On asymptotic behaviour of the increments of sums of i.i.d. random variables from domains of attraction of asymmetric stable laws.

Terterov M., St-Petersburg State University

March, 2010
Let $X, X_1, X_2, \ldots$ be a sequence of independent identically distributed (i.i.d.) random variables. Put $S_n = X_1 + \ldots + X_n$, $S_0 = 0$.

Let $a_n$ be a nondecreasing sequence of natural numbers. We will study the asymptotic behaviour of the increments of sums

$$T_n = S_{n+a_n} - S_n$$

as well as the maximal increments

$$U_n = \max_{0 \leq k \leq n-a_n} (S_{k+a_n} - S_k).$$

The aim is to describe a normalizing sequence $c_n$ such that

$$\limsup \frac{T_n}{c_n} = 1 \quad a.s.$$ $$\limsup \frac{U_n}{c_n} = 1 \quad a.s.$$
L. Shepp (1964) $T_n = \frac{S_{n+a_n}-S_n}{a_n}$, $a_n \nearrow \infty$, $a_n$ takes positive integer values. $M_t = \mathbb{E}e^{Xt} < \infty$. $T = \limsup T_n$ was determined in terms of the moment generating function of $X$ and the radius of convergence of $\sum x^{a_n}$ (denoted $r$).

$m(a) = \min M(t)e^{-at}$.

$T = a$ a.s., where $a = a(r)$ is the unique solution of $m(a) = r$. 
P. Erdős, A. Rényi (1970) $a_n = [c \log n]$. Theorem 1. Suppose that the moment generating function $M_t = \mathbf{E} e^{X_t}$ exists for $t \in I$, where $I$ is an open interval containing $t = 0$. Let us suppose that $\mathbf{E} X = 0$. Let $\alpha$ be any positive number such that the function $M(t)e^{-\alpha t}$ takes on its minimum in some point in the open interval $I$ and let us put

$$\min_{t \in I} M(t)e^{-\alpha t} = M(\tau)e^{-\alpha \tau} = e^{-1/c}.$$ 

Then

$$\mathbf{P}(\lim \max_{0 \leq k \leq n-[c \log n]} \frac{S_{k+[c \log n]} - S_k}{[c \log n]} = \alpha) = 1$$
Theorem 2. The functional dependence between $\alpha$ and $c = c(\alpha)$ determines the distribution of the random variables $X_n$ uniquely.

Practical implements.
1. The longest runs of pure heads.

Theorem 3 (special case of Theorem 1). Let $X_1, X_2, \ldots$ be independent Bernoulli random variables with $P(X_i = 1) = 0.5 = P(X_i = -1)$, $S_n = X_1 + \ldots + X_n$. Then for any $c \in (0, 1)$ there exists $n_0 = n_0(c)$ such that

$$\max_{0 \leq k \leq [c \log_2 n]} (S_k + [c \log_2 n] - S_k) = [c \log_2 n] \quad a.s.$$ 

if $n > n_0$.

This theorem guarantees the existence of a run of length $[c \log_2 n]$ when $n$ is large enough.
2. The stochastic geyser problem. $X_1, X_2, \ldots$ - i.i.d.r.v., $F(.)$ is their distribution function. Put $V_n = S_n + R_n$, where $R_n$ is also a r.v. sequence.

Theorem (Bártfai, 1966). Assume that the moment generating function of $X_1$ exists in a neibourhood of $t = 0$ and $R_n = o(\log n)$. Then, given the values of $\{V_n; n = 1, 2\ldots\}$, the distribution function $F(.)$ is determined with probability 1, i.e. there exists a r. v. $L(x) = L(V_1, V_2, \ldots, x)$, measurable with a respect of $\sigma$-algebra, generated by $V_1, V_2\ldots$ such that for any given real $x$, $L(x) = F(x)$.

Proof. For any $c > 0$ we have

$$\lim_{\max_{0 \leq k \leq n - \lfloor c \log n \rfloor}} \frac{V_{k + \lfloor c \log n \rfloor} - V_k}{\lfloor c \log n \rfloor} = \alpha(c) \quad a.s.$$
Improvements.

The existence of a moment generating function is a necessary condition. If $M(t) = \mathbb{E}e^{Xt} = \infty$ for all $t > 0$, then

$$\limsup_{0 \leq k \leq n - [c \log n]} \max_{0 \leq k \leq n - [c \log n]} \frac{S_{k+[c \log n]} - S_k}{[c \log n]} = \infty \quad a.s.$$ 


$$\max_{0 \leq k \leq n-a_n} \frac{S_{k+a_n} - S_k}{\gamma(c)a_n} \xrightarrow{a.s.} 1,$$

where $\gamma(c)$ is a constant depending on $c$ and $M(t)$ remains true when $a_n/\log n \to 0$. (Erdős and Rényi had $a_n/\log n \sim c$).
M. Csörgő and J. Steinebach (1981). Theorem. Suppose $\mathbb{E}X = 0$, $\mathbb{E}X^2 = 1$ and there exists a $t_0 > 0$ such that $M(t) = \mathbb{E}e^{Xt} < \infty$ if $|t| < t_0$. Then for the sums $S_n$ the following holds

$$
\lim_{n \to \infty} \max_{0 \leq k \leq n-a_n} \frac{S_{k+a_n} - S_k}{(2a_n \log(n/a_n))^{1/2}} = 1 \quad \text{a.s.,}
$$

where $\frac{a_n}{(\log n)^2} \to \infty$.

In this case the normalizing sequence depends only on the moment conditions on $X$. 

\[ T_n = S_{n+c_n} - S_n \]

\[ U_n = \max_{0 \leq k \leq n-a_n} (S_{k+a_n} - S_k), \quad \limsup \frac{U_n}{c_n} = 1 \quad \text{a.s.} \]

The asymptotic behavior of \( U_n \) and \( T_n \) strongly depends on the rate of the growth of \( a_n \) and the moment conditions on \( X \).

If \( a_n = O(\log n) \), the normalizing sequence \( c_n \) depends on the distribution of \( X \) (Erdős-Rényi laws).

If \( a_n / \log n \to \infty \) and \( \mathbb{E}X = 0, \mathbb{E}X^2 = 1 \), the normalising sequence does not depend on the distribution of \( X \) and is the same as the one for the Gaussian distribution. In this case \( c_n = \sqrt{2a_n(\log(n/a_n) + \log \log n)} \) (Csörgő-Révész laws).

For example: put \( a_n = n \), \( c_n = (2n \log \log n)^{1/2} \), \( U_n = S_n \),

\[ \limsup \frac{S_n}{\sqrt{2n \log \log n}} = 1 \quad \text{a.s.} \]
Frolov (2000).
It turned out, that these two types of behaviour are particular cases of the universal one. For variables with a finite moment generating function there exists an explicit formula for the normalizing sequence $c_n$. 
H. Lanzinger, U. Stadtmuller.
Let $X, X_1, X_2, \ldots$ be a sequence of i.i.d. random variables. Suppose $E X = 0$, $E X^2 = \sigma^2$. $E e^{t|X|^{1/p}} < \infty$ for all $t$ in a neighbourhood of 0.

$$t_0 = \sup\{t \geq 0 : E e^{g(tX)} < \infty\} \in (0, \infty)$$

$$\varphi(c) = \max\{x + y : \frac{x^2}{2c\sigma^2} + (t_0 y)^{\frac{1}{p}} \leq 1, x \geq 0, y \geq 0\}.$$  

Theorem.
Under assumptions made above, we have

$$\lim_{n \to \infty} \max_{0 \leq j < n} \max_{1 \leq k \leq n-j} \frac{S_{j+k} - S_j}{\varphi\left(\frac{k}{(\log n)^{2p-1}}\right)(\log n)^p} = 1 \text{ a.s.}$$
Corollary

\[
\limsup_{n \to \infty} \max_{0 \leq j < n} \frac{S_j + c(\log n)^{2p-1} - S_j}{\varphi(c)(\log n)^p} = 1 \quad \text{a.s.}
\]


Theorem.

\[
\limsup_{n \to \infty} \frac{S_n + (\log n)^p - S_n}{(\log n)^{(p+1)/2}} = \varphi(1) \quad \text{a.s.}
\]
Definition. Suppose that $X$ has a distribution $R$. The distribution $R$ is stable if for every $n$ there exist $c_n > 0$ and $\gamma_n$ such that $S_n = c_n X + \gamma_n$. $c_n = n^{1/\alpha} c$, $0 < \alpha \leq 2$. Normal distribution is stable with $\alpha = 2$ and $\gamma_n = 0$. The distribution function $G$ belongs to the domain of attraction of $R$ if there exist a sequence $B_n, B_n > 0$ and $A_n$, such that

$$\frac{S_n - A_n}{B_n} \xrightarrow{d} R.$$
There exists a canonical representation of the characteristic function of a stable law.

\[ f(t) = \exp(it\gamma - c|t|^\alpha(1 - i\frac{t}{|t|}\beta\omega(t, \alpha))), \]

where \( \gamma \in \mathbb{R}, \ c \geq 0, \ |\beta| \leq 1, \ \omega(t, \alpha) = \tan \frac{\pi \alpha}{2} \) if \( \alpha \neq 1 \) and \( \omega(t, \alpha) = (2/\pi) \log t \), if \( \alpha = 1 \).
Let $X, X_1, X_2, \ldots$ be a sequence of i.i.d. random variables, $EX = 0$, $F(x) = P(X < x)$. Suppose $F(x)$ to be from a domain of attraction of a stable law with index $\alpha \in (1, 2)$ and the characteristic function $\psi(t) = \exp\{-a|t|^\alpha(1 + i\frac{t}{|t|} \tan \frac{\pi}{2} \alpha)\}$, 

$a = \cos(\pi(2 - \alpha)/2)$. Let $B_n = n^{\frac{1}{\alpha}}$.

Define, further

$c_n = (\log n)^{\frac{p+\alpha-1}{\alpha}}, \quad t_0 = \sup\{t \geq 0 : Ee^{t(X^+)^{\frac{\alpha}{p+\alpha-1}}} < \infty\}$,

$\varphi(c) = \max\{x + y : \frac{(\alpha - 1)x^{\frac{\alpha}{\alpha-1}}}{\alpha c^{\frac{1}{\alpha-1}}} + t_0 y^{\frac{\alpha}{p+\alpha-1}} \leq 1, x \geq 0, y \geq 0\}$.

Theorem. Suppose $t_0 \in (0, \infty)$

Then

$$\limsup_{n \to \infty} \frac{S_{n+ca_n} - S_n}{c_n\varphi(c)} = 1 \quad \text{a.s.}$$
References.


