Tail and moment estimates for sums of independent random variables with logarithmically concave tails

by

D. GLUSKIN (Tel Aviv) and S. KWAPIĘŃ (Warszawa)

Abstract. The random variables $S = \sum_{i=1}^{\infty} \alpha_i \xi_i$, where $\{\xi_i\}$ is a sequence of symmetric, independent, identically distributed random variables such that $\ln P(|\xi_i| \geq t)$ is a concave function we give estimates from above and from below for the tail and moments of $S$. The estimates are exact up to a constant depending only on the distribution of $\xi$. They extend results of S. J. Montgomery-Smith [MS], M. Ledoux and M. Talagrand [LT, Chapter 4.1] and P. Hitczenko [H] for the Rademacher sequence.

Notations and definitions. If $N$ is a convex, nondecreasing function on $\mathbb{R}^+$ with $N(0) = 0$ and $a = (\alpha_i)$ is a sequence of real numbers we define the conjugate function $N^*: \mathbb{R}^+ \to \mathbb{R}^+$ by

$$N^*(t) = \sup\{st - N(s) : s \in \mathbb{R}^+\}$$

and

$$\|a\|_N = \inf\left\{ t : N^*(t) \leq \frac{1}{N(t)} \right\},$$

$$\|a\|^*_N = \sup\left\{ \sum_{i=1}^{\infty} \alpha_i \beta_i : \sum_{i=1}^{\infty} N(|\beta_i|) \leq 1 \right\}.$$ 

The following inequalities hold true (cf. [KR, Chapter 2.9, inequality (9.24)]):

$$\|a\|_{N^{-}} \leq \|a\|_{N} \leq \frac{1}{r} \|a\|_{N^{r}}.$$ 

If $N(t) = t^r$ then $\|a\|_N = (\sum_{i=1}^{\infty} |\alpha_i|^r)^{1/r}$ and it will be denoted by $\|a\|_{r}$. If $a = (\alpha_i)$ is a sequence converging to 0 then we denote by $a^*$ the nonincreasing rearrangement of $(|\alpha_i|)$.

Given any $s \geq 1$ and a sequence $a$ we denote by $a^*$ the sequence $(\beta_i)$ defined by $\beta_i = \alpha_i^*$ for $i \leq s$ and $\beta_i = 0$ for $i > s$ and by $a_+$ the sequence $(\delta_i)$ defined by $\delta_i = 0$ for $i \leq s$ and $\delta_i = \alpha_i^*$ for $i > s$.

1991 Mathematics Subject Classification: 60G50, 60E15.

Research of the first author partially supported by the U.S.-Israel Binational Sciences Foundation.
For a real number $s$ we denote by $\lfloor s \rfloor$ the largest integer which does not exceed $s$ and by $\lceil s \rceil$ the smallest integer which is not less than $s$.

From now on we fix a convex function $N$ as above and let $(\xi_i)$ be a sequence of symmetric, independent random variables each with distribution given by $P(\xi_i \geq t) = e^{-N(t)}$ for $t > 0$.

If $N(t) \equiv t^r$ we refer to the sequence $(\xi_i)$ as a symmetric Weibull sequence with exponent $r$. In the particular case of $r = 1$ the sequence $(\xi_i)$ will be denoted by $(\eta_i)$.

We denote by $(\varepsilon_i)$ a Bernoulli sequence, i.e., a sequence of i.i.d. symmetric random variables taking on values $\pm 1$. It can be viewed as the sequence $(\xi_i)$ corresponding to the function $N$ which is equal to 0 on $[0, 1]$ and $\infty$ on $(1, \infty)$.

For $s > 0$ we abbreviate $(s^{-1}N)^s$ by $M_s$, i.e. $M_s(t) = s^{-1}N^s(st)$, and we define

$$K(s, a) = \max\{\|a^s\|_{M_s}, \sqrt{s}\|a\|_2\}$$

for a sequence $a$.

Finally, let us define two constants depending on $N$:

$$\kappa_1 = \inf\left\{ c > 0 : \int_{s>0} e^{-N(t)} dt \leq \int_s e^{-ct} dt \text{ for all } s > 0 \right\},$$

$$\kappa_2 = E|\xi_1| = \int_0^\infty e^{-N(t)} dt.$$

By Karamata's Theorem (cf. [MO, Chapter 16.B.4.a]), $E\phi(\xi_1) \leq E\phi(\kappa_1\eta_1)$ for each nonnegative, convex function $\phi$ on $\mathbb{R}$ and hence, by [MO, Chapter 11.F]) (cf. also [B, Chapter 1], or [KW, Chapter 3.1]), we easily see that for each sequence $\eta_i$,

$$\left( E\left| \sum_{i \geq s} \delta_i \xi_i \right|^s \right)^{1/s} \leq \kappa_1 \left( E\left| \sum_{i \geq s} \delta_i \eta_i \right|^s \right)^{1/s}.$$

**Theorem 1.** For each $1 \leq s < \infty$ and each sequence $a = (a_i)$ we have for $S = \sum_{i=1}^\infty a_i \xi_i$,

$$cK(s, a) \leq (E|S|^s)^{1/s} \leq CK(s, a)$$

with $c = \min\{\kappa_1/2, 1/(2e)\}$ and $C = 3 + 2\kappa_1(N(1) + 1)$.

**Proof.** First we prove that if $s = 2k$ for some positive integer $k$ then for each sequence $\eta_i$,

$$\left( E\left| \sum_{i \geq s} \delta_i \eta_i \right|^s \right)^{1/s} \leq \max\{s\|a\|_\infty, \sqrt{s}\|a\|_2\}.$$

Indeed, since $E|\eta_i|^s = 1$ for each positive integer $i$, by simple computation we check that

$$E\left| \sum_{i \geq s} \delta_i \eta_i \right|^s = \frac{[2k]!}{k!} E\left( \sum_{i \geq s} \delta_i^2 \eta_i \right)^k.$$

Hence, the expression in (4) is a convex function of $(\delta_i^2)$ and therefore its maximum under the constraints $|\eta_i| \leq s^{-1}$ for all $i$, $\|\delta_i\|_2 \leq 1/\sqrt{s}$ is attained when $\delta_i = s^{-1}$ for $s$ different $i$'s and $\delta_i = 0$ for the remaining $i$'s. So this maximum raised to the power $1/s$ equals

$$\left( \frac{1}{k!} \sum_{i \geq s} \delta_i^2 \eta_i \right)^{1/s} = \left( \frac{[2k]!}{k!} E\left( \sum_{i \geq s} \delta_i^2 \eta_i \right)^k \right)^{1/(2k)} \leq \frac{1}{2k} \left( \frac{2k(k+1)(k+2)\ldots(3k-1)}{2k} \right)^{1/(2k)} \leq 1.$$

where the second equality holds since $\sum_{i \geq s} \delta_i^2 \eta_i$ has the distribution $\gamma_{2k,1}$.

Now, for $s \geq 1$, put $k = \lceil s/2 \rceil$ and assume that $K(s, a) \leq 1$. Then $\|a^s\|_{M_s} \leq 1$ and by the definition of $M_s$,

$$1 \geq \sum_{i \geq s} M_s(\beta_i) \geq \sum_{i \geq s} M_s(a_i) = \sum_{i \geq s} a_i(\beta_i - s^{-1}N(1)).$$

Hence, $|\eta_i| \leq (N(1) + 1)/|a_i|$ for each positive integer $i$, which gives $2k\|a\|_\infty \leq 2(\ln N(1) + 1)$. Moreover, $K(s, a) \leq 1$ implies that $\sqrt{2k}\|a\|_2 \leq 2$. Therefore, by (3) we get

$$\left( E\left| \sum_{i \geq s} \delta_i \xi_i \right|^s \right)^{1/s} \leq \left( E\left| \sum_{i \geq s} \delta_i \eta_i \right|^{2k} \right)^{1/(2k)} \leq 2(N(1) + 1).$$

Combining this with (2) we obtain

$$\left( E\left| \sum_{i \geq s} \delta_i \xi_i \right|^s \right)^{1/s} \leq \kappa_1 2(N(1) + 1)K(s, a).$$

Now, we will estimate the remaining part of $S$, namely we will prove that

$$\left( E\left| \sum_{i \geq s} \beta_i \xi_i \right|^s \right)^{1/s} \leq 3\|a^s\|_{M_s}.$$

Let $(\varepsilon_i)$ be a Bernoulli sequence independent of $(\eta_i)$. Then $\xi_i \sim \varepsilon_i N(|\eta_i|)^{-1}$ and therefore

$$\left( E\left| \sum_{i \geq s} \beta_i \varepsilon_i \xi_i \right|^s \right)^{1/s} = \left( E\left| \sum_{i \leq s} \beta_i \xi_i N(|\eta_i|)^{-1} \right|^s \right)^{1/s}. $$
Since $aN(y)^{-1} \leq M(x) + s^{-1}y$ for all $x, y > 0$, the Contraction Principle (cf. [LT, Chapter 4.2, Lemma 4.6]) yields

$$
\left( E \sum_{i \leq s} \alpha_i \xi_i \right)^{s} \leq \left( E \sum_{i \leq s} \beta_i \left( M_n(\beta_i) + s^{-1} \eta_i \right) \right)^{s} \leq \sum_{i \leq s} M_n(\beta_i) + s^{-1} \left( E \sum_{i \leq s} \eta_i \right)^{s} \leq \sum_{i \leq s} M_n(\beta_i) + \frac{2k}{s} \cdot \frac{1}{2k} s \cdot \left( E \sum_{i \leq s} \eta_i \right)^{2k/(2k)} \leq 3,
$$

by (5), if $\|a^a\|_{M_n} \leq 1$ and $k = \lfloor s/2 \rfloor$. Hence by homogeneity we get (7), which combined with (6) proves the right side inequality of theorem.

To prove the left side inequality let $(\gamma_i)$ be such that $\sum_{i=1}^{\infty} s^{-1} N(\gamma_i) = 1$ and

$$
\sum_{i=1}^{\infty} \alpha_i \gamma_i = \sum_{i=1}^{s} \beta_i \gamma_i = \|a^a\|_{-1:M_n}.
$$

Then

$$
\left( E \sum_{i=1}^{\infty} \alpha_i \xi_i \right)^{s} \geq \left( E \sum_{i=1}^{s} \beta_i \xi_i \right)^{s} \geq \left( E \sum_{i=1}^{s} \beta_i \xi_i \right)^{s} \geq \|a^a\|_{-1:M_n} 2^{-s} \exp \left( - \sum_{i=1}^{\infty} s^{-1} N(\gamma_i) \right).
$$

Hence, by (1) we obtain

$$
\left( E \sum_{i=1}^{\infty} \alpha_i \xi_i \right)^{s} \geq \frac{1}{2e} \|a^a\|_{M_n}.
$$

By Jensen’s Inequality and Remark 1 from [HK] we have

$$
\left( E \sum_{i=1}^{\infty} \alpha_i \xi_i \right)^{s} \geq E[\xi_1] \left( E \sum_{i=1}^{\infty} \alpha_i \xi_i \right)^{s} \geq \frac{\|a^a\|_{M_n}}{2s} E[\xi_1^2] \geq \frac{\|a^a\|_{M_n}}{2s} \sqrt{s} \|a^a\|_{2}.
$$

This together with (8) proves the left side inequality of theorem.

The estimates of Theorem 1 lead quickly to a tail estimate for $S$. Namely, for each $\lambda > 0$ by Chebyshev’s Inequality we get for all $s \geq 1$

$$
P(|S| \geq \lambda K(s, a)) \leq \lambda^{-s} E[|S|^{s}] / K(s, a)^{s} \leq (C/\lambda)^{s}.
$$

In particular, if we put $\lambda = Ce$ we obtain

$$
P(|S| \geq CeK(s, a)) \leq e^{-\alpha}.
$$

To prove estimates from below we need the inequality

$$
(10) \quad K(r, u) \leq 6K(s, a) \quad \text{for all } 1 \leq s < r \leq 2s.
$$

To see this, since obviously

$$
\sqrt{s} \|a^a\|_{2} \geq \frac{1}{\sqrt{2}} \|a^a\|_{2},
$$

it is enough to show that $\|a^a\|_{M_n} / \|a^a\|_{M_n} \geq 1/6$. This is proved as follows: by convexity of $N$, for all $t > 0$,

$$
\frac{1}{s} N(t) \geq \frac{1}{s} N \left( \frac{t}{s} \right),
$$

which yields $M_r(\xi^*) \leq M_s(\xi^*)$ for all $t > 0$ and implies that $\|b\|_{M_n} \leq \frac{1}{s} \|b\|_{M_n}$ for each sequence $\xi$. This together with the obvious inequality $\|a^a\|_{M_n} \leq 3 \|a^a\|_{M_n}$ proves (10).

Now, by the Paley-Zygmund inequality (cf. [K], Chapter 1.6), we obtain

$$
P \left( |S| \geq \frac{C}{2} K(s, a) \right) \geq P \left( |S| \geq \left( \frac{1}{2} \right)^{s} E[|S|] \right) \geq \left( 1 - \left( \frac{1}{2} \right)^{s} \right)^{2} \|a^a\|_{M_n} \left( E[|S|] \right)^{2} \geq \left( 1 - \left( \frac{1}{2} \right)^{s} \right)^{2} \left( \frac{C}{2} \right)^{2s} \left( K(s, a) \right)^{2s} \geq \left( \frac{c}{12C} \right)^{2s}.
$$

Hence, if we define $F(t, a) = \max \{ s : K(s, a) \leq t \}$ then we arrive at

**Corollary.** There are positive constants $c_1, c_2$ depending only on the constants $c, C$ from Theorem 1 such that for all $t \geq K(1, a)$,

$$
e^{-F(c_1 t, a)} \leq P \left( |S| > t \right) \leq e^{-F(c_2 t, a)}.
$$

**Proof.** The right inequality with $c_2 = 1/(Ct)$ follows by (9) and the definition of $F$. To prove the left inequality let $C(t) = \inf \{ s : K(s, a) \geq t \}$. Then by (11) we have

$$
P \left( |S| > t \right) \geq \left( \frac{a}{12C} \right)^{2C(2t/a)}
$$

and it is enough to show that

$$
2 \ln \frac{12C}{c} G \left( \frac{2t}{c} \right) \leq F(c_1 t)
$$
for some $c_1$ and all $t \geq K(1, a)$. This follows easily by (10) with

$$c_1 = \frac{12}{c} \left( \frac{12c}{c} \right)^{\frac{\log 6}{c}}.$$

**Remark 1.** If $(\xi_i)$ is a symmetric Weibull sequence with exponent $r$ and $N(t) \equiv t^r$ and it is easy to compute that

$$K(s, a) = \max \{ c, s^{1/r} \|a\|_{r, \infty}, \sqrt{s} \| a_s \|_2 \},$$

where $r = r/(r-1)$ and $c_0 = (1/r)^{1/r}(1/r)^{1/r'}$. In this case the function $K(s, a)$ and hence $F(t, a)$ are explicitly computable for sequences such as $(1/t^2)$.

Theorem 1 is of interest for interpolation theory because it gives an equivalent formula for the $K$-functional interpolating the norms $\| \cdot \|_{r, \infty}$ and $\| \cdot \|_2$.

**Remark 2.** In the general case, the function $K(s, a)$ is computable for the sequence $a = a^n$ which is given by $a_1 = a_2 = \ldots = a_n = 1$, $a_{i+1} = 0$ for $i > n$. And it is easy to see that

$$F(t, a) \sim \begin{cases} t^2/n & \text{if } t < n, \\ G(t, n) & \text{if } t \geq n, \end{cases}$$

where $G(x) = \inf \{ y : N^*(y) - N^*(y) > 0 \}$, and $F \sim H$ means that there are universal constants $d_1, d_2, d_3$ which depend on the distribution of $\xi$ and which do not depend on $n$ such that $H(d_1 t, n) \leq F(t, n) \leq H(d_2 t, n)$ for all $t > d_3$.

It is of interest to compare the function $F(t, n)$ with the function appearing in the Large Deviation Theorem.

**Remark 3.** If we replace $(\xi_i)$ by the sequence $(|\xi_i| - E|\xi_i|)$ then we obtain similar estimates for moments and tails. This follows by the result for the sequence $(\xi_i)$ and the symmetrization inequalities as in [LT, Chapter 6.1].

**Remark 4.** Also, in the above results the convexity condition on $N$ can be relaxed to the following one: there exists a constant $\kappa_3$ such that

$$N(\kappa_3 xy) \leq y N(x) \text{ for all } x > 0 \text{ and } 1 \geq y \geq 0.$$

In this case we have $\kappa_3 \|a\|_{N, \infty} \leq \|a\|_{N}$ instead of (1) and the proof of Theorem 1 repeats with $c$ modified to $c = \min\{\kappa_2/2, \kappa_3/(2c)\}$. Similarly we modify the proof of the Corollary to fit this case.

**References**
