Pr(|X| > a(r))\}}{\{s\Pr(|X| > a(s))\}}\left\{\frac{\Pr(|X| > (a(s)/a(r))a(r))}{\Pr(|X| > a(r))}\right\},$$

the conditions (7), (8) and (11) give rise to the useful implication:

(13) $a(r_n)/a(s_n)$ is bounded away from zero $\Rightarrow r_n/s_n$ is bounded away from zero for arbitrary positive sequences $r_n \to \infty$ and $s_n \to \infty$.

THEOREM 2. *Suppose X and Y possess norming functions $a(\cdot)$ and $b(\cdot)$, respectively, and that they each satisfy assumptions* A1—A3. *Further, suppose conditions* C1 *and* C2 *are satisfied. Then on some probability space there exist two sequences of* i.i.d. *random variables* $X_1, X_2, \cdots$ *and* $Y_1, Y_2, \cdots$ *whose members are distributed*

[7] The fact that A1 says somewhat more than that $\Pr(|X| > x)$ varies dominatedly explains why the equivalence of A1, A1′ and A1‴ does not follow from Theorems A.2, A.6 and A.7 of Seneta's (1976) appendix, which appears to be the case at first glance. Certain unspecified parameters appearing in these theorems must be pinned down precisely.

as X and Y, respectively, and which satisfy

$$\frac{\sum_{i=1}^{n} Y_i}{b(n)} = \frac{\sum_{i=1}^{n} X_i}{a(n)} + h(n) + o_p(1) \qquad \text{as} \quad n \to \infty \tag{14}$$

for some sequence of constants $h(n)$.

If X is a stable variable of index $\alpha \in (0, 2)$ and $a(n) = n^{1/\alpha}$, it immediately follows from (14) that Y is in the domain of attraction of X.

The proof of Theorem 2 depends on Lemma 2, which, in turn, depends on an important fact about slowly varying functions, described in Lemma 1 (see Seneta (1976), page 2). We now state these lemmas for easy reference, delaying the proof of Lemma 2 until after the proof of Theorem 2.

LEMMA 1. *(Uniform convergence theorem for slowly varying functions.) If $\beta(\cdot)$ is a slowly varying function on R^+, then the required convergence*

$$\beta(Ax)/\beta(x) \to 1 \qquad \text{as} \quad x \to \infty \tag{15}$$

occurs uniformly in A on each fixed finite interval $[c, d]$, $0 < c < d < \infty$.

LEMMA 2. *Assume the hypotheses of Theorem 2. Fix $\xi, \eta > 0$ and let*

$$c_3 = \min(F(\xi a(n)), G(\xi b(n))) \qquad \text{and} \qquad c_4 = 1 - \frac{\eta}{n},$$

where F and G are the distribution functions of X and Y (appearing in the statement of Theorem 2) respectively. Then for each $\delta > 0$, there exists a large integer N such that for $n \geqq N$ and $u \in (c_3, c_4)$ (possibly empty),

$$\left(\frac{F^{-1}(u)}{a(n)} - \frac{G^{-1}(u)}{b(n)}\right)^2 \leqq \delta\left(\frac{F^{-1}(u)}{a(n)}\right)^2.$$

PROOF OF THEOREM 2. Define the sequences $X_1, X_2, \cdots$ and $Y_1, Y_2, \cdots$ in terms of i.i.d. uniform variables $U_1, U_2, \cdots$ (on $(0, 1)$) by letting

$$X_i = F^{-1}(U_i) \qquad \text{and} \qquad Y_i = G^{-1}(U_i), \qquad i \geqq 1$$

where F and G are the distribution functions of X and Y, respectively. Set

$$Z_{ni} = \frac{Y_i}{b(n)} - \frac{X_i}{a(n)} \qquad \text{and} \qquad S_n = \sum_{i=1}^{n} Z_{ni}.$$

It must be shown that for a sequence of constants $h(n)$

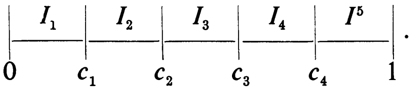
$$S_n - h(n) \to_p 0 \qquad \text{as} \quad n \to \infty. \tag{16}$$

An intuitive argument suggesting why each Z_{ni} and hence S_n (except for centering) is small with probability close to one is perhaps helpful before we plunge into technical details. Suppress the index i. There are three possible ranges for U. Either U is close to 1 or 0, U is moderately close to 1 or 0 or U is not close to 1 or 0. The first case happens with small probability and, hence, causes no problem. In the second case, $F^{-1}(U)/a(n) \approx G^{-1}(U)/b(n)$ and, hence,

$Z_n \approx 0$. In the third case, both $F^{-1}(U)/a(n)$ and $G^{-1}(U)/b(n)$ must be small and, hence, $Z_n \approx 0$. Thus, as desired, $\Pr(Z_n \approx 0) \approx 1$. The actual argument proceeds by truncation and centering. Although somewhat easier, our argument, in essence, amounts to verifying the conditions of the classical degenerate convergence theorem. (See Loève (1963), page 317.)

For each n, express the unit interval as the union of five subintervals $I_1 < I_2 < I_3 < I_4 < I_5$ where $I_1 = (0, c_1]$, $I_2 = (c_1, c_2)$, $I_3 = [c_2, c_3]$, $I_4 = (c_3, c_4)$ and $I_5 = [c_4, 1)$, where

$$c_1 = \eta/n\,, \qquad c_2 = \max(F(-\xi a(n)), G(-\xi b(n)))\,,$$
$$c_3 = \min(F(\xi a(n)), G(\xi b(n)))\,, \qquad c_4 = 1 - \eta/n\,.$$



The constants η and ξ, which will be chosen later, are intended to be small and positive. Certain of the intervals may be empty or overlap. For instance, if $c_4 \leqq c_3$, I_4 is empty and the two intervals I_3 and I_5 overlap.

The first step involves truncation. Let

$$S_n' = \sum_{i=1}^n Z'_{ni} \qquad \text{and} \qquad S''_n = S_n - S_n'\,,$$

where

$$Z'_{ni} = Z_{ni} \qquad \text{if} \quad U_i \in (I_2 \cup I_3 \cup I_4)\,,$$
$$= 0 \qquad \text{otherwise.}$$

The Z'_{ni}'s and, hence, S_n' are bounded random variables. Fix $\varepsilon > 0$. Then

$$\begin{aligned}\Pr(|S_n - ES_n'| > 2\varepsilon) &\leqq \Pr(|S_n' - ES_n'| > \varepsilon) + \Pr(|S''_n| > \varepsilon)\\ &\leqq \varepsilon^{-2}\operatorname{Var} S'_n + \Pr(\textstyle\bigcup_{i=1}^n [U_i \in (I_1 + I_5)])\\ &\leqq \varepsilon^{-2} n \operatorname{Var} Z'_{n1} + n \Pr(U_1 \in (I_1 + I_5))\\ &= \varepsilon^{-2}\, n \operatorname{Var} Z'_{n1} + 2\eta\,.\end{aligned}$$

It will be argued that

$$nE(Z'_{n1})^2 \to 0 \qquad \text{as} \quad n \to \infty\,, \tag{17}$$

so that

$$\limsup_{n\to\infty} \Pr(|S_n - ES_n'| > 2\varepsilon) \leqq 2\eta\,. \tag{18}$$

Even though $\eta > 0$ and $\varepsilon > 0$ are arbitrary, this does not prove (16) since the value of ES_n' depends on η. However, if $h(n)$ is defined as a median of S_n, then (16) follows from (18) and what Loève ((1963), page 245) calls a symmetrization inequality. Thus, truncation has reduced the task to showing (17).

Now

$$\begin{aligned}nE(Z'_{n1})^2 &= n \int_{c_1}^{c_4} \left(\frac{F^{-1}(u)}{a(n)} - \frac{G^{-1}(u)}{b(n)}\right)^2 du\\ &\leqq n\left(\int_{I_2} + \int_{I_3} + \int_{I_4}\right)\left(\frac{F^{-1}(u)}{a(n)} - \frac{G^{-1}(u)}{b(n)}\right)^2 du\,.\end{aligned}$$

The integrals over I_2, I_3 and I_4 must be treated separately. The integral over I_2 is treated like the integral over I_4 and, hence, will not be discussed. We shall discuss the integral over I_3 first.

$$\int_{I_3}\left(\frac{F^{-1}(u)}{a(n)}-\frac{G^{-1}(u)}{b(n)}\right)^2 du$$
$$\leqq 2\int_{I_3}\left\{\left(\frac{F^{-1}(u)}{a(n)}\right)^2+\left(\frac{G^{-1}(u)}{b(n)}\right)^2\right\} du$$
$$\leqq 2\int_{F(-\xi a(n))}^{F(\xi a(n))}\left(\frac{F^{-1}(u)}{a(n)}\right)^2 du + 2\int_{G(-\xi b(n))}^{G(\xi b(n))}\left(\frac{G^{-1}(u)}{b(n)}\right)^2 du .$$

We shall only discuss the integral

$$I(n)=\int_{F(-\xi a(n))}^{F(\xi a(n))}\left(\frac{F^{-1}(u)}{a(n)}\right)^2 du ;$$

the integral involving G can be handled the same way. Specifically, we shall show that

$$\limsup_{n\to\infty} nI(n)\to 0 \qquad \text{as} \quad \xi\downarrow 0 . \tag{19}$$

This will be sufficiently strong (to prove (17)) providing it can be shown that for each small but fixed $\xi>0$,

$$n\int_{I_4}\left(\frac{F^{-1}(u)}{a(n)}-\frac{G^{-1}(u)}{b(n)}\right)^2 du\to 0 \qquad \text{as} \quad n\to\infty . \tag{20}$$

Now, recalling (3) and using (10),

$$nI(n)=\frac{n}{a^2(n)}EX^2I(-\xi a(n)<X\leqq \xi a(n))\leqq \frac{n}{a^2(n)}EX^2I(|X|\leqq \xi a(n))$$
$$\leqq \frac{2n}{a^2(n)}\int_0^{\xi a(n)} v\,\Pr(|X|>v)\,dv .$$

Since X satisfies assumption A1 it satisfies A1′ and A1″. Using A1″, then A1′, it follows for some small $\gamma>0$ and some large M that

$$nI(n)\leqq \frac{2n}{a^2(n)}\xi^\gamma\int_0^{a(n)} v\,\Pr(|X|>v)\,dv$$
$$\leqq M\xi^\gamma n\Pr(|X|>a(n)) .$$

Since $a(n)$ is a norming function for X, (19) follows from (7).

It remains to prove (20). Using Lemma 2, we have, for each small $\delta>0$ and all sufficiently large n,

$$n\int_{I_4}\left(\frac{F^{-1}(u)}{a(n)}-\frac{G^{-1}(u)}{b(n)}\right)^2 du\leqq \delta n\int_{I_4}\left(\frac{F^{-1}(u)}{a(n)}\right)^2 du \tag{21}$$
$$\leqq \frac{\delta n}{a^2(n)}\int_{c_1}^{c_4}(F^{-1}(u))^2\,du .$$

By (7) and A2,

$$\sup_n n \Pr(|X| > Aa(n)) \to 0 \quad \text{as} \quad A \to \infty .$$

In turn, this implies

$$\sup_n n \Pr(|X| \geqq Aa(n)) \to 0 \quad \text{as} \quad A \to \infty .$$

Thus, for some large A,

$$c_1 = \eta/n \geqq \Pr(|X| \geqq Aa(n)) \geqq F(-Aa(n)) ,$$

and

$$c_4 = 1 - \eta/n \leqq 1 - \Pr(|X| \geqq Aa(n)) \leqq F(Aa(n)) . \tag{22}$$

Consequently, recalling (3),

$$\begin{aligned} \frac{\delta n}{a^2(n)} \int_{c_1}^{c_4} (F^{-1}(u))^2 \, du &\leqq \frac{\delta n}{a^2(n)} \int_{F(-Aa(n))}^{F(Aa(n))} (F^{-1}(u))^2 \, du \\ &= \frac{\delta n}{a^2(n)} EX^2 I(-Aa(n) < X \leqq Aa(n)) \\ &\leqq \frac{\delta n}{a^2(n)} EX^2 I(|X| \leqq Aa(n)) . \end{aligned} \tag{23}$$

But, by assumption A1 and (22), there is a large L such that

$$\frac{\delta n}{a^2(n)} EX^2 I(|X| \leqq Aa(n)) \leqq \delta A^2 L n P(|X| > Aa(n)) \leqq \delta A^2 L \eta . \tag{24}$$

Since $\delta > 0$ can be made arbitrarily small, (20) follows from (21), (23) and (24). □

PROOF OF LEMMA 2. Suppose the lemma is false. Then there exists a $\delta > 0$ and $u_n \in (c_3, c_4)$ such that

$$\left(\frac{F^{-1}(u_n)}{a(n)} - \frac{G^{-1}(u_n)}{b(n)}\right)^2 > \delta \left(\frac{F^{-1}(u_n)}{a(n)}\right)^2 \tag{25}$$

for each n in some infinite index set $\mathscr{N}$. This will be contradicted by showing that

$$\text{one is a limit point of} \quad \left\{\frac{F^{-1}(u_n)b(n)}{a(n)G^{-1}(u_n)} , \ n \in \mathscr{N}\right\} . \tag{26}$$

Since $u_n > c_3$, either $u_n > F(\xi a(n))$ or $u_n > G(\xi b(n))$ for infinitely many $n \in \mathscr{N}$. We consider only the case where $u_n > F(\xi a(n))$ for infinitely many $n \in \mathscr{N}$, the argument in the other case being very similar. By making $\mathscr{N}$ a smaller index set if necessary, we assume

$$u_n > F(\xi a(n)) \quad \text{for all} \quad n \in \mathscr{N} .$$

Thus, recalling (3),

$$F^{-1}(u_n) > \xi a(n) \to \infty \quad \text{as} \quad n \to \infty \quad \text{in} \quad \mathscr{N} . \tag{27}$$

Again, by deleting indices from $\mathscr{N}$ if necessary, we assume that $F^{-1}(u_n) > a(1)$ for all $n \in \mathscr{N}$, so that there exists a positive integer r_n for which

(28) $$a(r_n) < F^{-1}(u_n) \leqq a(r_n + 1)\,, \qquad n \in \mathscr{N}.$$

In the remainder of the proof, all statements concerning n will suppose $n \in \mathscr{N}$.

Clearly $r_n \to \infty$ as $n \to \infty$. In fact, r_n/n is bounded away from zero since

$$\xi a(n) < F^{-1}(u_n) \leqq a(r_n + 1)$$

(see, (27), (28) and then (13)). Likewise, r_n/n is bounded away from infinity because

(29) $$u_n < c_4 = 1 - \eta/n \leqq F(Aa(n))$$

for some large A (see, (22)), and hence, (using (3))

$$a(r_n) < F^{-1}(u_n) \leqq Aa(n)$$

(apply (13) again). Since $a(\cdot)/b(\cdot)$ is slowly varying and r_n/n is bounded away from zero and infinity, it follows from Lemma 1 that

(30) $$\frac{a(r_n)b(n)}{a(n)b(r_n)} \to 1\,.$$

Thus (26) holds providing

$$\frac{G^{-1}(u_n)}{b(r_n)} \to 1 \qquad \text{as} \quad n \to \infty$$

(see, (9) and (28)). But this must hold if, for an arbitrary $B \in (0, 1)$, one has

(31) $$\liminf_{n\to\infty} \frac{G^{-1}(u_n)}{b(r_n)} \geqq B$$

and

(32) $$\limsup_{n\to\infty} \frac{c_n}{b(r_n + 1)} \leqq B^{-1}\,,$$

where $c_n = G^{-1}(u_n) - 1$. (Note $b(r_n) \to \infty$ and, because of (9), $b(r_n + 1)/b(r_n) \to 1$. The sequence c_n is introduced for a technical reason: $U > u_n$ does not imply the strict inequality $G^{-1}(U) > G^{-1}(u_n)$, but it does imply the strict inequality $G^{-1}(U) > c_n$. Here and below, U is a uniform variable on (0, 1).) Rather than show (31) and (32) directly, we shall show that

(33) $$\frac{G^{-1}(u_n) \wedge Bb(r_n)}{Bb(r_n)} \to 1 \quad \text{and} \quad \frac{c_n \vee B^{-1}b(r_n + 1)}{B^{-1}b(r_n + 1)} \to 1 \quad \text{as} \quad n \to 1\,,$$

which are equivalent to (31) and (32) respectively ($x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$).

Now

$$
\begin{aligned}
r_n \Pr(Y > c_n) &= r_n \Pr(G^{-1}(U) > G^{-1}(u_n) - 1) \\
&\geqq r_n \Pr(U > u_n) = r_n(1 - u_n) \\
&\geqq r_n \Pr(F^{-1}(U) > F^{-1}(u_n)) \qquad (34) \\
&\geqq r_n \Pr(X > a(r_n + 1)) \\
&\geqq \frac{r_n}{r_n + 1}\{(r_n + 1)\Pr(Y > B^{-1}b(r_n + 1)) - o(1)\} \\
&\qquad \text{as} \quad n \to \infty ,
\end{aligned}
$$

on account of (28) and condition C2. But the quantity $r_n(1 - u_n)$, appearing in (34), is bounded away from zero because r_n/n is bounded away from zero and $u_n < 1 - \eta/n$ (see, (29)). Hence, $r_n \Pr(Y > c_n)$ is bounded away from zero and (34) yields

$$
\Pr(Y > B^{-1}b(r_n + 1)) \leqq \Pr(Y > c_n)(1 + o(1)) \qquad \text{as} \quad n \to \infty . \tag{35}
$$

In turn, this implies

$$
\frac{\Pr(Y > c_n \vee B^{-1}b(r_n + 1))}{\Pr(Y > B^{-1}b(r_n + 1))} \to 1 \qquad \text{as} \quad n \to \infty .
$$

Thus, either the second limit in (33) holds or the ratio $\Pr(Y > B'B^{-1}b(r_n + 1))/\Pr(Y > B^{-1}b(r_n + 1))$ has unity as a limit point for some $B' > 1$. By assumption A3 (for Y), this would require the ratio $\Pr(Y > B^{-1}b(r_n + 1))/\Pr(|Y| > B^{-1}b(r_n + 1))$ to have zero as a limit point. But $r_n \Pr(|Y| > B^{-1}b(r_n + 1))$ is bounded away from zero (see, (8) and (11)), and consequently, $r_n \Pr(Y > B^{-1}b(r_n + 1))$ has a limit point equal to zero. Since $r_n \Pr(Y > c_n)$ is bounded away from zero, as noted earlier, (35) implies that $B^{-1}b(r_n + 1) > c_n$ infinitely often. By deleting indices from $\mathscr{N}$ if necessary, one can make $B^{-1}b(r_n + 1) > c_n$ for all $n \in \mathscr{N}$. With this modification, the second limit in (33) holds. The proof of the first limit in (33) is similar but easier. The inequality analogous to (35) is

$$
\Pr(Y > G^{-1}(u_n)) \leqq \Pr(Y > Bb(r_n))(1 + o(1)) \qquad \text{as} \quad n \to \infty
$$

and, in the analogous argument, one concludes that $r_n \Pr(Y > Bb(r_n))$ is bounded away from zero. □

In some applications, it is helpful to be able to replace "$a(\cdot)/b(\cdot)$ is slowly varying" (i.e., condition C1), which appears in Theorem 2, by conditions more directly verifiable from F and G. Before stating these conditions, we make the following general observation: if $\beta(\cdot)$ is any slowly varying function and $b(\cdot)$ is any function of dominated variation, then $\beta(b(\cdot))$ is slowly varying. This follows immediately from Lemma 1. Recall that the norming functions $a(\cdot)$ and $b(\cdot)$ are of dominated variation by virtue of conditions (7), (8) and A2 (cf., (12)).

COROLLARY 2. *Theorem 2 remains valid if conditon* C1 *is replaced by the three conditions*

(36) $\beta(x) = \Pr(|Y| > x)/\Pr(|X| > x)$ *is slowly varying in* x ;

(37) $x\Pr(|Y| > b(x))$ *is slowly varying in* x ; *and*

(38) *a positive function* $\gamma(\cdot)$ *on* R^+ *is slowly varying whenever* $x\Pr(|X| > \gamma(x)a(x))$ *is slowly varying in* x .

PROOF. Since $\beta(b(\cdot))$ is slowly varying (because of (36) and the foregoing discussion) and

(39) $$\beta(b(x)) = \{x\Pr(|Y| > b(x))\}/\{x\Pr(|X| > \gamma(x)a(x))\}$$

when $\gamma(\cdot) = b(\cdot)/a(\cdot)$, condition C1 follows from (37) and (38). □

Observe that the conditions (37) and (38), entering into Corollary 2, are asymmetric in X and Y. Quite clearly Corollary 2 would remain valid if (37) and (38) were replaced by

(40) $x\Pr(|X| > a(X))$ is slowly varying in x ,

and

(41) a positive function $\gamma(\cdot)$ on R^+ is slowly varying whenever $x\Pr(|Y| > \gamma(x)b(x))$ is slowly varying in x .

4. Domains of attraction. Theorem 2 is specialized in this section to the study of stable (random) variables.

Recall that a random variable X is said to be stable if for each n there are independent random variables $X_1, X_2, \cdots, X_n$ with common distribution that of X, centering constants $e(n)$ and scaling constants $a(n) > 0$ such that

(42) $$\frac{\sum_{i=1}^n X_i}{a(n)} - e(n)$$

has the distribution of X. It is well known that, corresponding to each properly scaled nonnormal stable variable X, there exists a pair (α, p), $0 < \alpha < 2$, $0 \leqq p \leqq 1$, such that

(43) $$\Pr(|X| > x) \sim 1/x^\alpha , \qquad \Pr(X > x) \sim p/x^\alpha \quad \text{as} \quad x \to \infty .$$

Moreover, for such a random variable, $a(n)$ in (42) must assume the form $n^{1/\alpha}$. Observe, by direct substitution, that $a(x) = x^{1/\alpha}$ is a norming function for any random variable X satisfying (43). Any stable variable X which (as a result of proper scaling) satisfies (43) will be said to be of *type* (α, p).[8]

THEOREM 3. *Let X be a stable variable of type (α, p) and Y be an unbounded*

[8] If the word *type* were to be used as Loève does ((1963), page 202), then each pair (α, p) would correspond to exactly one type of stable law. This fact explains our terminology.

random variable with a distribution of the form

$$\Pr(|Y| > x) = \beta(x)/x^{\alpha}, \quad x > 0, \tag{44}$$

$$\Pr(Y > x)/\Pr(|Y| > x) \to p \qquad \textit{as} \quad x \to \infty, \tag{45}$$

where $0 < \alpha < 2$, $0 \leqq p \leqq 1$, *and* $\beta(x)$ *is a slowly varying function. Then, on some probability space, there exist two sequences of* i.i.d. *random variables* $X_1, X_2, \cdots$ *and* $Y_1, Y_2, \cdots$ *whose members are distributed as* X *and* Y, *respectively, and which satisfy*

$$\frac{\sum_{i=1}^{n} Y_i}{b(n)} = \frac{\sum_{i=1}^{n} X_i}{n^{1/\alpha}} + h(n) + o_p(1) \qquad \textit{as} \quad n \to \infty \tag{46}$$

for some positive function $b(x)$ *and some sequence of centering constants* $h(n)$. *Moreover,* (46) *holds for each choice of* $b(x)$ *satisfying*

$$\frac{x\beta(b(x))}{b^{\alpha}(x)} \to 1 \qquad \textit{as} \quad x \to \infty. \tag{47}$$

Theorem 3 yields an obvious but well-known corollary.

COROLLARY 3. *Under the hypotheses of Theorem* 3, *there exist constants* $b(n) > 0$ *and* $g(n)$ *such that*

$$\frac{\sum_{i=1}^{n} Y_i}{b(n)} - g(n) \to_L X, \tag{48}$$

where $Y_1, Y_2, \cdots$ *are* i.i.d. *random variables distributed as* Y. *Moreover,* (48) *holds for each positive function* $b(x)$ *satisfying* (47).

Corollary 3 is the strongest possible result in the direction it is stated. That is, a random variable Y is in the domain of attraction of a stable random variable X of type (α, p) ($0 < \alpha < 2$, $0 \leqq p \leqq 1$) iff Y satisfies the assumptions stated in the corollary. We have tried unsuccessfully to obtain the converse of Corollary 3 by probabilistic methods. Of course, a proof of the converse based on Fourier transforms is well known.

PROOF OF THEOREM 3. We are going to apply Corollary 2. Conditions (36) and (38) are immediate, and condition (37) follows from the fact that (44) and (47) combine to give

$$x \Pr(|Y| > b(x)) \to 1 \qquad \text{as} \quad x \to \infty.$$

This limit also implies (7) and (8) for the function $b(\cdot)$. Since (47) implies

$$\frac{b^{\alpha}(x+1)}{b^{\alpha}(x)} \sim \frac{\beta(b(x+1))}{\beta(b(x))} \qquad \text{as} \quad x \to \infty,$$

condition (9) holds for $b(\cdot)$ (and, hence, $b(\cdot)$ is a norming function) *if* $\beta(b(\cdot))$ is slowly varying. But $\beta(b(\cdot))$ is slowly varying because $\beta(\cdot)$ is slowly varying by assumption, $b(\cdot)$ satisfies (7) and (8), and Y (as will be shown) satisfies assumption

A2. (See the discussion preceding Corollary 2.) Clearly $a(x) = x^{1/\alpha}$ is a norming function for X. Thus it remains to verify assumptions A1—A3 for X and Y and condition C2. Of these, the only nontrivial conditions to check are A1—A3 for Y. It is easier to check assumption A1′ than its equivalent A1. Showing A1′ for Y amounts to showing

$$\limsup_{x\to\infty} \frac{\int_0^x v^{1-\alpha}\beta(v)\,dv}{x^{2-\alpha}\beta(x)} < \infty .$$

But this is immediate from a theorem appearing in Feller ((1966), page 273). Concerning assumptions A2 and A3, (44) yields

$$\frac{\Pr(|Y| > Ax)}{\Pr(|Y| > x)} = \frac{\beta(Ax)}{A^\alpha \beta(x)} .$$

From this and the fact that $\beta(x)$ is slowly varying, A2 is immediate for Y. Using (45), the proof of A3′ (and hence A3) for Y is similar. □

5. Discussion of the assumptions. A major objective of this section is to explain the significance and, to the extent possible, the *raison d'être* of the assumptions made in Theorem 2 (our main theorem). Recall that these assumptions restrict attention to random variables which have norming functions and satisfy assumptions A1—A3 and conditions C1 and C2. We will try to make a strong case for all of these requirements with the exception of assumption A3. Roughly speaking, assumption A3 requires a distribution function not to have locally nearly flat spots arbitrarily far out in its tails. Feller (1968), in a somewhat related context, has found it necessary to make a similar restriction. We have ambivalent feelings as to the necessity of A3.

The theory developed in Section 3 is predicated on the use of norming functions to normalize the sums appearing in (14). While other scalings might be considered, our decision to work with norming functions has two good reasons. To begin with, our methods are not suited for studying the classical central limit theorem (see the discussion at the end of Section 2), and consequently the usual normalization by $1/n^{\frac{1}{2}}$ is inappropriate. Secondly, norming functions provide exactly the right amount of scaling when one is working with random variables which are in the domain of attraction of a nonnormal stable variable; i.e., they are the *right* things to consider if one wants to study domains of attraction.

The following theorems explain why assumptions A1 and A2 are quite natural when norming functions are used to normalize sums. In both theorems it is assumed that $X, X_1, X_2, \cdots$ are i.i.d. that $T_n = X_1 + X_2 + \cdots + X_n$ and that X has a norming function $a(x)$.

THEOREM 4. *There exists a sequence of numbers c_n such that $\{(T_n - c_n)/a(n)\}_{n=1}^\infty$ is tight* iff *X satisfies assumptions* A1 *and* A2.

THEOREM 5. *If X satisfies* (11) (*which is implied by* A1), *then for every sequence* *of numbers c_n, the sequence of random variables $(T_n - c_n)/a(n)$ $(n \geqq 1)$ has no degenerate weak limit point. Specifically, for each $A > 0$,*

$$\liminf_{n\to\infty} \Pr(|T_n - c_n| > Aa(n)) > 0 . \tag{49}$$

Tightness for the sequence $\{(T_n - c_n)/a(n)\}_{n=1}^{\infty}$ guarantees that every subsequence has a weak limit point. Thus, by Theorems 4 and 5, A1 and A2 guarantee that such limit points must exist and be nondegenerate (for properly chosen c_n's). Further, Theorem 4 implies that A1 and A2 are necessary for tightness provided one normalizes by norming functions. Thus, Theorems 4 and 5 show that assumptions A1 and A2 are well suited to the study of domains of attraction.

There are interesting connections between Theorems 4 and 5 and the subject of domains of partial attraction. Doeblin (1940) shows that an unbounded random variable X belongs to *no* domain of partial attraction if

$$\lim_{A\to\infty} \liminf_{x\to\infty} \Pr(|X| > Ax)/\Pr(|X| > x) > 0 .$$

In contrast, it can be shown that assumption A2, which can be expressed as

$$\lim_{A\to\infty} \limsup_{x\to\infty} \Pr(|X| > Ax)/\Pr(|X| > x) = 0 ,$$

implies that X *is* in some domain of partial attraction. Theorems 4 and 5 make this clear when, in addition, assumption A1 holds and a norming function exists for X. Moreover, they tell how to properly scale T_n under the additional assumptions.

Theorem 2 says something about domains of partial attraction. Let $\mathscr{D}_X$ denote the collection of all nondegenerate infinitely divisible random variables Z such that X is in the domain of partial attraction of Z. According to the theory, $\mathscr{D}_X$ is empty, $\mathscr{D}_X$ consists of one type of random variable, or $\mathscr{D}_X$ consists of a nondenumerable set of types (Doeblin (1940), Gnedenko (1940)). The following result follows easily from Theorems 2, 4 and 5.

THEOREM 6. *If X and Y satisfy the hypotheses of Theorem* 2 *for some choice of norming functions $a(\cdot)$ and $b(\cdot)$, then $\mathscr{D}_X = \mathscr{D}_Y$.*

The proof of Theorem 6 (which will be deferred until later) shows that, for random variables which satisfy assumptions A1—A3, norming functions provide the normalizations appropriate for the study of domains of partial attraction.

Condition C1, which says that $a(\cdot)/b(\cdot)$ is slowly varying, is necessary in Theorem 2—at least for the way we constructed our random variables. Specifically, if (14) holds and the vectors (X_1, Y_1), (X_2, Y_2), $\cdots$ are i.i.d., then $a(\cdot)/b(\cdot)$ must be slowly varying.

In proving this claim, we shall assume, without loss of generality, that the random variables in (14) are symmetric and $h(n)$ is identically zero. (A standard symmetrization argument reduces the problem to this case.) Then (14) becomes

$$\frac{\sum_{i=1}^n Y_i}{b(n)} = \frac{\sum_{i=1}^n X_i}{a(n)} + o_p(1) \qquad \text{as} \quad n \to \infty .$$

Thus

$$\frac{\sum_{i=1}^{[x]} Y_i}{b([x])} = \frac{\sum_{i=1}^{[x]} X_i}{a([x])} + o_p(1) \quad \text{as} \quad x \to \infty, \tag{50}$$

where $[c]$ denotes the integer part of c. In view of (9), which holds for $a(\cdot)$ and $b(\cdot)$, it suffices to show that $a([\cdot])/b([\cdot])$ is slowly varying. For fixed $t \geqq 1$, (50) yields:

$$\begin{aligned} o_p(1) + \frac{\sum_{i=1}^{[xt]} X_i}{a([xt])} &= \frac{\sum_{i=1}^{[xt]} Y_i}{b([xt])} \\ &= \frac{b([x])}{b([xt])} \frac{\{\sum_{i=1}^{[x]} + \cdots + \sum_{i=[x]([t]-1)+1}^{[x][t]}\} Y_i}{b([x])} + \frac{\sum_{i=[x][t]+1}^{[xt]} Y_i}{b([xt])} \\ &= \frac{b([x])}{b([xt])} \left\{ \frac{\sum_{i=1}^{[x][t]} X_i}{a([x])} + o_p(1) \right\} + \frac{\sum_{i=[x][t]+1}^{[xt]} Y_i}{b([xt])} \end{aligned}$$

as $x \to \infty$. Thus

$$\begin{aligned} \left(1 - \frac{b([x])a([xt])}{a([x])b([xt])}\right) \frac{a([x][t])}{a([xt])} \frac{\sum_{i=1}^{[x][t]} X_i}{a([x][t])} \\ = \frac{\sum_{i=[x][t]+1}^{[xt]} Y_i}{b([xt])} - \frac{\sum_{i=[x][t]+1}^{[xt]} X_i}{a([xt])} + o_p(1) \quad \text{as} \quad x \to \infty. \end{aligned} \tag{51}$$

Since the random variables in the left hand side of (51) are independent of the random variables in the right hand side with the exception of the negligible $o_p(1)$ term, the two sides, after proper centering, must converge to zero in probability. Consequently,

$$\left(1 - \frac{b([x])a([xt])}{a([x])b([xt])}\right) \frac{a([x][t])}{a([xt])} \to 0 \quad \text{as} \quad x \to \infty,$$

or else Theorem 5 would be contradicted. Finally, since $a([x][t])/a([xt])$ is bounded away from zero as $x \to \infty$ (this follows from (12)),

$$\frac{b([x])a([xt])}{a([x])b([xt])} \to 1 \quad \text{as} \quad x \to \infty;$$

i.e., $a([\cdot])/b([\cdot])$ is slowly varying.

Finally, we claim C2 is a necessary condition (for the way we define our random variables in (14)). In order that (14) hold, it is necessary (c.f., Gnedenko and Kolmogorov (1965), page 128) that

$$n \Pr\left(\left|\frac{X}{a(n)} - \frac{Y}{b(n)}\right| > \varepsilon\right) \to 0 \quad \text{as} \quad n \to \infty$$

for each $\varepsilon > 0$. Thus for each $B \in (0, 1)$,

$$n \Pr(X > a(n)) \leqq n \Pr(Y > Bb(n)) + o(1) \quad \text{as} \quad n \to \infty, \tag{52}$$

which (because of (9)) can be upgraded to

$$x \Pr(X > a(x)) \leqq x \Pr(Y > Bb(x)) + o(1) \quad \text{as} \quad x \to \infty . \tag{53}$$

The other inequalities in C2 are derived similarly.

In spite of the fact that the assumptions of Theorem 2 are chosen with an eye toward the study of domains of attraction, they do not imply $\Pr(|X| > x)$ and $\Pr(|Y| > x)$ are regularly varying functions. (If X is in the domain of attraction of a nonnormal stable variable, $\Pr(|X| > x)$ is a regularly varying function.) This is demonstrated by the following example.

EXAMPLE. Let X be symmetric with

$$P(|X| > x) = \frac{2 + \sin(\log x)}{x}, \quad x > 3 .$$

Clearly X satisfies A1‴ with $\varepsilon = 1$ and $M = 3$. A2 is immediate. Consider A3′. We have

$$\frac{P(|X| > Bx)}{P(|X| > x)} = \frac{1}{B} \frac{2 + \sin(\log(Bx))}{2 + \sin(\log x)} .$$

Elementary calculations show that the above ratio is bounded away from one if $B \in (0, 1)$, establishing A3′. Also, $a(x) = x$ is a norming function. But as $x \to \infty$,

$$\frac{2 + \sin(\log xt)}{2 + \sin(\log x)}$$

clearly does not approach one for each fixed $t > 0$. Thus, $P[|X| > x]$ is not regularly varying. Clearly a Y can be constructed with distribution close enough to that of X such that C1 and C2 hold, as well as A1—A3, but with $P[|Y| > x]$ not regularly varying. □

The following lemmas are useful for proving Theorems 4 and 5. In both, it is assumed that the sequence $X, X_1, X_2, \cdots$ is i.i.d., that $T_n = X_1 + \cdots + X_n$, and that $a(x)$ is a norming function of X.

LEMMA 3. *There exist a sequence $\varepsilon_n \downarrow 0$ and constant $C < \infty$ such that for each $A \geqq 1$ and each sequence of numbers c_n,*

$$n \Pr(|X| > 6Aa(n)) \leqq C \Pr(|T_n - c_n| > Aa(n)) + \varepsilon_n . \tag{54}$$

LEMMA 4. *If* A1 *holds, then for some sequence of numbers c_n and some $M > 0$,*

$$\Pr(|T_n - c_n| > 4Aa(n)) \leqq Mn \Pr(|X| > Aa(n)) , \quad A > 0 . \tag{55}$$

PROOF OF LEMMA 3. For simplicity, the proof will be given for the special case $c_n \equiv 0$; the general case offers no additional difficulty. Let

$$E_i = [|T_n - X_i| \leqq 3Aa(n), |X_i| > 6Aa(n)] , \quad i = 1, \cdots, n .$$

Observe that $|T_n| > 3Aa(n)$ on each E_i and that $|T_n - X_i - X_j| > 3Aa(n)$ on each intersection $E_i \cap E_j$, $i \neq j$. Thus, by a Bonferroni inequality and the fact that the X_i's are i.i.d.,

$$\begin{aligned}\Pr(|T_n| > 3Aa(n)) &\geqq \Pr(\textstyle\bigcup_{i=1}^n E_i) \geqq \sum_{i=1}^n \Pr(E_i) - \sum_{1\leqq i\leqq j\leqq n} \Pr(E_i \cap E_j) \\ &\geqq n \Pr(|X| > 6Aa(n)) \Pr(|T_{n-1}| \leqq 3Aa(n)) \\ &\quad - \tbinom{n}{2} \Pr^2(|X| > 6Aa(n)) \Pr(|T_{n-2}| > 3Aa(n)) .\end{aligned}$$

Let $\gamma_n = n \Pr(|X| > a(n))$. Then, for $A \geqq \frac{1}{6}$,

$$\begin{aligned}(56) \qquad n \Pr(|X| > 6a(n)) &\leqq \Pr(|T_n| > 3Aa(n)) + \gamma_n \Pr(|T_{n-1}| > 3Aa(n)) \\ &\quad + \gamma_n{}^2 \Pr(|T_{n-2}| > 3Aa(n)) .\end{aligned}$$

Now, for $A \geqq 1$,

$$\begin{aligned}(57) \qquad \Pr(|T_{n-1}| > 3Aa(n)) &\leqq \Pr(|T_n| > 2Aa(n)) + \Pr(|X| > Aa(n)) \\ &\leqq \Pr(|T_n| > 2Aa(n)) + \gamma_n/n ,\end{aligned}$$

and similarly,

$$(58) \qquad \Pr(|T_{n-2}| > 3Aa(n)) \leqq \Pr(|T_n| > Aa(n)) + 2\gamma_n/n .$$

Since γ_n is bounded away from infinity (cf., (7)), the desired conclusion follows from (56), (57) and (58). Observe that the choice of ε_n and C does not need to depend on the value of A. □

Proof of Lemma 4. Fix $A > 0$ and let

$$X_{ni} = X_i I(|X_i| \leqq Aa(n)) \qquad \text{and} \qquad X'_{ni} = X_i - X_{ni} , \qquad i = 1, \cdots, n;\ n \geqq 1 .$$

Then

$$\begin{aligned}(59) \qquad \Pr(|\textstyle\sum_{i=1}^n X'_{ni}| > Aa(n)) &\leqq \Pr(\textstyle\bigcup_{i=1}^n [|X_i| > Aa(n)]) \\ &\leqq n \Pr(|X| > Aa(n)) .\end{aligned}$$

Assumption A1 says, for some large C, that

$$EX^2 I(|X| \leqq Aa(n)) \leqq CA^2a^2(n) \Pr(|X| > Aa(n)) .$$

Consequently, by Chebyshev's inequality,

$$\begin{aligned}(60) \qquad \Pr(|\textstyle\sum_{i=1}^n (X_{ni} - EX_{ni})| > Aa(n)) &\leqq \frac{nEX^2I(|X| \leqq Aa(n))}{A^2a^2(n)} \\ &\leqq Cn \Pr(|X| > Aa(n)) .\end{aligned}$$

From (59) and (60), we obtain

$$(61) \qquad \Pr(|T_n - \textstyle\sum_{i=1}^n EX_{ni}| > 2Aa(n)) \leqq (C+1)n \Pr(|X| > Aa(n)) .$$

This would prove the lemma were it not for the fact that the centering $\sum_{i=1}^n EX_{ni}$ depends on the value of A. If c_n is chosen to be a median of T_n then (55) follows from (61) and a symmetrization inequality (see Loève (1963), page 245). □

Proof of Theorem 4. By (7), we have that

$$n \Pr(|X| > Aa(n)) \leqq D \Pr(|X| > Aa(n))/\Pr(|X| > a(n))$$

where the constant D does not depend on A. Thus, the "if" part of Theorem 4 immediately follows from Lemma 4 and assumption A2.

Conversely, suppose $\{(T_n - c_n)/a(n)\}_{n=1}^{\infty}$ is tight. It is not difficult to show A2 (cf., (8)) if

$$\limsup_{n\to\infty} n \Pr(|X| > Aa(n)) \to 0 \qquad \text{as} \quad A \to \infty .$$

But then A2 follows from Lemma 3. It remains to show that tightness implies A1, i.e., that X is outside the domain of partial attraction of a normal variable. If X has a finite second moment, the central limit theorem says $T_n - ET_n$ should be scaled by the factor $1/n^{\frac{1}{2}}$ in order to obtain tightness. But when X has a finite second moment

$$n \Pr(|X| > \varepsilon n^{\frac{1}{2}}) \to 0 \qquad \text{as} \quad n \to \infty , \qquad \varepsilon > 0 . \tag{62}$$

Since $\liminf_{n\to\infty} n \Pr(|X| > a(n)) > 0$, $a(n)/n^{\frac{1}{2}} \to 0$ as $n \to \infty$. This precludes tightness for $\{(T_n - c_n\}/a(n)\}$ no matter how the c_n's are chosen. The same kind of contradiction arises if X is any random variable in the domain of partial attraction of a normal variable. The details can be worked out from the normal convergence criterion. (See Gnedenko and Kolmogorov (1954), page 128. The limit (62) corresponds to the first condition of their Theorem 2. In turn, if B_n is a suitable scaling factor (i.e., factor to multiply by), then $\liminf_n a(n)B_n = 0$. This implies $\{(T_n - c_n)/a(n)\}_{n=1}^{\infty}$ cannot be tight.) Thus, tightness implies A1. □

PROOF OF THEOREM 5. Inequality (49) only needs to be demonstrated for $A \geqq \frac{1}{6}$. For such A, (11) implies

$$\Pr(|X| > 6Aa(n)) \geqq \varepsilon \Pr(|X| > a(n)) , \qquad n \geqq 1 , \tag{63}$$

for some $\varepsilon > 0$, depending on A. Since (8) holds, (49) follows from (63) and Lemma 3. □

PROOF OF THEOREM 6. Suppose $Z \in \mathscr{D}_X$. Then there exist sequences of constants c_n and $d(n) > 0$ such that, on some subsequence $n' \to \infty$,

$$\frac{T_{n'} - c_{n'}}{d(n')} \to_L Z , \tag{64}$$

where $T_n = X_1 + \cdots + X_n$ ($n \geqq 1$). Let $a(\cdot)$ be a norming function for X and consider the ratios $a(n')/d(n')$. Since Z is not a degenerate random variable it follows from Theorem 4 that $a(n')/d(n')$ is bounded away from zero. In turn, (64) implies that $\{(T_{n'} - c_{n'})/a(n')\}$ is a tight sequence of random variables. Thus, for some $\alpha \geqq 0$ and real β and for some subsequence $n'' \to \infty$ in $\{n'\}$,

$$\frac{T_{n''} - c_{n''}}{a(n'')} \to_L \alpha Z + \beta .$$

But $\alpha = 0$ is impossible by Theorem 5. Thus $\alpha > 0$. Moreover, by (14), we have

$$\frac{\sum_{i=1}^{n''} Y_i}{b(n'')} - h(n'') - \frac{c_{n''}}{a(n'')} \to_L \alpha Z + \beta .$$

Thus, $Z \in \mathscr{D}_Y$, and we have shown that $\mathscr{D}_X \subset \mathscr{D}_Y$. By the same argument, $\mathscr{D}_Y \subset \mathscr{D}_X$. Hence, $\mathscr{D}_X = \mathscr{D}_Y$. □

It remains to prove the equivalence of A1′, A1″ and A1‴.

PROPOSITION 1. *Assumptions* A1′ *and* A1″ *are equivalent.*

PROOF. A1′ is equivalent to

$$h(x) \equiv \frac{x \Pr(|X| > x)}{\int_0^x v \Pr(|X| > v)\, dv} - \frac{\varepsilon}{x} \geqq 0\,, \qquad x > x_0\,,$$

for some $\varepsilon > 0$ and some $x_0 \geqq 0$. The value x_0 can be taken to be zero. For if $0 < x \leqq x_0$, then

$$h(x) \geqq \frac{x \Pr(|X| > x_0)}{\int_0^x v\, dv} - \frac{\varepsilon}{x} = \frac{2 \Pr(|X| > x_0) - \varepsilon}{x} \geqq 0$$

when ε is small enough. Since $h(x)$ is the derivative of

$$H(x) \equiv \log \{x^{-\varepsilon} \textstyle\int_0^x v \Pr(|X| > v)\, dv\}$$

whenever $\Pr(|X| = x) = 0$, it follows that A1′ is equivalent to saying that $H(x)$ is increasing. Thus A1′ is equivalent to A1″. □

PROPOSITION 2. *Assumptions* A1′ *and* A1‴ *are equivalent.*

PROOF. Assume A1′. Then A1″ holds as well, and, for some large M and for each $y > x > 0$,

$$\begin{aligned} \tfrac{1}{2} M y^2 \Pr(|X| > y) &\geqq \textstyle\int_0^y v \Pr(|X| > v)\, dv \\ &\geqq \left(\frac{y}{x}\right)^{\varepsilon} \textstyle\int_0^x v \Pr(|X| > v)\, dv \\ &\geqq \left(\frac{y}{x}\right)^{\varepsilon} \textstyle\int_0^x v\, dv \Pr(|X| > x) \\ &= \tfrac{1}{2} y^{\varepsilon} x^{2-\varepsilon} \Pr(|X| > x)\,. \end{aligned}$$

Thus A1′ implies A1‴. Conversely, if A1‴ holds, then for each $y > 0$,

$$\begin{aligned} \frac{y^2 \Pr(|X| > y)}{\int_0^y x \Pr(|X| > x)\, dx} &\geqq \frac{y^2 \Pr(|X| > y)}{\int_0^y x^{\varepsilon - 1} \cdot M y^{2-\varepsilon} \Pr(|X| > y)\, dx} \\ &= \frac{\varepsilon}{M} > 0\,. \end{aligned}$$

Thus A1‴ implies A1′. □

It is clear from the way that ε arises in the proof of Proposition 1, that the M appearing in assumption A1‴ can be taken to be as small as $2/\varepsilon$. Also, it is easily seen that the ε appearing in assumptions A1″ and A1‴ has to be smaller than two.

Acknowledgment. The authors wish to express their appreciation to the referee who carefully read the original manuscript and made many helpful comments and suggestions. His efforts have made us more cognizant of the modern approach to the subject of regularly varying functions. We trust that many readers of this revision will benefit from his efforts in the same way.

REFERENCES

[1] CSÖRGÖ, M. and RÉVÉSZ, P. (1975). A new method to prove Strassen type laws of invariance principle. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **31** 255–259.

[2] DOEBLIN, W. (1940). Sur l'ensemble de puissances d'une loi de probabilité. *Studia Math.* **9** 71–96.

[3] FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications* Vol. II. Wiley, New York.

[4] FELLER, W. (1968). An extension of the law of the iterated logarithm to variables without variance. *J. Math. Mech.* **18** 343–355.

[5] GNEDENKO, B. V. (1940). Some theorems on the powers of distribution functions. *Učen, Zap. Moskov. Gos. Univ. Mat.* **45** 61–72 (in Russian).

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

[7] HEYDE, C. C. (1967). On large deviation problems for sums of random variables which are not attracted to the normal law. *Ann. Math. Statist.* **38** 1575–1578.

[8] HEYDE, C. C. (1969). A note concerning behavior of iterated logarithm type. *Proc. Amer. Math. Soc.* **23** 85–90.

[9] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent RV's and CF. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** 111–131.

[10] LÉVY, PAUL (1954). *Théorie de l'Addition des Variables Aléatoires*, 2nd ed. Gauthier-Villars, Paris.

[11] LOÈVE, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, Princeton.

[12] PHILIPP, W. and STOUT, W. (1975). Almost sure invariance principles for partial sums of weakly dependent random variables. *Amer. Math. Soc. Mem.* No. 161, Providence.

[13] ROOT, D. and RUBIN, H. (1973). A probabilistic proof of the normal convergence criterion. *Ann. Probability* **1** 867–869.

[14] SENETA, E. (1976). *Regularly Varying Functions*, Springer, Berlin.

[15] SIMONS, G. and STOUT, W. (1975). A weak invariance principle with applications to domains of attraction. Institute of Statistics Mimeo Series No. 1044, Univ. of North Carolina at Chapel Hill.

[16] STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** 211–226.

[17] STRASSEN, V. (1965). Almost sure behavior of sums of independent random variables and martingales. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2** 315–343.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL
CHAPEL HILL, NORTH CAROLINA 27514

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
CHAMPAIGN-URBANA, ILLINOIS 61801