STUDIA MATHEMATICA 114 (3) (1995)

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

by

E. D. GLUSKIN (Tel Aviv) and S. KWAPIEŃ (Warszawa)

Abstract. For random variables $S = \sum_{i=1}^{\infty} \alpha_i \xi_i$, where (ξ_i) is a sequence of symmetric, independent, identically distributed random variables such that $\ln P(|\xi_i| \ge t)$ 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 ξ. 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

$$\|a\|_N = \inf\Big\{t : \sum_{i=1}^{\infty} N(t^{-1}|\alpha_i|) \le 1\Big\},$$

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

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

$$(1)\qquad \|a\|_{N^*} \le \|a\|_N^* \le 2\|a\|_{N^*}.$$

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 $a^* = (\alpha_i^*)$ the nonincreasing rearrangement of $(|\alpha_i|)$.

Given any $s \ge 1$ and a sequence a we denote by a^s the sequence (β_i) defined by $\beta_i = \alpha_i^*$ for $i \le s$ and $\beta_i = 0$ for $i > s$ and by a_s the sequence (δ_i) defined by $\delta_i = 0$ for $i \le s$ 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 (ξ_i) be a sequence of symmetric, independent random variables each with distribution given by $P(|\xi_i|\ge t)=e^{-N(t)}$ for $t>0$.

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

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

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

$$K(s,a)=\max\{\|a^s\|_{M_s},\sqrt{s}\|a_s\|_2\}\quad \text{for a sequence } a.$$

Finally, let us define two constants depending on N:

$$\kappa_1=\inf\Big\{c>0:\ \int_s^\infty e^{-N(t)}\,dt\le\int_s^\infty e^{-t/c}\,dt \text{ for all } s>0\Big\},$$

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

By Karamata's Theorem (cf. [MO, Chapter 16.B.4.a]), $E\phi(\xi_i)\le E\phi(\kappa_1\eta_i)$ for each nonnegative, convex function ϕ 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 (δ_i),

$$\Big(E\Big|\sum_{i>0}\delta_i\xi_i\Big|^s\Big)^{1/s}\le\kappa_1\Big(E\Big|\sum_{i>0}\delta_i\eta_i\Big|^s\Big)^{1/s}. \tag{2}$$

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

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

with $c=\min\{\kappa_2/2,1/(2e)\}$ 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 (δ_i),

$$\Big(E\Big|\sum_{i>0}\delta_i\eta_i\Big|^s\Big)^{1/s}\le\max\{s\|(\delta_i)\|_\infty,\sqrt{s}\|(\delta_i)\|_2\}. \tag{3}$$

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

$$E\Big|\sum_{i>0}\delta_i\eta_i\Big|^s=\frac{(2k)!}{k!}E\Big(\sum_{i>0}\delta_i^2|\eta_i|\Big)^k. \tag{4}$$

Hence, the expression in (4) is a convex function of (δ_i^2) and therefore its maximum under the constraints $|\delta_i|\le s^{-1}$ for all i, $\|(\delta_i)\|_2\le1/\sqrt{s}$ 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

$$\begin{aligned}\frac{1}{s}\Big(E\Big(\sum_{i\le s}\eta_i\Big)^s\Big)^{1/s}&=\frac{1}{2k}\Big(\frac{(2k)!}{k!}E\Big(\sum_{i=1}^{2k}|\eta_i|\Big)^k\Big)^{1/(2k)}\\&=\frac{1}{2k}(2k(k+1)(k+2)\ldots(3k-1))^{1/(2k)}\\&\le\frac{1}{2k}\cdot\frac{2k+(k+1)+\ldots+(3k-1)}{2k}=1,\end{aligned} \tag{5}$$

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

Now, for $s\ge1$, put $k=\lceil s/2\rceil$ and assume that $K(s,a)\le1$. Then $\|a^s\|_{M_s}\le1$ and by the definition of M_s,

$$1\ge\sum_{i\le s}M_s(\beta_i)\ge\lfloor s\rfloor M_s(\beta_{\lfloor s\rfloor})\ge\lfloor s\rfloor(\beta_{\lfloor s\rfloor}-s^{-1}N(1)).$$

Hence, $|\delta_i|\le(N(1)+1)/\lfloor s\rfloor$ for each positive integer i, which gives $2k\|a_s\|_\infty\le2(N(1)+1)$. Moreover, $K(s,a)\le1$ implies that $\sqrt{2k}\|a_s\|_2\le2$. Therefore, by (3) we get

$$\Big(E\Big|\sum_{i>s}\delta_i\eta_i\Big|^s\Big)^{1/s}\le\Big(E\Big|\sum_{i>s}\delta_i\eta_i\Big|^{2k}\Big)^{1/(2k)}\le2(N(1)+1).$$

Combining this with (2) we obtain

$$\Big(E\Big|\sum_{i>s}\delta_i\xi_i\Big|^s\Big)^{1/s}\le\kappa_12(N(1)+1)K(s,a). \tag{6}$$

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

$$\Big(E\Big|\sum_{i\le s}\beta_i\xi_i\Big|^s\Big)^{1/s}\le3\|a^s\|_{M_s}. \tag{7}$$

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

$$\Big(E\Big|\sum_{i\le s}\beta_i\xi_i\Big|^s\Big)^{1/s}=\Big(E\Big|\sum_{i\le s}\beta_i\varepsilon_iN(|\eta_i|)^{-1}\Big|^s\Big)^{1/s}.$$

Since $xN(y)^{-1} \le M_s(x) + s^{-1}y$ for all $x, y > 0$, the Contraction Principle (cf. [LT, Chapter 4.2, Lemma 4.6]) yields

$$\Big(E\Big|\sum_{i\le s}\beta_i\xi_i\Big|^s\Big)^{1/s} \le \Big(E\Big|\sum_{i\le s}\varepsilon_i(M_s(\beta_i)+s^{-1}|\eta_i|)\Big|^s\Big)^{1/s}$$
$$\le \sum_{i\le s}M_s(\beta_i) + s^{-1}\Big(E\Big|\sum_{i\le s}\eta_i\Big|^s\Big)^{1/s}$$
$$\le \sum_{i\le s}M_s(\beta_i) + \frac{2k}{s}\cdot\frac{1}{2k}\Big(E\Big|\sum_{i\le s}\eta_i\Big|^{2k}\Big)^{1/(2k)} \le 3,$$

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

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

$$\sum_{i=1}^{\infty}\beta_i\gamma_i = \sum_{i=1}^{s}\beta_i\gamma_i = \|a^s\|^*_{s^{-1}N}.$$

Then

$$\Big(E\Big|\sum_{i=1}^{\infty}\alpha_i\xi_i\Big|^s\Big)^{1/s} \ge \Big(E\Big|\sum_{i=1}^{\infty}\beta_i\xi_i\Big|^s\Big)^{1/s}$$
$$\ge \Big(\sum_{i=1}^{\infty}\beta_i\gamma_i\Big)P(\xi_i \ge \gamma_i \text{ for } i \le s)^{1/s}$$
$$= \|a^s\|^*_{s^{-1}N}2^{-\lfloor s\rfloor/s}\exp\Big(-\sum_{i=1}^{\infty}s^{-1}N(\gamma_i)\Big).$$

Hence, by (1) we obtain

$$\Big(E\Big|\sum_{i=1}^{\infty}\alpha_i\xi_i\Big|^s\Big)^{1/s} \ge \frac{1}{2e}\|a^s\|_{M_s}. \tag{8}$$

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

$$\Big(E\Big|\sum_{i=1}^{\infty}\alpha_i\xi_i\Big|^s\Big)^{1/s} \ge E|\xi_1|\Big(E\Big|\sum_{i=1}^{\infty}\alpha_i\varepsilon_i\Big|^s\Big)^{1/s} \ge \frac{\kappa_2}{2}\sqrt{s}\|a_s\|_2.$$

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 $s \ge 1$,

$$P(|S| \ge \lambda K(s,a)) \le \lambda^{-s}E|S|^s/K(s,a)^s \le (C/\lambda)^s.$$

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

$$P(|S| \ge CeK(s,a)) \le e^{-s}. \tag{9}$$

To prove estimates from below we need the inequality

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

To see this, since obviously

$$\frac{\sqrt{s}\|a_s\|_2}{\sqrt{r}\|a_r\|_2} \ge \frac{1}{\sqrt{2}},$$

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

$$\frac{1}{r}N(t) \ge \frac{1}{s}N\Big(\frac{s}{r}t\Big),$$

which yields $M_r(t) \le M_s(\frac{r}{s}t)$ for all $t > 0$ and this implies that $\|b\|_{M_r} \le \frac{r}{s}\|b\|_{M_s}$ for each sequence b; this together with the obvious inequality $\|a^r\|_{M_r} \le 3\|a^s\|_{M_r}$ proves (10).

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

$$P\Big(|S| \ge \frac{c}{2}K(s,a)\Big) \ge P\Big(|S|^s \ge \Big(\frac{1}{2}\Big)^s E|S|^s\Big) \tag{11}$$
$$\ge \Big(1-\Big(\frac{1}{2}\Big)^s\Big)^2\frac{(E|S|^s)^2}{E|S|^{2s}}$$
$$\ge \Big(1-\Big(\frac{1}{2}\Big)^s\Big)^2\Big(\frac{c}{C}\Big)^{2s}\Big(\frac{K(s,a)}{K(2s,a)}\Big)^{2s}$$
$$\ge \Big(\frac{c}{12C}\Big)^{2s}.$$

Hence, if we define $F(t,a) = \max\{s : K(s,a) \le t\}$ 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 $t \ge K(1,a)$,*

$$e^{-F(c_1t,a)} \le P(|S| > t) \le e^{-F(c_2t,a)}.$$

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

$$P(|S| > t) \ge \Big(\frac{c}{12C}\Big)^{2G(2t/c)}$$

and it is enough to show that

$$2\ln\frac{12C}{c}G\Big(\frac{2}{c}t\Big) \le F(c_1t)$$

for some c_1 and all $t \geq K(1,a)$. This follows easily by (10) with

$$c_1 = \frac{12}{c}\left(2\ln\frac{12C}{c}\right)^{\ln_2 6}.$$

Remark 1. If (ξ_i) is a symmetric Weibull sequence with exponent r then $N(t) \equiv t^r$ and it is easy to compute that $K(s,a) = \max\{c_r s^{1/r}\|a^s\|_{r^*}, \sqrt{s}\|a_s\|_2\}$, where $r^* = r/(r-1)$ and $c_r = (1/r)^{1/r}(1/r^*)^{1/r^*}$. In this case the function $K(s,a)$ and hence $F(t,a)$ are explicitly computable for sequences such as $(1/i^\lambda)$.

Theorem 1 is of interest for interpolation theory because it gives an equivalent formula for the K-functional interpolating the norms $\|\cdot\|_{r^*}$ 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 $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 1$, $\alpha_i = 0$ for $i > n$. And it is easy to see that

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

where $G(x) = \inf\{y : N^*(1)xy - N^*(y) > 0\}$, and $F \sim H$ means that there are universal constants d_1, d_2, d_3 which depend on the distribution of ξ 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 (ξ_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 (ξ_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 κ_3 such that $N(\kappa_3 xy) \leq yN(x)$ for all $x > 0$ and $1 \geq y > 0$.

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

References

[B] E. Berger, *Majorization, exponential inequalities and almost sure behavior of vector valued random variables*, Ann. Probab. 19 (1990), 1206–1226.

[H] P. Hitczenko, *Domination inequality for martingale transforms of Rademacher sequences*, Israel J. Math. 84 (1993), 161–178.

[HK] P. Hitczenko and S. Kwapień, *On the Rademacher series*, in: Probability in Banach Spaces, Proc. 9th Internat. Conf., Sandbjerg, 1993, Birkhäuser, 1994, 31–36.

[K] J.-P. Kahane, *Some Random Series of Functions*, Heath, 1968.

[KR] M. A. Krasnosel'skiĭ and Ya. B. Rutickiĭ, *Convex Functions and Orlicz Spaces*, Noordhoff, Groningen, 1961.

[KW] S. Kwapień and W. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple*, Birkhäuser, 1992.

[LT] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer, 1991.

[MO] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.

[MS] S. J. Montgomery-Smith, *The distribution of Rademacher sums*, Proc. Amer. Math. Soc. 109 (1990), 517–522.

THE RAYMOND AND BEVERLY SACKLER
FACULTY OF EXACT SCIENCES
TEL AVIV UNIVERSITY
RAMAT AVIV
69-978 TEL AVIV, ISRAEL

INSTITUTE OF MATHEMATICS
WARSAW UNIVERSITY
BANACHA 2
02-097 WARSZAWA, POLAND

Received August 8, 1994 (3322)