J. Austral. Math. Soc. (*Series A*) **30** (1981), 264–277

SOME PROPERTIES OF STOCHASTIC COMPACTNESS

R. A. MALLER

(Received 17 September 1979; revised 2 April 1980)

Communicated by R. L. Tweedie

Abstract

The aim of this paper is to show that some of the known properties of distributions in the domain of attraction of a stable law have counterparts for distributions which are stochastically compact in the sense of Feller. This enables us to unify the ideas of Feller and Doeblin, who first studied the concept of stochastic compactness, and give new characterizations of stochastic compactness and the domain of attraction of the normal distribution.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 60 F 05, 60 E 07, 60 G 50.

1. Results

Let X, X_i be independent and identically distributed random variables with distribution F. Suppose $P(|X| > x) > 0$ for $x > 0$. The terminology *stochastic compactness* was introduced by Feller (1965–66) to describe the following property of $S_n = X_1 + X_2 + \cdots + X_n$: there are sequences A_n and B_n with $B_n > 0$, $B_n \to +\infty$, such that, for every sequence $n'' \to +\infty$ of integers, there is a subsequence $n' \to +\infty$ for which $(S_{n'}/B_{n'}) - A_{n'}$ converges in distribution to a nondegenerate random variable. This is a generalization of the idea of attraction of normed and centred sums. Feller obtained elegant characterizations of stochastic compactness in terms of upper bounds on the tail of F and the truncated second moment $V(x) = \int_{-x}^{x+} u^2\, dF(u)$, which generalize regularly varying properties known to characterize the domains of attraction.

The concept of stochastic compactness had already been studied by Doeblin (1940, 1947), who called it 'compactness', but the simplifying ideas of the theory of regular variation, which Feller used to such good effect, were not known to

https://doi.org/10.1017/S144678870001716X Published online by Cambridge University Press

him. Doeblin introduced a further concept of 'strong compactness' and gave a characterization and some properties of it. One of the aims of the present paper is to elucidate Doeblin's results and relate them to those of Feller. We show in fact that 'compactness' and 'strong compactness' are the same, and thus obtain a unification of Doeblin's and Feller's ideas.

There has been much interest recently in the theory and application of stochastic compactness. A first application was given by Feller (1965–66), who proved a local limit theorem for stochastically compact sequences. A paper of Simons and Stout (1978) discussed stochastic compactness in relation to weak invariance principles, de Haan and Ridder (1979) extend the stochastic compactness idea to sample extremes, while Smythe (1974) and Thompson and Owen (1972) give other applications.

As regular variation is connected with domains of attraction, so a one-sided version of regular variation called by Feller (1969) 'dominated variation' is connected with stochastic compactness. Generalized regular variation was considered in the 1930's by Karamata, and has recently been reviewed and extended by Seneta (1976). Dominated variation has been connected with subexponentiality in the works of Goldie (1977, 1978); see also Embrechts *et al.* (1979). The mathematical duality between stochastic compactness and one-sided regular variation means that results obtained in one context apply in the other.

If F is stochastically compact, write $F \in SC$. The following are equivalent:

$$F \in SC; \tag{1.1}$$

$$\limsup_{x\to+\infty} x^2 P(|X| > x)/V(x) < +\infty; \tag{1.2}$$

$$\limsup_{x\to+\infty} V(x\lambda)/V(x) \leqslant c\lambda^{2-\alpha} \quad \text{for } \lambda \geqslant 1 \text{ for some } c \geqslant 1 \text{ and } \alpha \in (0, 2]; \tag{1.3}$$

$$\limsup_{x\to+\infty} V(x\lambda_0)/V(x) < \lambda_0^2 \quad \text{for some } \lambda_0 > 1. \tag{1.4}$$

The equivalence of (1.1), (1.2) and (1.3) follows from Theorem 2 and the theorem on page 387 of Feller (1965–66) (note that Feller's Theorem 2 is in error by the omission of an arbitrarily small $\varepsilon > 0$). Condition (1.3) is not actually Feller's; he uses instead $V(x\lambda)/V(x) \leqslant c\lambda^{2-\gamma}$ for $x > 1$ and $\lambda > \tau$ for some $c > 1$, $\tau > 1$ and $\gamma > 0$. However this uniform bound is implied by (1.3); one way to show this is by an argument like that of Letac (1970); see Seneta (1976, p. 97) (see also de Haan and Ridder (1979, p. 300) and Maller (1979)). We prefer to work with (1.3) since it is the direct generalization of the condition $V(x\lambda)/V(x) \to \lambda^{2-\alpha}$ for $\lambda > 0$, which is necessary for F to be in the domain of attraction of a stable law (Feller (1971, p. 577)); also necessary for this is $x^2P(|X| > x)/V(x) \to (2-\alpha)/\alpha$, see (1.2). The equivalence of (1.3) and (1.4) follows from regular variation theory (Feller (1969), Goldie (1977, p. 775)); note that Goldie's *indices*

https://doi.org/10.1017/S144678870001716X Published online by Cambridge University Press

of variation, and their properties due to Matuszewska, can be used to describe stochastic compactness; see Maller (1979).

We remark that, if (1.3) holds, it need not hold with $c = 1$. An example, due to de Haan, taking $V(x)$ equal to $e^{[\ln x]}$ for large x, shows this. However, de Haan and Ridder (1979) show that, under certain smoothness conditions on F, c may be taken as 1 in (1.3).

A second aim of this paper is to show that some of the well-known properties of distributions in a domain of attraction have counterparts for stochastically compact distributions. We give, for example, an analogue of the fact that the stable distributions are in their own domains of attraction.

In Theorem 6.2 de Haan and Ridder (1979) show that if

$$\lim_{\lambda \to +\infty} \limsup_{x \to +\infty} P(|X| > x\lambda)/P(|X| > x) = 0$$

then $F \in SC$ (their assumption of symmetry is easily seen to be unnecessary). (This strengthens a result of page 309 of Simons and Stout (1978); see also Siegel (1978), Theorem 3.2.) It can be shown by a similar method to that of de Haan and Ridder that if

$$\limsup_{x \to +\infty} P(|X| > x\lambda)/P(|X| > x) \leqslant c\lambda^{-2}$$

whenever $\lambda \geqslant 1$, for some $c \geqslant 1$, then $F \in D(2)$, the *domain of attraction of the normal distribution*. The converses of these results are not true (de Haan and Ridder (1979); Maller (1980)).

However, conditions on the tail function $P(|X| > x)$ can be necessary for stochastic compactness when F is not in *the domain of partial attraction of the normal distribution* (written $F \notin D_p(2)$) as was noticed by Simons and Stout (1978). Using their result on page 309 we can show that if $F \notin D_p(2)$, the following are equivalent:

(1.5) $$F \in SC;$$

(1.6) $$\lim_{\lambda \to +\infty} \limsup_{x \to +\infty} P(|X| > x\lambda)/P(|X| > x) < 1;$$

(1.7) *there are constants $c \geqslant 1$ and $\alpha > 0$ for which*

$$\limsup_{x \to +\infty} P(|X| > x\lambda)/P(|X| > x) \leqslant c\lambda^{-\alpha} \quad \textit{whenever } \lambda \geqslant 1.$$

The equivalence of (1.6) with (1.7) is again proved by standard methods of regular variation theory. We remark that $F \in D_p(2)$ if and only if

$$\liminf_{x \to +\infty} x^2 P(|X| > x)/V(x) = 0;$$

see Lévy (1937, p. 113) (see also Maller (1980, Theorem 1)).

It was shown in the last-mentioned paper that, if $F \in D_p(2)$, then

$$\liminf_{x \to +\infty} V(x\lambda)/V(x) = 1 \quad \text{for } \lambda \geqslant 1,$$

but that the converse is not true. Nevertheless the following useful result holds:

THEOREM 1.

(1.8) *If* $F \in SC$ *and* $\liminf_{x \to +\infty} V(x\lambda)/V(x) = 1$ *for* $\lambda > 1$ *then* $F \in D_p(2)$.

Let us call the random variables obtained as the weak limits of $(S_{n'}/B_{n'}) - A_{n'}$ for subsequences n', the *subsequential limit random variables* and their distributions the *subsequential limit distributions*. From Gnedenko and Kolmogorov (1968, pp. 70 and 116), these are infinitely divisible distributions whose characteristic functions have the Lévy canonical representation

$$(1.9)\qquad \log \xi(t) = i\gamma t - \tfrac{1}{2}\sigma^2 t^2 - \int_0^{\infty}\left[e^{itx} - 1 - itx/(1+x^2)\right]dN(x) + \int_{-\infty}^{0}\left[e^{itx} - 1 - itx/(1+x^2)\right]dM(x),$$

where σ^2 is the *normal component* of ξ and $N(x)$ and $M(-x)$ are *canonical measures*; they are nonincreasing functions on $(0, \infty)$ with

$$N(+\infty) = M(-\infty) = 0 \quad \text{and} \quad \int_{0^+}^{1} u^2|dN(u) + dM(-u)| < +\infty.$$

The distribution is nondegenerate if and only if either $\sigma^2 > 0$ or $N(x) + M(-x) > 0$ for some $x > 0$.

We now state our results on the limits of distribution on SC:

THEOREM 2. *The following are equivalent*:

(1.10) $F \in SC$ *and none of the subsequential limits has finite variance*;

(1.11) $F \in SC$ *and none of the subsequential limits is a normal distribution*;

(1.12) $F \in SC$ *and none of the subsequential limits has a positive normal component*;

(1.13) (1.3) *holds and* $\liminf_{x \to +\infty} V(x\lambda_0)/V(x) > 1$ *for some* $\lambda_0 > 1$.

According to (1.11) we can describe the situation in Theorem 2 as $F \in SC(\alpha) - D_p(2)$, $SC(\alpha)$ being the subclass of SC which satisfies (1.3) for the particular value of α. Clearly $SC(\alpha) - D_p(2)$ is characterized by (1.7) in conjunction with $F \notin D_p(2)$, and then $\alpha < 2$ (see the proof of (1.13)). $SC(\alpha) - D_p(2)$ is the analogue of the class of distributions attracted to nonnormal stable laws.

THEOREM 3.

(1.14) *If* $F \in SC$ *then* $I \in SC$, *where* I *is any subsequential limit distribution,*

(1.15) *if* $F \in SC(\alpha) - D_p(2)$ *then* $I \in SC(\alpha) - D_p(2)$.

and

(1.16) $F \in SC(\alpha) - D_p(2)$ *if and only if* $F \in SC$, $P(|X_I| > x) \leq cx^{-\alpha}$ *for* $x \geq 1$, *and* $I \notin D_p(2)$, *where* I *is any subsequential limit distribution,* X_I *has distribution* I, *and* c *is a constant independent of* I.

THEOREM 4. *$F \in SC$ and each subsequential limit distribution is normally distributed if and only if $F \in D(2)$.*

Doebelin (1940, 1947) introduced the term 'strong compactness' to describe stochastically compact distributions whose subsequential limit distributions are themselves stochastically compact. But we see from Theorem 3 that all stochastically compact distributions have this property, so the term 'strong compactness' is redundant. This result illustrates the power of the regular variation-like methods introduced by Feller. The equivalence of (1.1) and (1.2) was not discovered by Doeblin, and his discontinuous subsequential limit (Doeblin (1940, p. 89)) could not come from a distribution in SC. Following the proof of Theorem 4, in fact, we show that all subsequential limit distributions are absolutely continuous.

Doeblin (1940, Theorem X) gave necessary and sufficient conditions for F to be strongly compact, which we can exploit to obtain a new characterization of SC. Let $U^2(x) = V(x)/P(|X| > x)$ (recall that $P(|X| > x) > 0$ for $x > 0$).

THEOREM 5. *If*

$$\lim_{\lambda \to +\infty} \limsup_{x \to +\infty} P(|X| > \lambda U(x))/P(|X| > x) < 1,$$

then $F \in SC$, *while if* $F \in SC$,

$$\lim_{\lambda \to +\infty} \limsup_{x \to +\infty} P(|X| > \lambda U(x))/P(|X| > x) = 0.$$

Theorem 5 has the following counterpart for convergence to normality, whose proof, being similar to that of Theorem 5, is omitted:

THEOREM 6. *$F \in D(2)$ if and only if*

$$\lim_{x \to +\infty} P(|X| > \lambda U(x))/P(|X| > x) = 0 \quad \textit{for } \lambda > 0.$$

The characterization of $D(2)$ in Theorem 6 may be compared with those due to Lévy (1937) ($x^2 P(|X| > x)/V(x) \to 0$) and Feller (1971) ($V$ is slowly varying).

We conclude by mentioning that a characterization of the class of subsequential limit distributions of members of SC has not yet been obtained. Further unsolved problems include the conjecture that (1.10) and (1.12), with 'none' replaced by 'each', are equivalent to $SC(2)$.

2. Proofs

We use the abbreviation $H(x) = P(|X| > x)$ throughout. We can then write

$$V(x) = -\int_0^x u^2\, dH(u).$$

PROOF OF THEOREM 1. Suppose $F \notin D_p(2)$, so $x^2H(x)/V(x) \geqslant a > 0$ for $x \geqslant x_0$, or some a and $x_0 > 0$. By Seneta (1976, p. 97), (1.3) implies that, if $0 < \varepsilon < \alpha$, there is a $\lambda_0 > 1$ such that $V(x\lambda)/V(x) \leqslant \lambda^{2-\alpha+\varepsilon}$ whenever $\lambda \geqslant \lambda_0$ and $x \geqslant x_0$, if x_0 is large enough. We can assume λ_0 so large that

$$2\int_{\lambda_0}^{\infty} u^{-1-\alpha+\varepsilon}\, du = 2\lambda_0^{-\alpha+\varepsilon}/(\alpha - \varepsilon) < a.$$

Integrating by parts shows that

$$H(x) = -x^{-2}V(x) + 2\int_x^{\infty} u^{-3}V(u)\, du.$$

Now if $\liminf_x V(x\lambda)/V(x) = 1$ for $\lambda \geqslant 1$, we can take a sequence $x_i \to +\infty$ such that $V(x_i\lambda_0)/V(x_i) \to 1$; this means $V(x_i\lambda)/V(x_i) \to 1$ for $1 \leqslant \lambda \leqslant \lambda_0$ by monotonicity. Then by dominated convergence

$$\begin{aligned} x_i^2H(x_i)/V(x_i) &= 2\int_1^{\lambda_0} u^{-3}[V(ux_i)/V(x_i) - 1]\,du \\ &\quad + 2\int_{\lambda_0}^{\infty} u^{-3}[V(ux_i)/V(x_i) - 1]\,du \\ &\leqslant o(1) + 2\int_{\lambda_0}^{\infty} u^{-3}u^{2-\alpha+\varepsilon}\, du < o(1) + a, \end{aligned}$$

giving a contradiction which proves the theorem.

PROOF OF THEOREM 2. Suppose (1.11) does not hold; then $F \in D_p(2)$, and by Maller (1979a), $\liminf_x V(x\lambda)/V(x) = 1$ for $\lambda \geqslant 1$, which contradicts (1.13). Conversely suppose (1.13) does not hold; then $\liminf_{x\to+\infty} V(x\lambda)/V(x) = 1$ for $\lambda \geqslant 1$ so by Theorem 1, $F \in D_p(2)$, which contradicts (1.11). Thus (1.13) and (1.11) are equivalent (note that, by a standard argument, $\liminf_x V(x\lambda_0)/V(x) > 1$ for some $\lambda_0 > 1$ implies $V(x\lambda)/V(x) \geqslant \lambda^{\varepsilon}$ for some $\varepsilon > 0$ and x, λ, large enough, so (1.3) cannot hold with $\alpha = 2$). The remaining equivalences now follow from Theorem 3 of Maller (1980).

PROOF OF THEOREM 3. Let $F \in SC$ and let a limiting distribution of $(S_{n'}/B_{n'}) - A_{n'}$ be the infinitely divisible distribution I having canonical components σ^2, $N(x)$ and $M(-x)$. We have to show $I \in SC$. We can obviously assume that F has infinite variance. By Gnedenko and Kolmogorov (1968, p. 116), noting that our $N(x)$ is the negative of theirs, F must satisfy

$$n'[1 - F(xB_{n'})] \to N(x), \quad n'F(-xB_{n'}) \to M(-x)$$

and

$$\lim_{\varepsilon \to 0+} \lim \binom{\sup}{\inf} n' B_{n'}^{-2} V(\varepsilon B_{n'}) = \sigma^2,$$

for $x > 0$, using the fact that

$$\left[\int_{-x}^{x} u dF(u)\right]^2 = o[V(x)] \qquad (x \to +\infty),$$

(see Gnedenko and Kolmogorov (1968, p. 173)). In order that I be stochastically compact, it suffices by Gnedenko and Kolmogorov (1968, p. 88) that there is a $C_k \to +\infty$ such that every $k'' \to +\infty$ contains a subsequence k' for which

$$k'N(xC_{k'}) \to N_1(x), \qquad k'M(-xC_{k'}) \to M_1(-x)$$

and

$$\lim_{\varepsilon \to 0+} \lim_{k' \to +\infty} k' C_{k'}^{-2}\left[\sigma^2 - \int_0^{\varepsilon C_{k'}} u^2 \, dT(u)\right] = \tau^2,$$

where τ^2, $N_1(x)$ and $M_1(-x)$, are components of a nondegenerate infinitely divisible distribution, and $T(x) = N(x) + M(-x)$. Convergence is at points of continuity of the limits.

Since $y^{-2}\int_0^y uT(u)\,du \to 0$ as $y \to +\infty$, T being a canonical measure, and since $\sigma^2 + 2\int_0^y uT(u)\,du > 0$ for $y > 0$, I not being degenerate, we can define a sequence $C_k \uparrow +\infty$ by

$$C_k = \inf\left\{ y > 0 : \sigma^2 + 2\int_0^y uT(u)\,du > k^{-1}y^2 \right\};$$

this means $C_k^2 = k[\sigma^2 + 2\int_0^{C_k} uT(u)\,du]$. Take any $k'' \to +\infty$; by an extended version of Helly's theorem there is a subsequence k' of k'' for which

$$k'N(xC_{k'}) \to N_1(x), \qquad k'M(-xC_{k'}) \to M_1(-x)$$

and

$$k' C_{k'}^{-2}\left[\sigma^2 + 2\int_0^{xC_{k'}} uT(u)\,du\right] \to W(x),$$

for some nonincreasing $N_1(x)$ and $M_1(-x)$, and nondecreasing $W(x)$, on $x > 0$. We have to verify that N_1 and M_1 are canonical measures, that is, that $N_1(+\infty)$ and $M_1(-\infty) = 0$, and that $\int_0^1 u^2|dN_1(u)|$ and $\int_0^1 u^2|dM_1(-u)|$ are finite, equivalently, $uN_1(u)$ and $uM_1(-u)$ are integrable at 0.

Note that if $x > 0$ is a continuity point of T and $0 < \varepsilon < x$,

$$\limsup_{n'} n' B_{n'}^{-2} V(xB_{n'}) \leqslant \limsup_{n'} n' B_{n'}^{-2} V(\varepsilon B_{n'}) + \limsup\left[-\int_\varepsilon^x u^2 n'\, dH(uB_{n'})\right]$$
$$= \limsup_{n'} n' B_{n'}^{-2} V(\varepsilon B_{n'}) - \int_\varepsilon^x u^2\, dT(u)$$

since we can obviously apply dominated convergence to the second integral. Now letting $\varepsilon \downarrow 0+$ shows that

$$\limsup n' B_{n'}^{-2} V(xB_{n'}) \leqslant \sigma^2 - \int_0^x u^2\, dT(u),$$

and a reverse inequality for lim inf can be similarly obtained. Thus

$$n' B_{n'}^{-2} V(xB_{n'}) \to \sigma^2 - \int_0^x u^2\, dT(u)$$

at continuity points of T. Integrating by parts then gives

$$2n' B_{n'}^{-2} \int_0^{xB_{n'}} uH(u)\, du \to \sigma^2 + 2\int_0^x uT(u)\, du.$$

Also we have by (1.2) and (1.3) that for some $c > 0$ and $\lambda \geqslant 1$,

$$\limsup_x \int_0^{\lambda x} uH(u)\, du \Big/ \int_0^x uH(u)\, du \leqslant \limsup_x \left[V(x\lambda) + x^2\lambda^2 H(x\lambda)\right] / V(x)$$
$$= \limsup_x \left[1 + x^2\lambda^2 H(x\lambda)/V(x\lambda)\right] V(x\lambda)/V(x) \leqslant c\lambda^{2-\alpha}.$$

This means, for $x \geqslant 1$ and $y > 0$,

$$\frac{\sigma^2 + 2\int_0^{xy} uT(u)\, du}{\sigma^2 + 2\int_0^{y} uT(u)\, du} = \lim_{n'} \frac{n' B_{n'}^{-2} \int_0^{xyB_{n'}} uH(u)\, du}{n' B_{n'}^{-2} \int_0^{yB_{n'}} uH(u)\, du}$$
$$\leqslant \limsup_z \frac{\int_0^{xz} uH(u)\, du}{\int_0^{z} uH(u)\, du} \leqslant cx^{2-\alpha}$$

so

$$W(x) = \lim_{k'} \left[\sigma^2 + 2\int_0^{xC_{k'}} uT(u)\, du\right] \Big/ \left[\sigma^2 + 2\int_0^{C_{k'}} uT(u)\, du\right] \leqslant cx^{2-\alpha}$$

when $x \geqslant 1$. Since T is nonincreasing, then

$$x^2 k' T(xC_{k'}) \leqslant k' C_{k'}^{-2}\left[\sigma^2 + 2\int_0^{xC_{k'}} uT(u)\, du\right] \to W(x),$$

showing that $T_1(x) \equiv \lim k' T(xC_{k'}) = N_1(x) + M_1(-x)$ is finite for $x > 0$ and $\leqslant cx^{-\alpha}$ for $x \geqslant 1$. Thus $N_1(x)$ and $M_1(-x) \to 0$ as $x \to +\infty$. Furthermore,

$$1 = W(1) = \lim k' C_{k'}^{-2}\left[\sigma^2 + 2\int_0^{C_{k'}} uT(u)\, du\right]$$
$$\geqslant \liminf \int_0^1 uk' T(uC_{k'}) du \geqslant \int_0^1 uT_1(u)\, du,$$

showing that $uT_1(u)$ is integrable at 0, so the same is true of N_1 and M_1.

https://doi.org/10.1017/S144678870001716X Published online by Cambridge University Press

It remains to verify that the distribution defined by τ^2 (which we take to be $W(0+)$), N_1 and M_1, is nondegenerate. Clearly if $x > 0$,

$$W(x) = \lim k' C_{k'}^{-2}\left\{\sigma^2 + 2\left[\int_0^{C_{k'}} + \int_{C_{k'}}^{xC_{k'}}\right] uT(u)\,du\right\} = 1 + \int_1^x uT_1(u)\,du,$$

so $W(0+) = 1 - \int_0^1 uT_1(u)\,du$. If T_1 had no points of increase then T_1 would equal $T_1(+\infty) = 0$, in which case $W(0+) = \tau^2 = 1 > 0$. Thus the limit is nondegenerate and $I \in SC$, so that (1.14) is proved.

We now prove (1.15). Let $F \in SC(\alpha) - D_p(2)$, where $\alpha < 2$, and suppose $(S_{n'}/B_{n'}) - A_{n'}$ converges in distribution to X_I, where X_I has the inf.div. distribution I with canonical components σ^2 ($= 0$, by Theorem 2), $N(x)$ and $M(-x)$. We aim to show

$$\limsup_x P(|X_I| > x\lambda)/P(|X_I| > x) \leqslant c\lambda^{-\alpha} \quad \text{for } \lambda > 1.$$

Letting X_i^s be a symmetrization of X_i (Feller (1971, p. 147)) and

$$S_n^s = X_1^s + X_2^s + \cdots + X_n^s,$$

we have from the same reference and the inequality $1 - e^{-x} \geqslant xe^{-x}$, $x > 0$, that

$$\begin{aligned} P(|X_I| > x) &= \lim P(|S_{n'} - A_{n'}B_{n'}| > xB_{n'}) \\ &\geqslant \tfrac{1}{2} \liminf P(|S_{n'}^s| > 2xB_{n'}) \\ &\geqslant \tfrac{1}{2} \liminf n'P(|X^s| > 2xB_{n'})\exp\left[-n'P(|X^s| > 2xB_{n'})\right] \\ &\geqslant \tfrac{1}{4} \liminf n'P(|X - m| > 2xB_{n'})\exp\left[-2n'P(|X| > xB_{n'})\right], \end{aligned}$$

where m is a median of X. Since $m = o(B_n)$, we thus have

$$P(|X_I| > x) \geqslant \tfrac{1}{4} T(2x)e^{-2T(x)}. \tag{2.1}$$

Also, by a truncation argument and Chebychev's inequality (Feller (1971, p. 231)), and using the fact that A_n may be chosen as

$$nB_n^{-1}\int_{-xB_n}^{xB_n} u\,dF(u)$$

(see Gnedenko and Kolmogorov (1968, p. 117)), we have

$$P(|S_n - A_nB_n| > xB_n) \leqslant x^{-2}B_n^{-2}nV(xB_n) + nP(|X| > xB_n)$$

so letting $n \to +\infty$ through the subsequence n' gives

$$\begin{aligned} P(|X_I| > x) &= \lim_{n'} P(|X_{n'} - A_{n'}B_{n'}| > xB_{n'}) \\ &\leqslant -x^{-2}\int_0^x u^2\,dT(u) + T(x) \end{aligned} \tag{2.2}$$

because $n' B_{n'}^{-2} V(xB_{n'}) \to -\int_0^x u^2\, dT(u)$, where $T(x) = N(x) + M(-x)$ (recall $\sigma^2 = 0$). Now we need the following inequalities: by (1.7),

$$T(\lambda x)/T(x) = \lim n' H(x\lambda B_{n'})/n' H(xB_{n'}) \leqslant \limsup_y H(y\lambda)/H(y) \leqslant c_1 \lambda^{-\alpha}; \tag{2.3}$$

by the fact that $F \notin D_p(2)$,

$$x^2 T(x) - \int_0^x u^2\, dT(u) = \lim_{n'} (xB_{n'})^2 H(xB_{n'})/V(xB_{n'}) \geqslant \liminf_y y^2 H(y)/V(y) \geqslant a, \tag{2.4}$$

and again since $F \notin D_p(2)$, by the remark following the proof of Theorem 1 of Maller (1980),

$$T(\lambda x)/T(x) = \lim_{n'} H(x\lambda B_{n'})/H(xB_{n'}) \geqslant \liminf_y H(y\lambda)/H(y) \geqslant b\lambda^{2b-2}, \tag{2.5}$$

where $b = a/(a+1)$. In all of these we keep $x > 0$ and $\lambda \geqslant 1$. Putting together (2.1)–(2.5) gives

$$P(|X_I| > x\lambda)/P(|X_I| > x) \leqslant c_2 \lambda^{-\alpha} e^{2T(x)},$$

where c_2 does not depend on I. Letting $x \to +\infty$ now gives

$$\limsup P(|X_I| > x\lambda)/P(|X_I| > x) \leqslant c_2 \lambda^{-\alpha} \quad \text{if } \lambda \geqslant 1,$$

since $T(+\infty) = 0$. Our result will follow from (1.7) if $I \notin D_p(2)$. But by (2.1)–(2.5),

$$\liminf P(|X_I| > x\lambda)/P(|X_I| > x) \geqslant c_3 \lambda^{2b-2} \quad \text{if } \lambda \geqslant 1,$$

and $I \notin D_p(2)$ follows from Theorem 1 of Maller (1980). This proves (1.15).

Now (2.2), (2.3) and (2.4) imply

$$P(|X_I| > x) \leqslant -x^{-2} \int_0^x u^2\, dT(u) + T(x) \leqslant (a^{-1} + 1) T(x) \leqslant c_1 T(1) x^{-\alpha} \leqslant c_4 x^{-\alpha}$$

where c_4 does not depend on I, because by (1.2),

$$T(1) = \lim n' H(B_{n'}) \leqslant d \limsup B_{n'}^2 H(B_{n'})/V(B_{n'}) \leqslant cd$$

where by Feller (1965–66, p. 380), $\limsup nB_n^{-2} V(B_n) \leqslant d$ for some finite d. This proves the first part of (1.16).

Finally, suppose $F \in SC$, $P(|X_I| > x) \leqslant cx^{-\alpha}$ for $x \geqslant 1$, and $I \notin D_p(2)$. Certainly I itself is not the normal distribution, so $F \notin D_p(2)$. By (2.1) and (2.4),

$$-\int_0^x u^2\, dT(u) \leqslant cx^{2-\alpha},$$

https://doi.org/10.1017/S144678870001716X Published online by Cambridge University Press

while

$$-\int_0^1 u^2\, dT(u) = \lim n' B_{n'}^{-2} V(B_{n'}) \geqslant d$$

for some $d > 0$ by Feller (1965–66, p. 380). Fix $\lambda > 1$. For every $n'' \to +\infty$ there is a subsequence $n' \to +\infty$ for which

$$n' B_{n'}^{-2} V(\lambda B_{n'}) \to -\int_0^\lambda u^2\, dT(u),$$

so

$$\limsup V(\lambda B_{n'})/V(B_{n'}) \leqslant c\lambda^{2-\alpha},$$

where c does not depend on I. This means $\limsup V(\lambda B_n)/V(B_n) \leqslant c\lambda^{2-\alpha}$. Now B_{n+1}/B_n is bounded above (Feller (1965–66, p. 387)), $B_{n+1}/B_n \leqslant d$ say, so if $x > B_1$ and $n(x)$ is such that $B_n \leqslant x < B_{n+1}$,

$$\limsup_x V(x\lambda)/V(x) \leqslant \limsup_x V(d\lambda B_n)/V(B_n) \leqslant cd^{2-\alpha}\lambda^{2-\alpha},$$

which is the required result. This completes the proof of (1.16).

Remarks. It is not difficult to show from the estimates used in the proof of Theorem 1 that, if $F \in SC(\alpha)$, then F has finite moments of all orders less than α, and the same is true of all subsequential limit distributions. That the latter are also absolutely continuous follows from an argument like Lemma 2 of Maller (1978), wherein the characteristic function of a limit distribution is shown to be absolutely integrable.

The following inequality also follows from the above estimates:

$$P(|S_n - A_n B_n| > xB_n) \leqslant cx^{-\alpha+\varepsilon}$$

for $\varepsilon > 0$, $n \geqslant n_0(\varepsilon)$ and $x \geqslant x_0(\varepsilon)$, if $F \in SC(\alpha)$. This may be compared with inequalities due to Thompson *et al.* (1971, Lemma 3.2) and Owen (1973).

Proof of Theorem 4. Suppose $F \in SC$ and every n'' contains a subsequence n' for which $(S_{n'}/B_{n'}) - A_{n'} \to N(0, (\sigma')^2)$ for some $\sigma' > 0$. Then by Gnedenko and Kolmogorov (1968, p. 128), ignoring terms like $[\int_{-x}^x u\, dF(u)]^2$, as we may, we have $n' B_{n'}^{-2} V(\lambda B_{n'}) \to (\sigma')^2$, so $V(\lambda B_{n'})/V(B_{n'}) \to 1$ if $\lambda > 0$. Thus $V(\lambda B_n)/V(B_n) \to 1$ for $\lambda > 0$. Since $B_{n+1} < dB_{n'}$ if $x > B_1$ and $n(x)$ is such that $B_n \leqslant x < B_{n+1}$, and $\lambda > 1$,

$$1 \leqslant \frac{V(\lambda x)}{V(x)} \leqslant \frac{V(\lambda d B_n)}{V(B_n)} \to 1$$

so $V(\lambda x)/V(x) \to 1$ for $\lambda > 0$, V is slowly varying, and $F \in D(2)$.

For the proof of Theorem 5, we require the following:

LEMMA 1. *If* $F \in SC$ *there are constants* $c_0 > 0$, $x_0 > 0$, *for which*

$$\liminf_{n\to+\infty} P(|S_n - A_n B_n| > x_0 B_n) \geq c_0.$$

PROOF OF LEMMA 1. If the condition of the lemma did not hold, for any sequence $x_i \downarrow 0$ there would be a sequence $m_j = m_j(i)$ of integers $\uparrow +\infty$ such that for some $\varepsilon_i \downarrow 0$,

$$P\big(|X_{m_j} - A_{m_j} B_{m_j}| > x_i B_{m_j}\big) \leq \varepsilon_i$$

whenever $j \geq j_0(i)$. Thus if $n_i = m_{j_0}(i)$,

$$P\big(|S_{n_i} - A_{n_i} B_{n_i}| > x_i B_{n_i}\big) \leq \varepsilon_i$$

for $i \geq 1$. By stochastic compactness there would be a subsequence n_i' of n_i for which

$$P\big(|S_{n_i} - A_{n_i} B_{n_i}| > x B_{n_i}\big) \to P(|X_I| > x) \quad \text{for } x > 0,$$

where X_I is nondegenerate. But clearly $P(|X_I| > x) = 0$ for $x > 0$, so X_I would be degenerate at 0, which is impossible.

PROOF OF THEOREM 5. Suppose $\lim_{\lambda\to+\infty} \limsup_{x\to+\infty} H(\lambda U(x))/H(x) < 1$ and there is a sequence $x_i \to +\infty$ for which $x_i^2 H(x_i)/V(x_i) = y_i^2 \to +\infty$. Then $y_i^2 U^2(x_i) = x_i^2$, so

$$1 = \frac{H(x_i)}{H(x_i)} = \frac{H(y_i U(x_i))}{H(x_i)} \leq \frac{H(\lambda U(x_i))}{H(x_i)} \leq \limsup_x \frac{H(\lambda U(x))}{H(x)}$$

where $\lambda > 1$ and i is large enough. Letting $\lambda \to +\infty$ now gives a contradiction.

Conversely, suppose $F \in SC$ and apply Lemma 1 to the symmetrized sum S_n^s, which is clearly also stochastically compact. Let c_0 and x_0 be the resulting constants. Suppose by way of contradiction that there are sequences $\lambda_n \to +\infty$, $x_n \to +\infty$ for which $H(\lambda_n U(x_n))/H(x_n) \geq \beta > 0$ and let k_n be the integer part of $c_0/[8H(x_n)]$. By truncating at $2x_n$,

$$\begin{aligned}\tfrac{1}{2}c_0 &\leq P\big(|S_{k_n}^s| > x_0 B_{k_n}\big) \leq x_0^{-2} k_n B_{k_n}^{-2} V(2x_n) + 2k_n H(x_n)\\ &\leq 2^\gamma x_0^{-2} k_n B_{k_n}^{-2} V(x_n) + \tfrac{1}{4}c_0\\ &\leq 2^{\gamma-2} c_0 x_0^{-2} B_{k_n}^{-2} U^2(x_n) + \tfrac{1}{4}c_0\end{aligned}$$

for some $\gamma < 2$ as a result of Feller's uniform bound. This means $U^2(x_n) \geq 2^{-\gamma} x_0^2 B_{k_n}^2$, so if $x > 0$ and n is large enough for $2^{-\gamma} x_0^2 \lambda_n^2 \geq x^2$, we deduce that

$$\beta \leq \frac{H(\lambda_n U(x_n))}{H(x_n)} \leq \frac{k_n H(x B_{k_n})}{k_n H(x_n)} \sim 8c_0^{-1} k_n H(x B_{k_n}).$$

By stochastic compactness k_n contains a subsequence for which $k_n H(xB_{k_n}) \to T(x)$, where $T(x)$ is a canonical measure; but then $T(x) > \beta c_0/8$, which is impossible since $T(+\infty) = 0$. This completes the proof.

REMARK. We mention that Kesten (1972) gives a necessary and sufficient condition for stochastic compactness in terms of the dispersion function of S_n, related to conditions of Doeblin's, but not closely related to the conditions of the present paper.

Acknowledgements

I am grateful to Dr. C. C. Heyde and Dr. E. Seneta for discussions on different aspects of stochastic compactness.

References

L. de Haan and G. Ridder (1979), 'Stochastic compactness of sample extremes'. *Ann. Probability* **7**, 290–303.

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

W. Doeblin (1947), 'Sur l'ensemble de puissances d'une loi de probabilité', *Ann. Sci. École Norm. Sup* **63**, 317–350.

P. Embrechts, C. M. Goldie and N. Veraverbeke (1979), 'Subexponentiality and infinite divisiblity', *Z. Wahrsch. Verw. Gebiete* **49**, 335–347.

W. Feller (1965–66), 'On regular variation and local limit theorems'. *Proc. V Berkeley Symp. Math. Stats. Prob.* II, part I, pp. 373–388.

W. Feller (1969), 'One sided analogues of Karamata's regular variation'. *Enseignement Math.* **15**, 107–121.

W. Feller (1971), *An introduction to probability theory and its applications II* (2nd ed., Wiley, New York).

B. V. Gnedenko and A. N. Kolmogorov (1968), *Limit distributions for sums of independent random variables* (2nd ed., Addison-Wesley, Reading, Mass.).

C. M. Goldie (1977), 'Convergence theorems for empirical Lorentz curves and their inverses', *Advances in Appl. Probability* **9**, 765–791.

C. M. Goldie (1978), 'Subexponential dsitributions and dominated variation tails', *J. Appl. Probability* **15**, 440–442.

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

G. Letac (1970), 'On slow variation', *Proc. Amer. Math. Soc.* **24**, 745–746.

P. Lévy (1937), *Théorie de l'addition des variables aléatoires* (Gauthier-Villars, Paris).

R. A. Maller (1978), 'A local limit theorem for independent random variables', *Stochastic Processes Appl.* **7**, 101–111.

R. A. Maller (1980), 'A note on domains of partial attraction', *Ann. Probability* **8**, 576–583.

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

W. L. Owen (1973), 'An estimate for $E|S_n|$ for variables in the domain of normal attraction of a stable law of index α, $1 < \alpha < 2$', *Ann. Probability* **1**, 1071–1073.

E. Seneta (1976), *Regularly varying functions* (Lecture Notes in Mathematics 508, Springer-Verlag, Berlin).

V. G. Siegel (1978), 'Zero-one laws and weak convergences for sums of independent random variables', *Math. Nachr.* **86**, 333–346.

G. Simons and W. Stout (1978), 'A weak invariance principle with applications to domains of attraction', *Ann. Probability* **6**, 294–315.

R. T. Smythe (1974), 'Sums of independent random variables on partially ordered sets', *Ann. Probability* **2**, 906–917.

M. E. Thompson, A. K. Basu and W. L. Owen (1971), 'On the existence of the optimal stopping rule in the S_n/n problem when the second moment is infinite', *Ann. Math. Statist.* **42**, 1936–1942.

M. E. Thompson and W. L. Owen (1972), 'The structure of the optimal stopping rule in the S_n/n problem', *Ann. Math. Statist.* **43**, 1110–1121.

CSIRO Division of Mathematics and Statistics
Perth, Western Australia