Probab. Th. Rel. Fields 83, 331–353 (1989)

Probability Theory and Related Fields

Almost Sure and Weak Invariance Principles for Random Variables Attracted by a Stable Law

I. Berkes[1,*] and H. Dehling[2]

[1] Mathematical Institute of the Hungarian Academy of Sciences, Reáltanoda u. 13–15, H-1053 Budapest, Hungary
[2] Groningen University, Department of Mathematics and Computing Science, P.O. Box 800, NL-9700 AV Groningen, The Netherlands

Summary. Let (X_n) be i.i.d. random variables belonging to the domain of normal attraction of a symmetric stable law with parameter $0<p<2$. We study the a.s. and weak approximation of the partial sum process $S(t) = \sum_{n \leq t} X_n (t \geqq 0)$ by a symmetric stable process $G_p(t)$. Stout proved an upper bound for the optimal remainder term in this approximation; we prove here a lower bound, leaving only a small gap between the upper and lower estimates. We also give a new method to obtain upper bounds. Finally, we prove analogues of these results in the case when a.s. approximation is replaced by approximation in probability.

§ 1. Introduction

Let $X_1, X_2, \ldots$ be independent r.v.'s with common symmetric distribution function $F(x)$ satisfying

$$1-F(x) \sim c x^{-\alpha} \quad \text{as } x \to +\infty\,^1 \tag{1.1}$$

where $c>0$, $0<\alpha<2$. By a well known theorem,

$$n^{-1/\alpha}(X_1 + \ldots + X_n) \xrightarrow{D} G_{\alpha,c} \tag{1.2}$$

where $G_{\alpha,c}$ is the symmetric stable distribution with characteristic function $\exp(-\rho|t|^\alpha)$, $\rho = 2c \int_0^\infty \sin y/y^\alpha\, dy$. Conversely, for i.i.d. r.v.'s $X_1, X_2, \ldots$ with sym-

* Research supported by Hungarian National Foundation for Scientific Research, grant no. 1808

[1] $f(x) \sim g(x)$, $f(x) \ll g(x)$, $f(x) \approx g(x)$ $(x \to \infty)$ mean, as usual, $\lim_{x\to\infty} f(x)/g(x) = 1$, $\overline{\lim}_{x\to\infty} |f(x)/g(x)| < +\infty$, $0 < \underline{\lim}_{x\to\infty} f(x)/g(x) \leqq \overline{\lim}_{x\to\infty} f(x)/g(x) < +\infty$, respectively

metric distribution function $F(x)$, (1.1) is necessary for (1.2). Let $Y_1, Y_2, \ldots$ be independent r.v.'s with common distribution function $G_{\alpha,c}$. The purpose of the present paper is to study the strong approximation of $\sum_{i=1}^{n} X_i$ by $\sum_{i=1}^{n} Y_i$. More precisely, we shall investigate to which functions $\varphi(n)$ holds

$$\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i = O(\varphi(n)) \qquad \text{a.s. as } n \to \infty \tag{1.3}$$

in the sense that the i.i.d. sequences $\{X_i\}$, $\{Y_i\}$ can be defined on a common probability space such that (1.3) is valid. We shall also investigate the analogous question when "a.s." in (1.3) is replaced by "in probability".

Write (1.1) in the form

$$1 - F(x) = c x^{-\alpha} + \beta(x) x^{-\alpha} \qquad x > 0 \tag{1.4}$$

where $\beta(x) \to 0$ as $x \to +\infty$. Stout proved in [13], under certain regularity conditions on $\beta(x)$, that if a_n is a positive numerical sequence satisfying

$$\sum_{n=1}^{\infty} a_n^{-\alpha} < +\infty \tag{1.5}$$

then the approximation

$$\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i = O(a_n |\beta(a_n)|) \qquad \text{a.s.} \tag{1.6}$$

holds. In [2] this result was improved to

$$\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i = O(a_n |\beta(a_n)|^{1/\alpha}) \qquad \text{a.s.} \tag{1.7}$$

for $\alpha < 1$. Theorem 1 of the present paper shows that if $\{c_n\}$ is a positive numerical sequence satisfying

$$\beta^2(c_N) \sum_{k \leqq N} c_k^{-\alpha} \to +\infty \qquad \text{as } N \to \infty \tag{1.8}$$

then almost surely

$$\left| \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i \right| \geqq K c_n |\beta(c_n)| \qquad \text{infinitely often} \tag{1.9}$$

for any versions of $\{X_i\}$ and $\{Y_i\}$ defined on a common probability space; K is a positive constant depending on α and c. It is easy to see (cf. the Remark after the proof of Theorem 1) that there exist sequences $\{a_n\}$, $\{c_n\}$ satisfying (1.5) and (1.8), respectively, such that $a_n/c_n \ll |\beta(n)|^{-3/\alpha}$. Hence, the gap between the upper estimates $a_n |\beta(a_n)|$, $a_n |\beta(a_n)|^{1/\alpha}$ and the lower estimate $c_n |\beta(c_n)|$ is a power of $|\beta(n)|$ and thus all three estimates are fairly precise if $\beta(x)$ tends to zero very slowly. Unfortunately, we were not able to determine the optimal remainder term $\varphi(n)$ in (1.3).

For $\beta(x)$ tending to zero sufficiently rapidly, (1.6) and (1.7) are useful invariance principles for limit theorems for i.i.d. r.v.'s in the domain of attraction of a stable law. For example, if $|\beta(x)| \leqq (\log x)^{-\gamma}$ for $\gamma > 1/\alpha$ then choosing $a_n = (n \log^{1+\varepsilon} n)^{1/\alpha}$ with a small enough $\varepsilon > 0$, the right side (1.6) becomes $O(n^{1/\alpha}(\log n)^{-\lambda})$ with $\lambda > 0$. This remainder term is strong enough to carry over most known limit theorems from $\{Y_n\}$ to $\{X_n\}$. On the other hand, Theorem 1 of the present paper shows that given any positive numerical sequence $\{b_n\}$ with $\sum b_n^{-\alpha} = +\infty$, the approximation

$$\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i = O(b_n) \quad \text{a.s.} \tag{1.10}$$

cannot hold if $\beta(x)$ tends to zero sufficiently slowly. For example, if $\beta(x) \geqq (\log x)^{-\gamma}$ for $\gamma < 1/(\alpha+2)$ then (1.10) cannot hold with $b_n = n^{1/\alpha}$. Hence for such $\beta(x)$ even the weak limit theorem (1.2) cannot be obtained via an a.s. invariance principle of the type (1.3). This difficulty can be overcome by adapting an idea of Major [9] and approximating $\sum_{i=1}^{n} X_i$ not by $\sum_{i=1}^{n} Y_i$ but by $\sum_{i=1}^{n} (Y_i + \tau_i)$ where τ_i are small "correcting" r.v.'s. A natural choice is $\tau_i = \varepsilon_i Y_i$ for some numerical sequence $\varepsilon_i \to 0$ and Mijnheer showed ([10], Examples 3 and 4) that for some classes of F, this choice leads to an improvement of Stout's remainder term in (1.6). A still better remainder term can be obtained by choosing the τ_n independent of the sequence $\{Y_n\}$. In fact, as is shown in [2], this choice leads to a remainder term $O(n^{1/\alpha - \lambda})$, $\lambda > 0$ for all positive β's such that $\beta(x) x^{-\alpha} \downarrow 0$. Thus we get an a.s. invariance principle applicable for arbitrarily slowly decreasing $\beta(x)$'s. In Sect. 2 we give a new, simplified proof of the just mentioned "perturbed" a.s. invariance principle of [2] and extend it to the case when $\beta(x)$ can take both positive and negative values. (For additional information on the remainder term in the case $\tau_n = \varepsilon_n Y_n$, ε_n numerical, see the remark at the end of Sect. 2.) In Sect. 3 we investigate the "in probability" analogue of (1.3). Simons and Stout showed in [12] that for any $\beta(x) \to 0$ the approximation

$$\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i = o(n^{1/\alpha}) \quad \text{in probability} \tag{1.11}$$

holds. In other words, there is a joint construction of $\{X_i\}$ and $\{Y_i\}$ and a numerical sequence $\varepsilon_n \to 0$ such that

$$P\left(n^{-1/\alpha} \left| \sum_{i=1}^{n} (X_i - Y_i) \right| \geqq \varepsilon_n \right) \leqq \varepsilon_n \qquad n = 1, 2, \dots \tag{1.12}$$

In Sect. 3 we give upper and lower estimates for the (optimal) ε_n in (1.12). As in the a.s. case, these estimates will be quite precise for slowly decreasing $\beta(x)$, the gap between the upper and lower estimates being again a power of $|\beta(n)|$.

The technical conditions made on $\beta(x)$ in our theorems could be weakened but we preferred to give conditions making the formulations and proofs of our theorems simple. (Note that most known results in the field, e.g. those in [5], [10], [13], are rather complicated.) For example, in each of Theorems 2, 3 and 4 we assume that $\beta(x)$ is positive for $x \geqq x_0$. This assumption makes the arguments particularly simple but it can be easily removed, as it is shown in the remark following the formulation of Theorem 2.

It is worth noting that all known a.s. invariance principles for r.v.'s in the domain of attraction of a stable law G_α, $0<\alpha<2$ use constructions with independent $X_n - Y_n$. (The argument in [2] is an exception but the result of [2] is reproved in this paper with i.i.d. $X_n - Y_n$, cf. the proof of Theorem 2.) That constructions of this simple type can give rather good estimates is in sharp contrast with the case of finite variances (Brownian approximation) where no nontrivial bound can be given using independent $X_n - Y_n$. The small gap between the upper bonds (1.6), (1.7) given by independent $X_n - Y_n$ and the lower bound (1.9) valid for all constructions shows that no essential improvement of the remainder term can be obtained by constructions using dependent $X_n - Y_n$. Whether independent $X_n - Y_n$ give actually the best remainder term remains open.

We finally note that Fisher [5] and Mijnheer [10] proved a.s. invariance principles for i.i.d. sequences $\{X_n\}$ in the general (non-normal) domain of attraction of a stable law G_α. Specifically, they showed that in this case the partial sums $\sum_{i=1}^{n} X_i$ can be approximated by weighted sums $\sum_{i=1}^{n} \lambda_i Y_i$ of i.i.d. r.v.'s Y_i having distribution G_α. (Clearly if X_n are outside of the domain of normal attraction of G_α then the upper classes belonging to $\sum_{i=1}^{n} X_i$ and $\sum_{i=1}^{n} Y_i$ are different and thus no nontrivial approximation of $\sum_{i=1}^{n} X_i$ by unweighted sums $\sum_{i=1}^{n} Y_i$ is possible.) Specialized to the case of the domain of normal attraction, Fisher's result reduces to Stout's theorem. Mijnheer's result, on the other hand, gives weighted approximation also in the domain of normal attraction; a comparison of this result with our theorems will be given in Sect. 2. No lower estimates are proved either in [5] or [10].

§ 2. A.s. Invariance Principles – Upper and Lower Bounds

Before we formulate our theorems, we note that the distribution function $G = G_{\alpha,c}$ of the stable limit distribution in (1.2) satisfies

$$1 - G(x) = c x^{-\alpha} + O(x^{-2\alpha}) \quad \text{as } x \to +\infty \tag{2.1}$$

(see e.g. [4] p. 549, Lemma 1.) As a matter of fact, all what we shall use about the i.i.d. sequences $\{X_n\}$ and $\{Y_n\}$ in our study of the approximation (1.3) is that their distribution functions F and G satisfy (1.1) and (2.1), respectively. Hence our results will cover more than just the case of exactly stable Y_n's.

Theorem 1. *Let F and G be symmetric distribution functions satisfying* (1.4) *and* (2.1), *respectively. Let $\{c_n\}$ be a positive numerical sequence such that $|\beta(c_n)|$ is non-increasing, $c_n|\beta(c_n)|$ is nondecreasing, $c_n^\alpha|\beta(c_n)| \to +\infty$ and*

$$\beta^2(c_N) \sum_{k \leqq N} c_k^{-\alpha} \to +\infty \qquad \text{as } N \to \infty. \tag{2.2}$$

Then for any i.i.d. *sequences $\{X_n\}$ and $\{Y_n\}$ defined on a common probability space with respective distribution functions F and G we have*

$$P\left\{\left|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i\right| \geqq \frac{1}{4\alpha c} c_n |\beta(c_n)| \text{ infinitely often}\right\} = 1. \tag{2.3}$$

The technical assumptions made on β and c_n in Theorem 1 could be weakened or dropped at the cost of changing (2.2) and (2.3) slightly. For example, the monotonicity of $|\beta(c_n)|$ can be dropped if in (2.2) we replace $\beta^2(c_N)$ by $\min_{1 \leqq k \leqq N} \beta^2(c_k)$. Similarly, the assumption that $c_n|\beta(c_n)|$ is nondecreasing can be weakened to $\liminf_{n\to\infty} |c_{n+1}\,\beta(c_{n+1})/c_n\,\beta(c_n)| > 0$ provided we replace the constant $(4\alpha c)^{-1}$ in (2.3) by another constant. Actually, the monotonicity of $c_n|\beta(c_n)|$ can be dropped completely if we replace the lower bound $(4\alpha c)^{-1} c_n|\beta(c_n)|$ in (2.3) by

$$(4\alpha c)^{-1} \min(c_n|\beta(c_n)|, c_{n+1}|\beta(c_{n+1})|).$$

Finally, $c_n^\alpha|\beta(c_n)| \to +\infty$ can be replaced by the condition $\sum c_n^{-2\alpha} < +\infty$, not containing β.

Proof of Theorem 1. Assume first $\beta(x) > 0$ $(x > 0)$. Then $\beta(c_n)$ is decreasing i.e.

$$\sum_{k \leqq N} \frac{\beta(c_k)}{c_k^\alpha} \geqq \beta(c_N) \sum_{k \leqq N} \frac{1}{c_k^\alpha}$$

and thus (2.2) implies

$$b_N = \left(\sum_{k \leqq N} \frac{\beta(c_k)}{c_k^\alpha}\right) \Bigg/ \left(\sum_{k \leqq N} \frac{1}{c_k^\alpha}\right)^{1/2} \to +\infty. \tag{2.4}$$

Let $\{X_n\}$ and $\{Y_n\}$ be i.i.d. sequences defined on a common probability space with distribution functions F and G, respectively. Let

$$d_k = \frac{1}{2\alpha c} c_k \beta(c_k)$$

$$\xi_k = I(X_k > c_k), \qquad \eta_k = I(Y_k > c_k - d_k)$$

where $I(\cdot)$ denotes the indicator function of the set in brackets. Using (1.4), (2.1), $\beta(c_k)\downarrow 0$ and $c_k^\alpha \beta(c_k)\to\infty$ we get

$$\begin{aligned}
E\xi_k &= P(X_k>c_k)=(c+\beta(c_k))\,c_k^{-\alpha}\\
E\eta_k &= P(Y_k>c_k-d_k)=c(c_k-d_k)^{-\alpha}+O((c_k-d_k)^{-2\alpha})\\
&= c\,c_k^{-\alpha}\left(1-\frac{d_k}{c_k}\right)^{-\alpha}+O(c_k^{-2\alpha})\\
&= c\,c_k^{-\alpha}\left(1+\alpha\frac{d_k}{c_k}+O\left(\left(\frac{d_k}{c_k}\right)^2\right)\right)+O(c_k^{-2\alpha})\\
&= c\,c_k^{-\alpha}+\tfrac{1}{2}\beta(c_k)\,c_k^{-\alpha}+O(c_k^{-\alpha}\beta^2(c_k))+O(c_k^{-2\alpha})\\
&= c\,c_k^{-\alpha}+\tfrac{1}{2}\beta(c_k)\,c_k^{-\alpha}+o(\beta(c_k)\,c_k^{-\alpha}).
\end{aligned}$$

Thus

$$E\xi_k-E\eta_k\sim\tfrac{1}{2}c_k^{-\alpha}\beta(c_k)\qquad\text{as } k\to\infty. \tag{2.5}$$

For the variances we find

$$\begin{aligned}
\mathrm{Var}(\xi_k)&\leqq P(X_k>c_k)\ll c_k^{-\alpha}\\
\mathrm{Var}(\eta_k)&\leqq P(Y_k>c_k-d_k)\ll c_k^{-\alpha}
\end{aligned}$$

and thus by the Chebisev inequality

$$P\{|\sum_{i\leqq N}(\xi_i-E\xi_i)|\geqq b_N^{1/2}(\sum_{i\leqq N}c_i^{-\alpha})^{1/2}\}\ll b_N^{-1}. \tag{2.6}$$

Let $\{N_k\}$ be a sequence of integers tending of infinity so rapidly that $b_{N_k}\geqq k^2$. Then (2.6) and the Borel-Cantelli lemma imply

$$\sum_{i\leqq N_k}\xi_i=\sum_{i\leqq N_k}E\xi_i+O(b_{N_k}^{1/2}(\sum_{i\leqq N_k}c_i^{-\alpha})^{1/2})\qquad\text{a.s.} \tag{2.7}$$

and similarly

$$\sum_{i\leqq N_k}\eta_i=\sum_{i\leqq N_k}E\eta_i+O(b_{N_k}^{1/2}(\sum_{i\leqq N_k}c_i^{-\alpha})^{1/2})\qquad\text{a.s.} \tag{2.8}$$

By (2.4) and (2.5) we have

$$\sum_{i\leqq N_k}(E\xi_i-E\eta_i)/(\sum_{i\leqq N_k}c_i^{-\alpha})^{1/2}b_{N_k}^{1/2}\to+\infty$$

and thus by (2.7) and (2.8) we find

$$\sum_{i\leqq N_k}(\xi_i-\eta_i)\sim\sum_{i\leqq N_k}(E\xi_i-E\eta_i)\to+\infty\qquad\text{a.s.} \tag{2.9}$$

where the validity of the second relation follows from (2.4) and (2.5). Since ξ_i and η_i are indicator variables, (2.9) implies $P(A_k \text{ i.o.})=1$ where

$$A_k=\{\xi_k=1\text{ and }\eta_k=0\}=\{X_k>c_k\text{ and }Y_k\leqq c_k-d_k\}$$

and thus

$$P\{X_k - Y_k > d_k \text{ infinitely often}\} = 1. \tag{2.10}$$

But $X_k - Y_k > d_k$ implies at least one of the inequalities

$$\left|\sum_{i=1}^{k} (X_i - Y_i)\right| > \tfrac{1}{2} d_k, \qquad \left|\sum_{i=1}^{k-1} (X_i - Y_i)\right| > \tfrac{1}{2} d_k$$

and since $d_k > d_{k-1}$ for $k \geqq k_0$ by our assumptions, (2.10) yields

$$P\left\{\left|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i\right| \geqq \tfrac{1}{2} d_n \text{ infinitely often}\right\} = 1$$

proving Theorem 1 in the case $\beta(x) > 0$.

Let us drop now the assumption $\beta(x) > 0$ and let $H_1 = \{n\colon \beta(c_n) > 0\}$, $H_2 = \{n\colon \beta(c_n) < 0\}$. Clearly (2.4) remains valid with $\beta(c_k)$ replaced by $|\beta(c_k)|$ and thus at least one of the sequences

$$b_N^{(H_1)} = \left(\sum_{\substack{k \leqq N \\ k \in H_1}} \frac{|\beta(c_k)|}{c_k^\alpha}\right) \Big/ \left(\sum_{\substack{k \leqq N \\ k \in H_1}} \frac{1}{c_k^\alpha}\right)^{1/2} \qquad N = 1, 2, \ldots$$

and

$$b_N^{(H_2)} = \left(\sum_{\substack{k \leqq N \\ k \in H_2}} \frac{|\beta(c_k)|}{c_k^\alpha}\right) \Big/ \left(\sum_{\substack{k \leqq N \\ k \in H_2}} \frac{1}{c_k^\alpha}\right)^{1/2} \qquad N = 1, 2, \ldots$$

is unbounded. If $b_N^{(H_1)}$ is unbounded then the proof given in the case $\beta(x) > 0$ remains valid provided we define d_k, ξ_k, η_k only for $k \in H_1$ and replace all sums $\sum_{k \leqq N}$, $\sum_{i \leqq N_k}$ in the proof by $\sum_{k \leqq N, k \in H_1}$, $\sum_{i \leqq N_k, i \in H_1}$, respectively. A similar argument holds if $b_N^{(H_2)}$ is unbounded. This completes the proof of Theorem 1 in the general case.

Remark. Since the lower estimate (2.3) is given only indirectly in terms of $\beta(x)$ (via the sequence $\{c_n\}$ depending on $\beta(x)$), it is not clear how large the gap is between (2.3) and the upper estimates (1.6), (1.7). To see this, we prove the following proposition where, for the sake of simplicity, we assume $\beta(x) > 0$.

Proposition. *Assume that $\beta(x) \downarrow 0$ and*

$$\beta(x^{1/\alpha}) \approx \beta(x) \qquad \text{as } x \to +\infty \tag{2.11}$$

$$1/\beta(x)^{2+\delta} \text{ is concave in } x \text{ for small enough } \delta > 0. \tag{2.12}$$

Then for every $\varepsilon > 0$ there exist positive sequences $a_n \uparrow \infty$, $c_n \uparrow \infty$ such that (1.5) *and* (2.2) *hold and*

$$\frac{a_n \beta(a_n)}{c_n \beta(c_n)} \ll \left(\frac{1}{\beta(n)}\right)^{(2/\alpha)+\varepsilon} \tag{2.13}$$

Remark. Conditions (2.11) (2.12) could be weakened but they cover typical cases like $\beta(x)=(\log x)^{-\lambda}$, $\beta(x)=(\log\log x)^{-\lambda}$, etc. The proposition shows that by choosing properly the sequence $\{c_n\}$ in Theorem 1 and $\{a_n\}$ in the upper estimate (1.6), the gap between the upper and lower estimates is at most a power of $\beta(n)$. A similar comparison holds between (1.9) and the upper estimate (1.7).

Proof of the Proposition. Let δ_0 be a small positive number to be chosen later and set $\gamma(x)=\beta(x)^{-(2+\delta_0)}$. By $\beta(x)\downarrow 0$ and (2.12), $\gamma(x)$ is concave for large x if δ_0 is small enough and $\gamma(x)\uparrow\infty$ as $x\uparrow\infty$. By concavity, $\gamma(x+1)-\gamma(x)$ is decreasing and thus the nonnegative finite limit $l=\lim_{x\to\infty}(\gamma(x+1)-\gamma(x))$ exists. $l>0$ is impossible since then $\gamma(x)\sim lx$ and $\beta(x)\sim(lx)^{-1/(2+\delta_0)}$ would follow, contradicting to (2.11). Thus

$$\gamma(x+1)-\gamma(x)\to 0 \quad \text{as } x\to+\infty. \tag{2.14}$$

Set

$$s_k=\min\{i:\gamma(i)\geqq k\}, \quad \rho(k)=s_k-s_{k-1}, \quad I_k=[s_{k-1}, s_k-1]$$

(2.14) and the concavity of $\gamma(x)$ imply that $\rho(k)\uparrow\infty$ and thus I_k is not empty for $k\geqq k_0$. Define two sequences $\{a_n\}$ and $\{c_n\}$ by

$$a_i=\begin{cases}\rho(k)^{1/\alpha}(k\log^2 k)^{1/\alpha} & \text{for } i\in I_k, k\geqq k_0\\ \lambda_0 & \text{otherwise}\end{cases}$$

$$c_i=\begin{cases}\rho(k)^{1/\alpha} & \text{for } i\in I_k, k\geqq k_0\\ \lambda_0 & \text{otherwise}\end{cases}$$

where λ_0 is chosen so small that $\lambda_0\leqq(k_0\log^2 k_0)^{1/\alpha}$. By $\rho(k)\uparrow\infty$, a_n and c_n tend monotonically to ∞. Obviously

$$\sum_{i\in I_k} a_i^{-\alpha}=(k\log^2 k)^{-1} \qquad k\geqq k_0$$

and thus (1.5) is true. To prove (2.2) assume $N\in I_k$ (i.e. $s_{k-1}\leqq N\leqq s_k-1$), then

$$\begin{aligned}\sum_{i\leqq N} c_i^{-\alpha}&\geqq\sum_{i\leqq s_{k-1}-1} c_i^{-\alpha}\geqq\sum_{k_0\leqq j\leqq k-1}\sum_{i\in I_j} c_i^{-\alpha}\\ &=\sum_{k_0\leqq j\leqq k-1} 1=k-k_0\geqq\gamma(s_k-1)-k_0\geqq\gamma(N)-k_0\\ &\geqq\tfrac{1}{2}\gamma(N)\end{aligned} \tag{2.15}$$

for $N\geqq N_0$. On the other hand, for $N\in I_k$, $N\geqq N_0$ we get, using (2.11)

$$\begin{aligned}\gamma(c_N)&=\gamma(\rho(k)^{1/\alpha})\ll\gamma(\rho(k))=\gamma(s_k-s_{k-1})\\ &\leqq\gamma(s_k)\leqq\gamma(s_{k-1})+2\leqq\gamma(N)+2\leqq 2\gamma(N)\end{aligned}$$

and thus (2.15) implies

$$\sum_{i\leqq N} c_i^{-\alpha}\geqq\tfrac{1}{4}\gamma(c_N)$$

whence we get

$$\beta^2(c_N) \sum_{i \leqq N} c_i^{-\alpha} \geqq \tfrac{1}{4}\beta(c_N)^{-\delta_0} \to +\infty$$

establishing (2.2). Finally, notice that for $N \in I_k$, $k \geqq k_0$ we have

$$a_N/c_N = (k \log^2 k)^{1/\alpha}, \quad k-1 \leqq \gamma(N) \leqq k$$

and thus

$$a_N/c_N \sim (\gamma(N) \log^2 \gamma(N))^{1/\alpha} \quad \text{as } N \to \infty.$$

Using the definition of $\gamma(x)$ and choosing δ_0 small enough, the last relation yields

$$\frac{a_N}{c_N} \ll \left(\frac{1}{\beta(N)}\right)^{(2/\alpha)+\varepsilon}$$

whence (2.13) follows since $a_N \geqq c_N$ and thus $\beta(a_N) \leqq \beta(c_N)$ for $N \geqq N_0$.

It is worth writing out the upper and lower estimates given by (1.6), (1.7) and (1.9) in the special case $\beta(x) = (\log x)^{-\gamma}$, $\gamma > 0$. Choosing $a_n = (n \log^{1+\varepsilon} n)^{1/\alpha}$ with $\varepsilon > 0$, (1.5) holds and thus (1.6), (1.7) give for any $\varepsilon > 0$

$$\left|\sum_{i=1}^n X_i - \sum_{i=1}^n Y_i\right| = O(n^{1/\alpha} (\log n)^{1/\alpha - \gamma + \varepsilon}) \quad \text{a.s.} \tag{2.16a}$$

$$\left|\sum_{i=1}^n X_i - \sum_{i=1}^n Y_i\right| = O(n^{1/\alpha} (\log n)^{1/\alpha - \gamma/\alpha + \varepsilon}) \quad \text{a.s.} \tag{2.16b}$$

for suitable versions of $\{X_i\}$, $\{Y_i\}$. On the other hand, $c_n = (n \log^\rho n)^{1/\alpha}$ satisfies (1.8) if $\rho < 1 - 2\gamma$ and thus Theorem 1 shows that with probability one

$$\left|\sum_{i=1}^n X_i - \sum_{i=1}^n Y_i\right| \geqq K n^{1/\alpha} (\log n)^{1/\alpha - 2\gamma/\alpha - \gamma - \varepsilon} \text{ infinitely often} \tag{2.16c}$$

for any $\varepsilon > 0$ and any versions of $\{X_i\}$, $\{Y_i\}$ defined on the same probability space. Note that the gap between the upper estimates (2.16a), (2.16b) and the lower estimate (2.16c) is a power of $\beta(n)$ and the gap is decreasing as $\gamma \to 0$. In [10] Mijnheer showed that the remainder term in Stout's approximation can be improved in certain cases if we approximate $\sum_{i=1}^n X_i$ not by $\sum_{i=1}^n Y_i$ but by $\sum_{i=1}^n \lambda_i Y_i$ where $\lambda_n \to 1$ is a suitable numerical sequence. His result in the case $\beta(x) = (\log x)^{-\gamma}$ is

$$\left|\sum_{i=1}^n X_i - \sum_{i=1}^n \lambda_i Y_i\right| = O(n^{1/\alpha} (\log n)^{1/\alpha - \gamma - 1 + \varepsilon}) \quad \text{a.s.} \tag{2.17a}$$

for suitable versions of $\{X_i\}$, $\{Y_i\}$ provided $\gamma>1$, $\alpha(\gamma+1)>1$. On the other hand, the remark at the end of this section shows that for any joint construction of $\{X_i\}$, $\{Y_i\}$ and any numerical sequence $\lambda_n\to 1$ we have with probability one

$$\left|\sum_{i=1}^{n} X_i-\sum_{i=1}^{n} \lambda_i Y_i\right|\geqq K n^{1/\alpha}(\log n)^{-1/\alpha-2\gamma/\alpha-\gamma-1-\varepsilon} \quad \text{i.o.} \tag{2.17b}$$

The gap between the upper estimate (2.17a) and the lower estimate (2.17b) is again a power of $\log n$. When both apply, (2.17a) gives a better remainder term than Stout's (2.16a) and for $\alpha(\gamma+1)>\gamma$ (2.17a) improves (2.16b) as well. However, for $1<\alpha(\gamma+1)<\gamma$ the remainder term of (2.17a) is worse than that of the unweighted approximation (2.16b).

If $\gamma>1/\alpha$ then for sufficiently small $\varepsilon>0$ the right side of (2.16a) is $O(n^{1/\alpha}(\log n)^{-\lambda})$ where $\lambda>0$. This remainder term is strong enough to carry over most known limit theorems from $\{Y_i\}$ to $\{X_i\}$. On the other hand, if $\gamma<1/(\alpha+2)$ then choosing $\varepsilon>0$ small enough,

$$\left|\sum_{i=1}^{n} X_i-\sum_{i=1}^{n} Y_i\right|\geqq K n^{1/\alpha}(\log n)^{\lambda} \quad \text{infinitely often}$$

where $\lambda>0$. (In fact $\lambda\to 1/\alpha$ if $\gamma\to 0$.) Hence for such $\beta(x)$ even the weak limit theorem (1.2) cannot be obtained via an a.s. invariance principle of the type (1.3). This difficulty can be overcome, as it was shown in [2], by adapting an idea of Major [9] and approximating $\sum_{i=1}^{n} X_i$ not by $\sum_{i=1}^{n} Y_i$ but by $\sum_{i=1}^{n} (Y_i+Z_i)$ where Z_i are small "correcting" r.v.'s, independent of each other and of the Y_i's. In what follows, we give a new, simpler proof of the just mentioned "perturbed" a.s. invariance principle of [2]. We shall namely prove the following

Theorem 2. *Let* $\beta(x)\to 0$ *satisfy*

$$\beta(x)\, x^{-\alpha} \text{ decreases for } x\geqq x_0 \tag{2.18}$$

$$|\beta(x+h)-\beta(x)|\leqq K x^{-1} \quad \text{for } x\geqq x_0, |h|\leqq 1 \tag{2.19}$$

Let F, G and H be symmetric distribution functions satisfying (1.4), (2.1) *and*

$$1-H(x)=\beta(x)\, x^{-\alpha} \qquad x\geqq x_0 \tag{2.20}$$

respectively. Assume also that G is continuous. Then there exist i.i.d. *sequences* $\{X_n\}$, $\{Y_n\}$, $\{Z_n\}$ *defined on a suitable probability space with distribution functions F, G and H, respectively, such that the sequences* $\{Y_n\}$ *and* $\{Z_n\}$ *are independent of each other and*

$$\sum_{i=1}^{n} (X_i-Y_i-Z_i)=O(n^{1/\alpha-\lambda}) \qquad \text{a.s.} \tag{2.21}$$

for every $0<\lambda<1/(\alpha+8)$.

In [2] this theorem is proved for $G=G_{\alpha,c}$ but without the assumption (2.19). (2.19) is satisfied e.g. if $\beta(x)$ and $\beta(x)-\beta(x+1)$ are decreasing; in particular if $\beta(x)$ is decreasing and convex. We note that the proof of Theorem 2 will yield r.v.'s X_n, Y_n, Z_n such that the vectors (X_n, Y_n, Z_n) are independent. The assumption that $\beta(x)x^{-\alpha}$ is decreasing (and consequently $\beta(x)\geqq 0$) is clearly necessary in Theorem 2 since otherwise there is no distribution function H satisfying (2.20). However, Theorem 2 can be easily modified to cover β's which take both positive and negative values. Let e.g. $\beta^*(x)=\sup_{t\geqq x}|\beta(t)|$, $\beta_1(x)=\beta(x)+\sup_{t\geqq x}|\beta(t)|$ $(x>0)$; clearly $\beta_1(x)\geqq 0$ and $\beta_1(x)\to 0$ as $x\to+\infty$. Assume that $\beta_1(x)x^{-\alpha}$ is decreasing and $\beta(x)$, $\beta_1(x)$ satisfy (2.19). Let $\{X_n\}$, $\{T_n\}$ be i.i.d. sequences, independent of each other, defined on a suitable probability space such that X_n has distribution F and T_n is symmetric with $P(T_n>x)=\beta^*(x)x^{-\alpha}$ $x\geqq x_0$. (Note that $\beta^*(x)$ is decreasing.) The proof of Lemma (2.1) shows that

$$P(X_n+T_n>x)=(c+\beta(x)+\beta^*(x))\,x^{-\alpha}+O(x^{-\alpha-\lambda}) \qquad x\to+\infty$$

for some $\lambda>0$. Thus Theorem 2 applies for the sequence X_n+T_n with $\beta(x)$ replaced by $\beta_1(x)$ (the extra term $O(x^{-\alpha-\lambda})$ in the last formula does not cause any difficulty) and we get that on a suitable probability space there exist sequences $\{X_n\}$, $\{Y_n\}$, $\{Z_n\}$ satisfying (2.21) such that $\{X_n\}$ is i.i.d. with distribution function F, $\{Y_n\}$ is i.i.d. with distribution function G and Z_n are independent with $P(|Z_n|\geqq x)\leqq 24(\sup_{t\geqq x/2}|\beta(t)|)\,x^{-\alpha}$, $x\geqq x_0$. Specifically, $Z_n=Z_n^{(1)}+Z_n^{(2)}$ where $Z_n^{(1)}$ are i.i.d. symmetric with $P(Z_n^{(1)}>x)=\beta_1(x)x^{-\alpha}$, $Z_n^{(2)}$ are i.i.d. symmetric with $P(Z_n^{(2)}>x)=\beta^*(x)x^{-\alpha}$. Since the main use of Theorem 2 is to get information on $\sum_{i=1}^{n}(X_i-Y_i)$, the more complicated form of Z_n in the present situation does not cause any difficulty in applications.

For the proof of Theorem 2 we need two lemmas.

Lemma 2.1. *Let X, Y and Z be random variables with symmetric distribution functions F, G and H satisfying* (1.4), (2.1) *and* (2.20), *respectively. Assume also that Y and Z are independent. Then we have*

$$|P(X>t)-P(Y+Z>t)|\ll t^{-5\alpha/4} \quad \textit{as } t\to\infty. \tag{2.22}$$

Proof. We use the inequalities

$$\begin{aligned}P\{Y+Z>t\}\geqq P\{Y>t(1+\varepsilon)\}\,P\{|Z|<t\varepsilon\}\\ +P\{Z>t(1+\varepsilon)\}\,P\{|Y|<t\varepsilon\}\end{aligned} \tag{2.23}$$

and

$$\begin{aligned}P\{Y+Z>t\}\leqq P\{Y>t(1-\varepsilon)\}+P\{Z>t(1-\varepsilon)\}\\ +P\{Y>t\varepsilon\}\,P\{Z>t\varepsilon\}\end{aligned} \tag{2.24}$$

valid for $t>1$, $\varepsilon>0$. (See [4], p. 271.) Now let $\varepsilon_t=t^{-\alpha/4}$, then

$$(t(1\pm\varepsilon_t))^{-\alpha}=t^{-\alpha}(1+O(\varepsilon_t))=t^{-\alpha}+O(t^{-5\alpha/4}) \quad \text{as } t\to\infty$$

By repeated application of (2.19) we get

$$\beta(t\pm t\varepsilon_t)-\beta(t)=O(t^{-1}\cdot t\varepsilon_t)=O(t^{-\alpha/4}).$$

Hence using (2.1), (2.20), (2.23) and $t\varepsilon_t=t^{1-\alpha/4}\geqq t^{1/2}$ we obtain

$$\begin{aligned}P\{Y+Z>t\}\geqq&(c(t(1+\varepsilon_t))^{-\alpha}-O((t(1+\varepsilon_t))^{-2\alpha}))\left(1-\frac{2\beta(t\varepsilon_t)}{(t\varepsilon_t)^\alpha}\right)\\&+\frac{\beta(t(1+\varepsilon_t))}{(t(1+\varepsilon_t))^\alpha}(1-O((t\varepsilon_t)^{-\alpha}))\\\geqq&(ct^{-\alpha}-O(t^{-5\alpha/4}))(1-O(t^{-\alpha/2}))\\&+(\beta(t)-O(t^{-\alpha/4}))(t^{-\alpha}-O(t^{-5\alpha/4}))(1-O(t^{-\alpha/2}))\\=&ct^{-\alpha}+\beta(t)\,t^{-\alpha}-O(t^{-5\alpha/4}).\end{aligned}$$

A similar calculation yields, using (2.24),

$$P\{Y+Z>t\}\leqq ct^{-\alpha}+\beta(t)\,t^{-\alpha}+O(t^{-5\alpha/4})$$

and Lemma 2.1 follows since

$$P\{X>t\}=ct^{-\alpha}+\beta(t)\,t^{-\alpha}.$$

Lemma 2.2. *Let X, Y and Z be random variables with symmetric distribution functions F, G and H satisfying* (1.4), (2.1) *and* (2.20), *respectively. Assume also that Y and Z are independent and denote by L the distribution function of $Y+Z$. Then we have*

$$|u-F^{-1}(L(u))|\ll u^{1-\alpha/4} \quad \textit{for } u\geqq u_0 \tag{2.25}$$

where $F^{-1}(x)=\inf\{t\colon F(t)\geqq x\}$.

Proof. By Lemma 2.1 and (1.4) we have

$$1-L(x)=cx^{-\alpha}+\beta(x)\,x^{-\alpha}+O(x^{-\alpha-\gamma}) \quad x>0 \tag{2.26}$$

where $\gamma=\alpha/4>0$. We next prove that there exists a constant $K=K(\alpha,c)$ such that

$$F(u)-F(u-Ku^{1-\gamma})>K_1\,u^{-\alpha-\gamma} \quad \text{for } u\geqq u_0 \tag{2.27}$$

where K_1 is the constant implied by O in (2.26). Indeed, note that for $1/2<x<3/2$ we have $x^\alpha-1\geqq K_\alpha(x-1)$ and thus, using (1.4) and the monotonicity of $\beta(x)\,x^{-\alpha}$,

$$\begin{aligned}F(u)-F(u-Ku^{1-\gamma})&=(1-F(u-Ku^{1-\gamma}))-(1-F(u))\\&\geqq\frac{c}{(u-Ku^{1-\gamma})^\alpha}-\frac{c}{u^\alpha}=\frac{c}{u^\alpha}\left(\frac{1}{(1-Ku^{-\gamma})^\alpha}-1\right)\\&\geqq\frac{c}{u^\alpha}K_\alpha\left(\frac{1}{1-Ku^{-\gamma}}-1\right)=\frac{c}{u^\alpha}K_\alpha\frac{Ku^{-\gamma}}{1-Ku^{-\gamma}}\geqq cKK_\alpha u^{-\alpha-\gamma}.\end{aligned}$$

for $u\geqq u_0$. Hence any $K>K_1/cK_\alpha$ satisfies (2.27).

Now we can easily prove (2.25). Assume $F(u)\geqq L(u)$; the other case goes similarly. (1.4) and (2.26) imply $|F(u)-L(u)|\leqq K_1u^{-\alpha-\gamma}$ and thus by (2.27) we get $F(u-Ku^{1-\gamma})<L(u)$. By the definition of F^{-1} this means

$$F^{-1}(L(u))\geqq u-Ku^{1-\gamma}. \tag{2.28}$$

By $F(u)\geqq L(u)$ we have $F^{-1}(L(u))\leqq u$ which, together with (2.28), implies

$$0\leqq u-F^{-1}(L(u))\leqq Ku^{1-\gamma}$$

and Lemma (2.2) is proved.

We can now easily complete the proof of Theorem 2. Let $(\Omega,\mathscr{F},P)$ be a probability space large enough to carry two i.i.d. sequences $\{Y_n\}$ and $\{Z_n\}$ with distribution functions G and H, respectively, such that $\{Y_n\}$ and $\{Z_n\}$ are also independent of each other. Let $L(x)$ be the distribution function of Y_1+Z_1 and put

$$X_n=F^{-1}(L(Y_n+Z_n))\qquad n=1,2,\ldots$$

Since G is continuous, so is L and thus $\{X_n\}$ is an i.i.d. sequence with distribution function F. We show that (2.21) holds. Set

$$\eta_n=X_n-(Y_n+Z_n)$$

then

$$\eta_n=f(Y_n+Z_n)$$

where

$$f(u)=F^{-1}(L(u))-u.$$

Since F and L are symmetric, $f(u)$ is an odd function and thus η_n has a symmetric distribution. By Lemma 2.2 we have

$$|\eta_n|\ll|Y_n+Z_n|^{1-\alpha/4} \tag{2.29}$$

for $Y_n+Z_n\geqq u_0$ and, for reasons of symmetry, for $Y_n+Z_n\leqq -u_0$. Since $0<L(u)<1$ for all u (see (2.26)), $f(u)$ is bounded in $[-u_0,u_0]$ and thus (2.29) holds also for $|Y_n+Z_n|\leqq u_0$, provided we add 1 to the right side. Hence we get for $t\geqq t_0$

$$\begin{aligned} P\{|\eta_n|\geqq t\} &\leqq P\{|Y_n+Z_n|\geqq c_1\, t^{1/(1-\alpha/4)}\} \\ &=2(1-L(c_1\, t^{1/(1-\alpha/4)}))=O(t^{-\alpha/(1-\alpha/4)})=O(t^{-\alpha-\delta}) \end{aligned} \tag{2.30}$$

where $\delta=\alpha^2/(4-\alpha)\geqq\alpha^2/4$. Now $\{\eta_n\}$ is an i.i.d. symmetric sequence and by (2.30) we have $E|\eta_n|^{\alpha+\delta/2}<+\infty$. Thus by Marczinkiewicz's strong law (see e.g. [8] p. 243)

$$\eta_1+\ldots+\eta_n=O(n^{1/(\alpha+\delta/2)})=O(n^{1/\alpha-\lambda}) \quad \text{a.s.} \tag{2.31}$$

for any $0<\lambda<1/(\alpha+8)$; hence (2.21) is proved.

Remark. Under the conditions of Theorem 2 one can also show that for some $p>\alpha$

$$\sup_n \left\| \frac{\sum_{i=1}^{n} (X_i-Y_i-Z_i)}{n^{1/\alpha-\lambda}} \right\|_p <+\infty. \tag{2.32}$$

This follows immediately from (2.31) and the following result which is implicit in [1] (see the proof of Theorem (6.1) on p. 226 with $a_n=n^\gamma$):

Lemma 2.3. *Let* $\xi_1, \xi_2, \ldots$ *be* i.i.d. *symmetric r.v.'s,* $S_n=\sum_{i\leqq n}\xi_i$ *and assume that* S_n/n^γ *is stochastically bounded for some* $\gamma>0$. *Then* $\sup_n\|S_n/n^\gamma\|_{1/\gamma-\varepsilon}<+\infty$ *for every* $\varepsilon>0$.

Simple applications of Theorem 2 are given in Sect. 4 of [2]. In particular, Lemma (4.1) of [2] shows that (1.5) implies

$$\sum_{i\leqq n} Z_i \ll a_n\,\beta(a_n)^{1/\alpha} \quad \text{a.s.}$$

and thus Theorem 2 implies (1.7). Further applications of the theorem (for example for Chung's law of the iterated logarithm) will be given elsewhere.

Remark. Mijnheer showed (see [10], Examples 3 and 4) that in certain cases the remainder term in (1.6) can be improved if we replace $\sum_{i=1}^{n} Y_i$ by $\sum_{i=1}^{n} \lambda_i Y_i$ where λ_i is a suitable numerical sequence tending to 1. E.g. for $\beta(x)=(\log x)^{-\gamma}$, $\gamma>0$ he obtained for any $\varepsilon>0$

$$\left|\sum_{i=1}^{n} X_i-\sum_{i=1}^{n}\lambda_i Y_i\right|=O(n^{1/\alpha}(\log n)^{1/\alpha-\gamma-1+\varepsilon}) \quad \text{a.s.}$$

provided $\gamma>1$, $\alpha(\gamma+1)>1$. We show now that an easy modification of the proof of Theorem 1 yields lower estimates for $\left|\sum_{i=1}^{n} X_i-\sum_{i=1}^{n}\lambda_i Y_i\right|$ for any joint construction of $\{X_n\}$, $\{Y_n\}$ and any numerical sequence $\lambda_n \to 1$. For the sake of simplicity, let us consider the case $\beta(x)=(\log x)^{-\gamma}$ $\gamma>0$. Let F and G be distribution functions satisfying (1.4) and (2.1), respectively, where $\beta(x)=(\log x)^{-\gamma}$, $\gamma>0$. Let $c_n \to +\infty$ and $\lambda_n \to 1$ be numerical sequences and put, for $t>0$,

$$\psi_k(t) =: P(X_k>t)-P(\lambda_k Y_k>t)$$
$$=(c(1-\lambda_k^\alpha)+\beta(t))\, t^{-\alpha}+O(t^{-2\alpha}).$$

Since $|\beta(2c_k)-\beta(c_k)| \geqq \text{const}\cdot(\log c_k)^{-\gamma-1}$ by the mean value theorem, the inequality $|c(1-\lambda_k^\alpha)+\beta(t)| \geqq \text{const}\cdot(\log c_k)^{-\gamma-1}$ holds at least for one of the values $t=c_k$ and $t=2c_k$ which implies that one of $|\psi_k(c_k)|$ and $|\psi_k(2c_k)|$ exceeds $\text{const}\cdot(\log c_k)^{-\gamma-1} c_k^{-\alpha}$. In other words, there is a sequence c_k^* such that $c_k^*=c_k$ or $c_k^*=2c_k$ for each $k\geqq 1$ and

$$|P(X_k>c_k^*)-P(\lambda_k Y_k>c_k^*)| \geqq \text{const}\cdot(\log c_k^*)^{-\gamma-1}(c_k^*)^{-\alpha}.$$

Set $d_k^*=\rho c_k^*(\log c_k^*)^{-\gamma-1}$, then using (2.1), $\lambda_n\to 1$ and the mean value theorem we get

$$|P(\lambda_k Y_k>c_k^*)-P(\lambda_k Y_k>c_k^*\pm d_k^*)|$$
$$\leqq \text{const}\cdot(\log c_k^*)^{-\gamma-1}(c_k^*)^{-\alpha}+O((c_k^*)^{-2\alpha})$$

where the constant can be made as small a desired by choosing ρ small. Thus for ρ small enough we get

$$|P(X_k>c_k^*)-P(\lambda_k Y_k>c_k^*\pm d_k^*)| \geqq \text{const}\cdot(\log c_k^*)^{-\gamma-1}(c_k^*)^{-\alpha}$$
$$=\text{const}\cdot\beta(c_k^*)^{1+1/\gamma}(c_k^*)^{-\alpha}. \tag{2.33}$$

We also see that the sign of $P(X_k>c_k^*)-P(\lambda_k Y_k>c_k^*)$ is the same as that of $P(X_k>c_k^*)-P(\lambda_k Y_k>c_k^*\pm d_k^*)$ for both choices of the $\pm$. The estimate (2.33) is the same as we get in the unweighted case $\lambda_k=1$ except that $\beta(x)$ is replaced by $\beta(x)^{1+1/\gamma}$. Hence the proof of Theorem 1 can be repeated (the sets H_1 and H_2 should be defined in the present case as the sets of those k such that $P(X_k>c_k^*)-P(\lambda_k Y_k>c_k^*)$ is positive resp. negative) and we get that if

$$c_k^*\to\infty, \qquad \min_{1\leqq k\leqq N}\beta(c_k^*)^{2+2/\gamma}\sum_{k\leqq N}(c_k^*)^{-\alpha}\to+\infty \tag{2.34}$$

and

$$\liminf_{k\to\infty}|c_{k+1}^*\,\beta(c_{k+1}^*)^{1+1/\gamma}/c_k^*\,\beta(c_k^*)^{1+1/\gamma}|>0 \tag{2.35}$$

then

$$P\left\{\left|\sum_{i=1}^n X_i-\sum_{i=1}^n \lambda_i Y_i\right|\geqq K c_n^*\,\beta(c_n^*)^{1+1/\gamma}\text{ i.o.}\right\}=1 \tag{2.36}$$

for any joint construction of $\{X_n\}$, $\{Y_n\}$ and any numerical sequence $\lambda_n\to 1$. Since $c_k\leqq c_k^*\leqq 2c_k$ and $\beta(x)=(\log x)^{-\gamma}$, in (2.34), (2.35) and (2.36) we can replace

c_k^* by c_k. Choosing $c_n=(n\log^\rho n)^{1/\alpha}$ with $\rho=-1-2\gamma-\varepsilon$ $(\varepsilon>0)$ we get the lower bound

$$\left|\sum_{i=1}^n X_i-\sum_{i=1}^n \lambda_i Y_i\right|\geqq K n^{1/\alpha}(\log n)^{-1/\alpha-2\gamma/\alpha-\gamma-1-\varepsilon} \quad \text{i.o.}$$

valid for any versions of $\{X_i\}$, $\{Y_i\}$ and any $\lambda_i\to 1$.

The above argument extends, without any difficulty, for a large class of β's. We leave the details to the reader.

§ 3. Probabilistic Invariance Principles – Upper and Lower Bounds

In this section we prove analogues of the results of Sect. 2 in the case when "a.s." in (1.3) is replaced by "in probability". Simons and Stout proved in [12] that for any $\beta(x)\to 0$ the sequences $\{X_n\}$ and $\{Y_n\}$ can be defined on a common probability space such that

$$n^{-1/\alpha}\sum_{i=1}^n (X_i-Y_i)\xrightarrow{P}0. \tag{3.1}$$

In other words, there exists an $\varepsilon_n\to 0$ such that

$$P\left\{n^{-1/\alpha}\left|\sum_{i=1}^n (X_i-Y_i)\right|\geqq\varepsilon_n\right\}\leqq\varepsilon_n \qquad n=1,2,\ldots \tag{3.2}$$

In this section we give upper and lower estimates for the optimal ε_n in (3.2).

Theorem 3. *Let $\beta(x)\downarrow 0$ satisfy the following conditions:*

$$\beta(x)\gg x^{-\lambda} \quad \text{as } x\to\infty \quad \text{for some } \lambda>0 \tag{3.3}$$

$$|\beta(x+1)-\beta(x)|\ll x^{-1} \quad \text{as } x\to\infty. \tag{3.4}$$

Let F and G be symmetric distribution functions satisfying (1.4) *and* (2.1), *respectively, and assume that G is continuous. Then there exist* i.i.d. *sequences $\{X_n\}$ and $\{Y_n\}$ defined on a suitable probability space with distribution functions F and G, respectively, such that*

$$P\left\{n^{-1/\alpha}\max_{k\leqq n}\left|\sum_{i=1}^k (X_i-Y_i)\right|\gg\beta(n^\gamma)^{1/(1+\alpha)}\right\}\ll\beta(n^\gamma)^{1/(1+\alpha)} \qquad n=1,2,\ldots \tag{3.5}$$

for some constant $\gamma>0$. If β satisfies (3.3) *for every $\lambda>0$ then* (3.5) *holds for every $\gamma<1/\alpha$.*

Remarks. 1. If β satisfies, in addition to the above conditions, also

$$\beta(x^\rho)\ll\beta(x)^\tau \quad \text{for some } 0<\rho<1, 0<\tau\leqq 1 \tag{3.6}$$

then (3.5) can be replaced by

$$P\left\{n^{-1/\alpha}\max_{k\leqq n}\left|\sum_{i=1}^{k}(X_i-Y_i)\right|\gg\beta(n^{1/\alpha})^{\tau/(1+\alpha)}\right\}\ll\beta(n^{1/\alpha})^{\tau/(1+\alpha)}\qquad n=1,2,\dots \quad (3.7)$$

Typical cases are $\beta(x)=x^{-\lambda}$, $\beta(x)=(\log x)^{-\lambda}$, $\beta(x)=(\log\log x)^{-\lambda}$, ..., $(\lambda>0)$; $\beta(x)=\exp\{-C(\log x)^{\lambda}\}$ $(C>0, 0<\lambda<1)$, etc.

2. If $\{X_n\}$ and $\{Y_n\}$ are i.i.d. sequences with distribution functions F and G (not necessarily satisfying (1.4) and (2.1)) and $F\neq G$ then (3.2) is impossible for any $\varepsilon_n=o(n^{-1/\alpha})$. Indeed, if (3.2) holds for some $\varepsilon_n=o(n^{-1/\alpha})$ then $\sum_{i=1}^{n}(X_i-Y_i)\xrightarrow{P}0$ and consequently $X_n-Y_n\xrightarrow{P}0$ which is clearly impossible.

This observation shows that if $\beta(x)$ tends to zero more rapidly than any negative power of x then (3.5) cannot hold. In other words, without condition (3.3), Theorem 3 is false.

3. Theorem 3, as formulated above, applies only for positive β's. However, exactly as Theorem 2, it can be extended for more general β's by a simple smoothing argument (cf. the remark following Theorem 2).

For the proof of Theorem 3 we shall need the following

Corollary of Lemma 2.3. *Let* $\xi_1, \xi_2, \dots$ *be* i.i.d. *symmetric* r.v.'s *with* $P\{|\xi_1|\geqq t\}\ll t^{-\alpha-\delta}$ *for some* $0<\alpha<2$ *and* $0<\delta<2-\alpha$. *Then*

$$P\left\{n^{-1/\alpha}\left|\sum_{i=1}^{n}\xi_i\right|\geqq\varepsilon_n\right\}\ll\varepsilon_n \quad \textit{where } \varepsilon_n=n^{-\delta/12}. \quad (3.8)$$

Proof. The assumptions of the corollary imply $E|\xi_1|^{\alpha+\delta/2}<+\infty$ and thus by Marczinkiewicz's strong law

$$n^{-1/(\alpha+\delta/2)}\sum_{i=1}^{n}\xi_i\to 0 \qquad \text{a.s.}$$

Hence by Lemma 2.3

$$n^{-1/(\alpha+\delta/2)}\left\|\sum_{i=1}^{n}\xi_i\right\|_{\alpha}=O(1)$$

and thus the Markov inequality yields (3.8) with $\varepsilon_n=n^{-\lambda}$,

$$\lambda=\frac{\delta}{(\alpha+1)(2\alpha+\delta)}.$$

Obviously $\lambda\geqq\delta/12$, completing the proof.

We shall also use the following large deviation result of Heyde, implicit in [6]:

Lemma 3.1. *Let $\xi, \xi_1, \xi_2, \ldots$ be* i.i.d. *symmetric* r.v.'s *not belonging to the domain of partial attraction of the normal law. Let $S_n = \sum_{i=1}^{n} \xi_i$, then*

$$\overline{\lim_{n\to\infty}}\, P\{|S_n| \geqq t_n\}/nP\{|\xi| \geqq t_n\} < +\infty$$

for any numerical sequence $t_n \to +\infty$.

Proof of Theorem 3. Since $\beta(x) \downarrow 0$ and (3.4) imply the assumptions of Theorem 2, we can apply the proof of Theorem 2. Let $\{X_n\}$, $\{Y_n\}$, $\{Z_n\}$, $\{\eta_n\}$ be the i.i.d. symmetric sequences introduced there and put

$$U_k = \sum_{i=1}^{k} \eta_i, \qquad V_k = \sum_{i=1}^{k} Z_i.$$

Then

$$\max_{k \leqq n} \left| \sum_{i=1}^{k} (X_i - Y_i) \right| \leqq \max_{k \leqq n} |U_k| + \max_{k \leqq n} |V_k|$$

and thus by Lévy's maximal inequality (see e.g. [8] p. 247) we get

$$P\left\{\max_{k \leqq n} \left| \sum_{i=1}^{k} (X_i - Y_i) \right| \geqq t\right\} \leqq P\{\max_{k \leqq n} |U_k| \geqq t/2\} + P\{\max_{k \leqq n} |V_k| \geqq t/2\}$$

$$\leqq 2P\{|U_n| \geqq t/2\} + 2P\{|V_n| \geqq t/2\} \tag{3.9}$$

for every $t > 0$. Let

$$c_n = \beta(n^\gamma)^{1/(1+\alpha)}. \tag{3.10}$$

Using (3.3) and the monotonicity of β it follows easily that

$$n^{1/\alpha} c_n \to \infty \quad \text{and} \quad \beta(n^{1/\alpha} c_n) \ll c_n^{1+\alpha} \tag{3.11}$$

provided γ is small enough. Now $P\{|\eta_n| \geqq t\} \ll t^{-\alpha-\delta}$ for some $\delta > 0$ (see (2.30)) and choosing γ small enough we can guarantee also that $c_n \gg n^{-\delta/12}$. (If β satisfies (3.3) for every $\lambda > 0$ then c_n will satisfy the above requirements for every $\gamma < 1/\alpha$.) Hence applying the Corollary above we get

$$P\{|U_n| \gg n^{1/\alpha} c_n\} \ll c_n. \tag{3.12}$$

If Z_1 has a finite moment of order $> \alpha$ then $P\{|Z_1| \geqq t\} \ll t^{-\alpha-\delta_1}$ for some $\delta_1 > 0$ and thus the argument leading to (3.12) yields

$$P\{|V_n| \gg n^{1/\alpha} c_n\} \ll c_n. \tag{3.13}$$

If $E|Z_1|^p = +\infty$ for every $p > \alpha$ then Z_1 has an infinite moment of order < 2 and thus it does not belong to the domain of partial attraction of the normal law (see [7] p. 117). Hence Lemma 3.1 applies to the sequence $\{Z_n\}$ and we get, using (3.11) and $P\{|Z_1|>t\}=2\beta(t)/t^{\alpha}$,

$$P\{|V_n|\geqq n^{1/\alpha}c_n\}\ll\frac{n\beta(n^{1/\alpha}c_n)}{n c_n^{\alpha}}\ll c_n$$

i.e. (3.13) holds in this case too. Now (3.9), (3.10), (3.12) and (3.13) imply (3.5), completing the proof of Theorem 3.

We turn now to lower estimates for the optimal ε_n in (3.2).

Theorem 4. *Let F and G be symmetric distribution functions satisfying* (1.4) *and* (2.1), *respectively. Assume that $\beta(x)\to 0$ and*

$$\beta(x) \text{ is decreasing for } x\geqq x_0 \tag{3.14}$$

$$\beta(x) \text{ is regulary varying at infinity with exponent } -\gamma \text{ where } 0\leqq\gamma<\min(\alpha, 2-\alpha). \tag{3.15}$$

Then there is a constant $K>0$ such that if $\{X_n\}$ and $\{Y_n\}$ are i.i.d. *sequences defined on a common probability space with distribution functions F and G, respectively, the relation*

$$P\left\{n^{-1/\alpha}\left|\sum_{i=1}^{n}(X_i-Y_i)\right|\geqq K\beta(n^{1/\alpha})\right\}\leqq K\beta(n^{1/\alpha}) \tag{3.16}$$

is impossible for every n.

Comparing Theorem 3 and Theorem 4 we see that the optimal ε_n in (3.2) satisfies

$$\beta(n^{1/\alpha})\ll\varepsilon_n\ll\beta(n^{\gamma})^{1/(1+\alpha)} \tag{3.17}$$

for some constant $\gamma>0$ provided β satisfies (3.4), (3.14) and (3.15). If β is slowly varying then (3.17) holds for every $\gamma<1/\alpha$. If β satisfies also (3.6) then we have

$$\beta(n^{1/\alpha})\ll\varepsilon_n\ll\beta(n^{1/\alpha})^{\tau/(1+\alpha)} \tag{3.18}$$

(see Remark 1 after the Theorem 3). As a matter of fact, the regular variation of β in Theorem 4 is assumed only to guarantee a lower estimate for ε_n in (3.17), (3.18) which is directly comparable with the upper estimates. It is easy to get various (less precise) bounds on ε_n under weaker assumptions on β. For example, Remark 2 after Theorem 3 shows that (3.2) cannot hold with $\varepsilon_n=o(n^{-1/\alpha})$ if the distribution functions F and G of $\{X_n\}$ and $\{Y_n\}$ are different. Hence if β satisfies $\beta(x)\ll x^{-\lambda}$ for some $\lambda>0$ then the optimal ε_n in (3.2) satisfies $\varepsilon_n\geqq c\beta(n^{\rho})$ for infinitely many n where c and ρ are positive constants. The proof of Theorem 4 also yields lower bounds for ε_n in the absence of (3.15); these involve, however, the function $\beta_1(t)$ defined by (3.20) and thus are less explicit in terms of β.

For the proof of Theorem 4 we need some lemmas.

Lemma 3.2. *(a) Let F be a symmetric distribution function satisfying* (1.4) *such that $\beta(x)\to 0$ and* (3.14) *holds. Extend $\beta(x)$ for all $x\geqq 0$ by defining $\beta(x)=0$ for $0\leqq x<x_0$. Then the characteristic function φ of F satisfies*

$$\varphi(t)=1-c_1 t^\alpha-\beta_1(t)\, t^\alpha+O(t^2) \qquad \text{as } t\downarrow 0 \tag{3.19}$$

where

$$c_1=2c\int_0^\infty \sin y/y^\alpha\, dy, \qquad \beta_1(t)=2\int_0^\infty \beta(y/t)\sin y/y^\alpha\, dy \tag{3.20}$$

(Here both integrals are convergent as Riemann-integrals.) Moreover, $\beta_1(t)\to 0$ as $t\downarrow 0$.

b) Let G be a symmetric distribution function satisfying (2.1). *Then the characteristic function ψ of G satisfies*

$$\psi(t)=1-c_1 t^\alpha+O(t^{2\alpha}) \qquad \text{as } t\downarrow 0$$

where c_1 is defined by (3.20).

For the proof of (a) see [2], Lemma 3.1. Part (b) is implicit in [3].

Lemma 3.3. *Let $\beta(x)\to 0$ be a function satisfying* (3.14), (3.15) *and $\beta(x)=0$ for $0\leqq x<x_0$. Define $\beta_1(t)$ by* (3.20). *Then we have*

$$\beta_1(t)\sim A\beta\left(\frac{1}{t}\right) \qquad \text{as } t\downarrow 0$$

where

$$A=2\int_0^\infty \frac{\sin y}{y^{\alpha+\gamma}}\, dy.$$

Proof. The integral defining A is convergent in the Riemann sense since $\alpha+\gamma<2$; clearly $A>0$. By (3.15), $\beta(x)=x^{-\gamma}L(x)$ where L is slowly varying at infinity and $0\leqq\gamma<\min(\alpha, 2-\alpha)$. Also $L(x)=0$ for $0\leqq x<x_0$. Hence

$$\frac{\beta_1(t)}{\beta(1/t)}=2\int_0^\infty \frac{L(y/t)}{L(1/t)}\frac{\sin y}{y^{\alpha+\gamma}}\, dy. \tag{3.21}$$

The integrand in (3.21) converges to $\sin y/y^{\alpha+\gamma}$ for any $y>0$ if $t\downarrow 0$. We now find an integrable bound for this integrand. Choose $\delta>0$ so small that $\alpha+\gamma+\delta<2$. It is easy to see that for $0<t\leqq 1/x_0$

$$\frac{L(y/t)}{L(1/t)}\leqq\begin{cases} y^\gamma & \text{if } y\geqq 1\\ C y^{-\delta} & \text{if } y<1\end{cases} \tag{3.22}$$

where C is a constant depending on $\beta(x)$ and δ. The upper line of (3.22) is equivalent to $\beta(y/t)\leqq\beta(1/t)$ and thus it is valid by the monotonicity of β. Using the representation theorem for slowly varying functions (see e.g. [11], pp. 2–3) it follows immediately that the lower inequality of (3.22) holds if we assume in addition that $y/t \geqq C$. If $y/t < C$ then we have

$$\beta\left(\frac{1}{t}\right) \geqq \beta\left(\frac{C}{y}\right) = \left(\frac{C}{y}\right)^{-\gamma} L\left(\frac{C}{y}\right) \gg \left(\frac{C}{y}\right)^{-\gamma-\delta} \qquad (t \leqq t_0) \tag{3.23}$$

by the monotonicity of β and a well known property of slowly varying functions (see e.g. [11], p. 18, 1°). Since β is bounded from above, (3.23) yields

$$\frac{\beta(y/t)}{\beta(1/t)} \ll \left(\frac{C}{y}\right)^{\gamma+\delta}$$

and thus the second inequality of (3.22) follows in this case too. We thus proved that the absolute value of the integrand in (3.21) is bounded by $C \sin y/y^{\alpha+\gamma+\delta}$ for $0 < y \leqq 1$ and by $1/y^{\alpha}$ for $y > 1$. Since $\int_0^1 \sin y/y^{\lambda} dy$ converges for $\lambda < 2$, the dominated convergence theorem yields

$$\int_0^K \frac{L(y/t)}{L(1/t)} \frac{\sin y}{y^{\alpha+\gamma}} dy \to \int_0^K \frac{\sin y}{y^{\alpha+\gamma}} dy \qquad \text{for every } 0 < K < \infty. \tag{3.23}$$

On the other hand, the monotonicity of $\beta(x)$ for $x \geqq x_0$ implies that for $0 < t \leqq 1/x_0$ the series

$$\sum_{n=1}^{\infty} \int_{n\pi}^{(n+1)\pi} \frac{\beta(y/t)}{\beta(1/t)} \frac{\sin y}{y^{\alpha}} dy$$

is alternating with decreasing summands. Hence

$$\left| \int_{n\pi}^{\infty} \frac{L(y/t)}{L(1/t)} \frac{\sin y}{y^{\alpha+\gamma}} dy \right| = \left| \int_{n\pi}^{\infty} \frac{\beta(y/t)}{\beta(1/t)} \frac{\sin y}{y^{\alpha}} dy \right|$$
$$\leqq \int_{n\pi}^{(n+1)\pi} \left| \frac{\beta(y/t)}{\beta(1/t)} \frac{\sin y}{y^{\alpha}} \right| dy \leqq \int_{n\pi}^{(n+1)\pi} y^{-\alpha} dy \leqq 4n^{-\alpha} \qquad (t \leqq 1/x_0) \tag{3.24}$$

where the integrals with upper limit ∞ are meant in the Riemann sense. Similarly,

$$\int_{n\pi}^{\infty} \frac{\sin y}{y^{\alpha+\gamma}} dy \leqq 4n^{-(\alpha+\gamma)} \tag{3.25}$$

Relations (3.23)–(3.25) show that the convergence relation in (3.23) holds also for $K = \infty$ which, in view of (3.21), implies the statement of the lemma.

Lemma 3.4. *Let* $\{X_n\}$ *and* $\{Y_n\}$ *be* i.i.d. *symmetric sequences with distribution functions F and G, respectively. Assume that* (1.4) *and* (2.1) *hold and the function*

$\beta(x)\to 0$ satisfies (3.14) and (3.15). Then the characteristic functions φ_n and ψ_n of the normalized sums $n^{-1/\alpha}\sum_{i=1}^{n} X_i$ and $n^{-1/\alpha}\sum_{i=1}^{n} Y_i$ satisfy

$$|\varphi_n(1)-\psi_n(1)|\geqq C\beta(n^{1/\alpha}) \tag{3.26}$$

where C is positive constant.

Proof. Let φ and ψ denote the characteristic functions of F and G, respectively. Using part (a) of Lemma 3.2 and the expansions $1+x=\exp(x+O(x^2))$, $e^x=1+x+O(x^2)$ $(x\to 0)$ we get

$$\begin{aligned}\varphi_n(1)&=\varphi^n(n^{-1/\alpha})=(1-c_1 n^{-1}-\beta_1(n^{-1/\alpha})n^{-1}+O(n^{-2/\alpha}))^n\\&=\exp\{-c_1-\beta_1(n^{-1/\alpha})+O(n^{1-2/\alpha})+O(n^{-1})\}\\&=\exp(-c_1)\{1-\beta_1(n^{-1/\alpha})+O(\beta_1(n^{-1/\alpha})^2)\}\{1+O(n^{1-2/\alpha})+O(n^{-1})\}\\&=\exp(-c_1)\{1-\beta_1(n^{-1/\alpha})\}+O(\beta_1(n^{-1/\alpha})^2+n^{1-2/\alpha}+n^{-1}).\end{aligned} \tag{3.27}$$

On the other hand, using part (b) of Lemma 3.2 we get

$$\begin{aligned}\psi_n(1)&=\psi^n(n^{-1/\alpha})=(1-c_1 n^{-1}+O(n^{-2}))^n\\&=\exp\{-c_1+O(n^{-1})\}=\exp(-c_1)+O(n^{-1}).\end{aligned} \tag{3.28}$$

Comparing (3.27) and (3.28) we obtain

$$|\varphi_n(1)-\psi_n(1)|\geqq K_1(|\beta_1(n^{-1/\alpha})|-O(n^{-\frac{2-\alpha}{\alpha}})-O(n^{-1})) \tag{3.29}$$

where K_1 is a positive constant. By Lemma 3.3 we have $\beta_1(n^{-1/\alpha})\sim A\beta(n^{1/\alpha})$ as $n\to\infty$. Also, $\beta(x)\gg x^{-\gamma-\varepsilon}$ for every $\varepsilon>0$ by (3.15) and a known property of slowly varying functions (see [11] p. 18, 1°). Thus $\gamma<\min(\alpha, 2-\alpha)$ yields

$$n^{1-2/\alpha}+n^{-1}=o(\beta(n^{1/\alpha}))$$

and consequently (3.29) implies (3.26).

Theorem 4 is an immediate consequence of Lemma 3.4. Assume that the i.i.d. sequences $\{X_n\}$ and $\{Y_n\}$ with distribution functions F and G can be defined on a common probability space such that (3.16) holds for some $n\geqq 1$. Setting

$$X=n^{-1/\alpha}\sum_{i=1}^{n} X_i,\qquad Y=n^{-1/\alpha}\sum_{i=1}^{n} Y_i,\qquad \varepsilon_n=K\beta(n^{1/\alpha})$$

it follows that $P\{|X-Y|\geqq\varepsilon_n\}\leqq\varepsilon_n$ which implies, via the inequality $|e^{ix}-e^{iy}|\leqq|x-y|\wedge 2$ (x, y are reals) that the characteristic function $\varphi_X(t)$ and $\varphi_Y(t)$ of X and Y satisfy $|\varphi_X(1)-\varphi_Y(1)|\leqq 3\varepsilon_n$. The last relation obviously contradicts to Lemma 3.4 if K is small enough.

References

1. deAcosta, A., Giné, E.: Convergence of moments and related functionals in the general central limit theorem in Banach spaces. Z. Wahrscheinlichkeitstheor. Verw. Geb. **48**, 213–231 (1979)
2. Berkes, I., Dabrowski, A., Dehling, H., Philipp, W.: A strong approximation theorem for sums of random vectors in the domain of attraction to a stable law. Acta. Math. Hungar. **48**, 161–172 (1986)
3. Cramér, H.: On the approximation to a stable probability distribution. In: Szegö, G. (ed.) Studies in mathematical analysis and related topics (Essays in honor of G. Pólya), pp. 70–76. Stanford, Calif.: Stanford University Press 1962
4. Feller, W.: An introduction to probability theory and its applications, vol. II. New York: Wiley 1966
5. Fisher, E.: An almost sure invariance principle for random variables in the domain of attraction of a stable law. Z. Wahrscheinlichkeitstheor. Verw. Geb. **67**, 461–471 (1984)
6. Heyde, C,C.: On large deviation problems for sums of random variables which are not attracted by the normal law. Ann. Math. Statist. **38**, 1575–1578 (1967)
7. Lévy, P.: Théorie de l'addition des variables aléatoires. Paris: Gauthiers-Villars 1937
8. Loève, M.: Probability theory. 2nd ed. New York: Van Nostrand 1963
9. Major, P.: An improvement of Strassen's invariance principle. Ann. Probab. **7**, 55–61 (1979)
10. Mijnheer, J.: Strong approximations of partial sums of i.i.d. random variables in the domain of attraction of a symmetric stable distribution. Trans. of the 9th Prague conference on information theory, stat. decision functions, random processes. Prague: Academia 1983
11. Seneta, E.: Regularly varying functions. (Lect. Notes. Math., vol. 508) Berlin Heidelberg New York: Springer 1976
12. Simons, G., Stout, W.F.: A weak invariance principle with applications to domains of attraction. Ann. Probab. **6**, 294–315 (1978)
13. Stout, W.F.: Almost sure invariance principles when $EX_1^2 = \infty$. Z. Wahrscheinlichkeitstheor. Verw. Geb. **49**, 23–32 (1979)

Received January 28, 1986; in revised form October 17, 1988