ELSEVIER

Statistics & Probability Letters 55 (2001) 367–376

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

The best constant in the Rosenthal inequality for nonnegative random variables

R. Ibragimov[a,*], Sh. Sharakhmetov[b]

[a] *Department of Economics, Yale University, 28 Hillhouse Ave., New Haven, CT 06511, USA*
[b] *Tashkent State Economics University, Tashkent, Uzbekistan*

Received 1 December 1999; received in revised form 1 November 2000

Abstract

In the present paper, we obtain the explicit expression for the best constant in the Rosenthal inequality for nonnegative random variables.

Keywords: Moment inequalities; Nonnegative random variables; Best constant

1. Introduction

Rosenthal (1970) proved the following inequalities (in what follows, $A(t)$, $B(t)$ denote constants depending on t only, L, L_i, $i=1,2$, denote absolute constants, not necessarily the same in different places):

$$E\left(\sum_{k=1}^{n} X_k\right)^t \leqslant A(t)\max\left(\sum_{k=1}^{n} EX_k^t, \left(\sum_{k=1}^{n} EX_k\right)^t\right) \tag{1}$$

for all independent nonnegative random variables $X_1,\ldots,X_n$ with finite tth moment, $1 \leqslant t < \infty$;

$$E\left|\sum_{k=1}^{n} X_k\right|^t \leqslant B(t)\max\left(\sum_{k=1}^{n} E\left|X_k\right|^t, \left(\sum_{k=1}^{n} E\left|X_k\right|^2\right)^{t/2}\right) \tag{2}$$

for all independent mean zero random variables $X_1,\ldots,X_n$ with finite tth moment, $2 \leqslant t < \infty$.

Several studies have focused on problems related to estimation of the constants in inequalities (1) and (2) and their extensions. For example, using Sazonov's (1974) estimate, one can obtain (2) with a constant $B(t)=L^t 2^{t^2/4}$, while from results in Nagaev and Pinelis (1977) and Pinelis (1980) it follows that one can take

* Corresponding author.
E-mail address: rustam.ibragimov@yale.edu (R. Ibragimov).

0167-7152/01/$ - see front matter
PII: S0167-7152(01)00134-1

$B(t) = L^t t^t$. Concerning refinements and extensions of relations (1) and (2) and similar inequalities see also Pinelis (1994), Nagaev (1990, 1997).

Denote by $A^*(t)$ and $B^*(t)$ the best constants in inequalities (1) and (2). Jonhson et al. (1985) showed that the constants $A^*(t)$ and $B^*(t)$ satisfy the inequalities $L_1^t(t/\log t)^t \leqslant A^*(t)$, $B^*(t) \leqslant L_2^t(t/\log t)^t$ (see also Talagrand, 1989; Kwapien and Szulga, 1991). Ibragimov and Sharakhmetov (1998a, b) proved that $B^*(2m) = E(Z(1) - 1)^{2m}$, $m \in N$, where $Z(1)$ is a Poisson random variable with parameter 1. Figiel et al. (1997) and Ibragimov and Sharakhmetov (1995, 1997) independently obtained that the best constant $B^*_{\text{sym}}(t)$ in inequality (2) for symmetric random variables is given by $B^*_{\text{sym}}(t) = 1 + E|N|^t$, $2 < t < 4$, $B^*_{\text{sym}}(t) = E|Z_1(0.5) - Z_2(0.5)|^t$, $t \geqslant 4$, where N is the standard normal random variable, $Z_1(0.5)$ and $Z_2(0.5)$ are independent Poisson random variables with parameter 0.5. Ibragimov and Sharakhmetov (1995, 1997) found the exact asymptotics of the constant $B^*(t)$ as $t \to \infty$.

The present paper deals with determining the best constant $A^*(t)$ in Rosenthal's inequality (1) for nonnegative random variables. The obtained results were announced in Ibragimov and Sharakhmetov (1998a).

The problem of determining the best constants in inequalities (1) and (2) is closely connected with problems of finding the exact estimates for $E|\sum_{i=1}^n X_i|^t$ in terms of moment characteristics of the random variables $X_1, \ldots, X_n$. These problems were considered in Prokhorov (1962), Pinelis and Utev (1984) and Utev (1985). The proof of the main results of the present paper uses some ideas and methods presented in Utev (1985), Hoeffding (1955), Karr (1983) and Zolotarev (1997).

2. Main results

For $d > 0$, let $Z(d)$ denote a random variable with Poisson distribution with parameter d: $P(Z(d) = k) = e^{-d} d^k / k!, \quad k = 0, 1, 2, \ldots$.

Theorem 1. *The best constant $A^*(t)$ in inequality* (1) *is given by* $A^*(t) = 2$, $1 < t < 2$; $A^*(t) = EZ^t(1)$, $t \geqslant 2$.

The proof of Theorem 1 is based on Theorem 2 below which extends the results obtained in Utev (1985) and Ibragimov and Sharakhmetov (1997) to the case of sums of nonnegative random variables and gives explicit bounds on moments of those objects in terms of their particular components.

Let $X_1, \ldots, X_n$ be independent nonnegative r.v.'s with finite tth moment, $1 \leqslant t < \infty$. Fix $a_i > 0$, $b_i > 0$, $a_i^t \leqslant b_i$, $i = 1, \ldots n$, A_t, D, $M_t > 0$. Set

$$M_1(n, t, a, b) = \{(X, n) : EX_i = a_i,\ EX_i^t = b_i,\ i = 1, \ldots, n\},$$

$$M_2(n, t, a, b) = \{(X, n) : EX_i \leqslant a_i,\ EX_i^t \leqslant b_i,\ i = 1, \ldots, n\},$$

$$U_1(A_t, D) = \left\{(X, n) : n \geqslant 1, \sum_{i=1}^n EX_i = D, \sum_{i=1}^n EX_i^t = A_t\right\},$$

$$U_2(A_t, D) = \left\{(X, n) : n \geqslant 1, \sum_{i=1}^n EX_i \leqslant D, \sum_{i=1}^n EX_i^t \leqslant A_t\right\},$$

$$U(M_t) = \left\{(X, n) : n \geqslant 1, \max\left(\left(\sum_{i=1}^n EX_i\right)^t, \sum_{i=1}^n EX_i^t\right) = M_t\right\}.$$

Denote by $U_3(A_t, D)$ and $U_4(A_t, D)$ the subsets of $U_1(A_t, D)$ and $U_2(A_t, D)$, respectively, consisting of identically distributed r.v.'s. Let $\bar{Z}(A_t, D) = Z((D^t/A_t)^{1/(t-1)})$, and let $V_1(t, a_1, b_1), \ldots, V_n(t, a_n, b_n)$ be independent random variables with distributions

$$P(V_i(t,a,b)=0) = 1-(a^t/b)^{1/(t-1)}, \qquad P(V_i(t,a,b)=(b/a)^{1/(t-1)}) = (a^t/b)^{1/(t-1)}.$$

Denote

$$F(n,t,a,b) = E\left(\sum_{i=1}^n V_i(t,a_i,b_i)\right)^t, \quad G(n,t,a,b) = \sum_{i=1}^n (b_i - a_i^t) + \left(\sum_{i=1}^n a_i\right)^t.$$

Theorem 2. *If* $t \geqslant 2$, *then*

$$\sup_{(X,n)\in M_k(n,t,a,b)} E\left(\sum_{i=1}^n X_i\right)^t = F(n,t,a,b), \quad k=1,2, \tag{3}$$

$$\sup_{(X,n)\in U_k(A_t,D)} E\left(\sum_{i=1}^n X_i\right)^t = (A_t/D)^{t/(t-1)} E\bar{Z}^t(A_t,D), \quad k=1,2,3,4. \tag{4}$$

If $1 < t < 2$, *then*

$$\sup_{(X,n)\in M_k(n,t,a,b)} E\left(\sum_{i=1}^n X_i\right)^t = G(n,t,a,b), \quad k=1,2, \tag{5}$$

$$\sup_{(X,n)\in U_k(A_t,D)} E\left(\sum_{i=1}^n X_i\right)^t = A_t + D^t, \quad k=1,2,3,4. \tag{6}$$

3. Preliminaries

Let us formulate some auxiliary results needed for the proof of the theorems.

Let $\mathbb{R}_+ = [0,\infty)$, $\mathfrak{I}$ be the σ-algebra of Borel subsets of $\mathbb{R}_+$, Λ be a class of finite positive σ-additive measures λ on $\mathfrak{I}$ such that $\lambda(\{0\}) = 0$. For a measure $\lambda \in \Lambda$ denote by $T(\lambda)$ a random variable with the characteristic function $E e^{itT(\lambda)} = \exp\left(\int_0^\infty (e^{itx}-1)\,d\lambda(x)\right)$.

Let $\lambda \in \Lambda$. Set (here and in what follows (X,n) denotes a vector of random variables $(X_1,\ldots,X_n)$)

$$W_1(\lambda) = \left\{(X,n)\colon n \geqslant 1, X_i \text{ is nonnegative},\ i=1,\ldots,n, \sum_{i=1}^n P(X_i \in B\backslash\{0\}) = \lambda(B), B \in \mathfrak{I}\right\}.$$

Denote by $W_2(\lambda)$, the subset of $W_2(\lambda)$ consisting of identically distributed r.v.'s. The following theorem refines and complements the results obtained by Utev (1985).

Theorem 3. *Let* $\lambda \in \Lambda$ *and* $f\colon \mathbb{R}_+ \to \mathbb{R}$ *be a continuous and convex function. Let there exist a constant* $C = C(f)$ *such that*

$$|f(a_1+a_2)| \leqslant C(1+|f(a_1)|)(1+|f(a_2)|), \quad a_1, a_2 \in \mathbb{R}_+. \tag{7}$$

If $\int_0^\infty |f(x)|\,\mathrm{d}\lambda(x) < \infty$, *then*

$$E\,|f(T(\lambda))| < \infty, \quad \sup_{(X,n)\in W_1(\lambda)} Ef\left(\sum_{i=1}^n X_i\right) \leqslant Ef(T(\lambda)). \tag{8}$$

If, in addition to that, the function f is nonnegative, then

$$\sup_{(X,n)\in W_k(\lambda)} Ef\left(\sum_{i=1}^n X_i\right) = Ef(T(\lambda)), \quad k=1,2. \tag{9}$$

Let us formulate two lemmas needed for the proof of Theorem 1. For a vector $a\in\mathbb{R}^n$, denote by $a_{(1)} \geqslant \cdots \geqslant a_{(n)}$ its components arranged in non-increasing order.

Definition 1 (Marshall and Olkin (1979)). Let $x, y\in\mathbb{R}^n$. The vector x is majorized by the vector $y (x \prec y)$, if $\sum_{i=1}^k x_{[i]} \leqslant \sum_{i=1}^k y_{[i]}$, $k=1,\ldots,n-1$, $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$.

Definition 2 (Marshall and Olkin (1979)). Let $A \subseteq \mathbb{R}^n$. A function $f: A \to \mathbb{R}$ is called S-convex (resp. S-concave) on A, if $(x \prec y) \Rightarrow (f(x) \leqslant f(y))$ (resp. $(x \prec y) \Rightarrow (f(x) \geqslant f(y))$) for all $x, y\in A$.

Lemma 1. *A continuous function* $f: \mathbb{R}_+ \to \mathbb{R}$ *is convex on* $\mathbb{R}_+$ *if and only if*

$$(n-1)f(x) + f\left(\sum_{i=1}^n a_i + x\right) \geqslant \sum_{i=1}^n f(a_i + x), \quad a_1,\ldots,a_n, \quad x\in\mathbb{R}_+, \quad n \geqslant 1. \tag{10}$$

Proof. Let $f: \mathbb{R}_+ \to \mathbb{R}$ be a continuous and convex function on $\mathbb{R}_+$. From Proposition 3.C.1 in Marshall and Olkin (1979) it follows that $\sum_{i=1}^n f(x_i)$ is S-convex on $\mathbb{R}_+^n$. Since $(a_1+x,\ldots,a_n+x) \prec (x,\ldots,x,\sum_{i=1}^n a_i + x)$, this implies inequality (10). Let now a continuous function $f: \mathbb{R}_+ \to \mathbb{R}$ satisfy inequality (10), and let $0 \leqslant y \leqslant z$. Setting in (10) $n=2$, $x=y$, $a_1=a_2=(z-y)/2$, we obtain that $(f(y)+f(z))/2 \geqslant f((y+z)/2)$, that is the function f is convex. □

Lemma 2. *Let* X *and* Y *be nonnegative random variables*, $f: \mathbb{R}_+ \to \mathbb{R}$ *be a continuous convex function satisfying condition* (7). *Suppose that the random variable* X *has a distribution* $\lambda\in\Lambda$, *the random variables* X, Y *and* $T(\lambda)$ *are independent*, $E|f(X)| < \infty$, $E|f(Y)| < \infty$. *Then* $E|f(T(\lambda) + Y)| < \infty$ *and* $Ef(X+Y) \leqslant Ef(T(\lambda)+Y)$.

Proof. The distribution of the random variable $T(\lambda)$ is the same as the distribution of the random variable $\sum_{i=1}^{Z(1)} X_i$, where $X_1, X_2, \ldots,$ is a sequence of independent random variables with the distribution λ, which are independent of Y and $Z(1)$. According to Utev (1985), from (7) it follows that for all $a_1,\ldots,a_n \in \mathbb{R}$ $|f(\sum_{i=1}^n a_i)| \leqslant q^{n-1}\prod_{i=1}^n (1+|f(a_i)|)$, where $q=\max(1, 2C(f))$. Consequently,

$$E|f(T(\lambda)+Y)| = \mathrm{e}^{-1}\sum_{k=0}^\infty E\left|f\left(\sum_{i=1}^k X_i + Y\right)\right| \Big/ k! \leqslant \mathrm{e}^{-1}(1+E|f(Y)|)\sum_{k=0}^\infty (q(1+E|f(X)|))^k/k!$$

$$= (1+E|f(Y)|)\exp(q(1+E|f(X)|)-1) < \infty.$$

From Lemma 1 it follows that

$$\mathrm{e}^{-1}\sum_{k=1}^{\infty} Ef\left(\sum_{i=1}^{k} X_i + Y\right)\Big/ k!$$

$$\geqslant \mathrm{e}^{-1}\sum_{k=1}^{\infty}\left(\sum_{i=1}^{k} Ef(X_i+Y)-(k-1)Ef(Y)\right)\Big/ k! = Ef(X+Y)-\mathrm{e}^{-1}Ef(Y).$$

Therefore,

$$Ef(T(\lambda)+Y)=\mathrm{e}^{-1}\sum_{k=0}^{\infty} Ef\left(\sum_{i=1}^{k} X_i + Y\right)\Big/ k! \geqslant Ef(X+Y).$$

The proof is complete. □

Proof of Theorem 3. Inequalities (8) are evident consequences of Lemma 2. Let us prove relations (9). Let $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous convex function satisfying condition (7), and let $\lambda \in \Lambda$. It suffices to prove the exactness of upper bounds in (9). Take $n \geqslant \lambda(\mathbb{R}_+)$. Let $X_{1n}, \ldots, X_{nn}$ be independent nonnegative random variables such that $P(X_{in} \in B\backslash\{0\}) = n^{-1}\lambda(B)$ for $B \in \Im$, $i = 1, \ldots, n$. Then $\sum_{i=1}^{n} P(X_{in} \in B\backslash\{0\}) = \lambda(B)$ and the characteristic function of the random variable

$$\sum_{i=1}^{n} X_{in} \text{ is } \left(1 + n^{-1}\int_0^{\infty}(\mathrm{e}^{\mathrm{i}tx}-1)\,\mathrm{d}\lambda(x)\right)^n \underset{n\to\infty}{\to} \exp\left(\int_0^{\infty}(\mathrm{e}^{\mathrm{i}tx}-1)\,\mathrm{d}\lambda(x)\right).$$

Since the function f is continuous from here it follows that $f(\sum_{i=1}^{n} X_{in}) \to f(T(\lambda))$ (in distribution), as $n \to \infty$. By Fatou's lemma,

$$\liminf_{n\to\infty} Ef\left(\sum_{i=1}^{n} X_{in}\right) \geqslant Ef(T(\lambda)).$$

The proof is complete. □

Let d be an arbitrary positive number, $n \geqslant d$. Set

$$H(d) = \left\{(p_1, p_2, \ldots, p_n)\colon 0 \leqslant p_i \leqslant 1,\ i = 1, \ldots, n, \sum_{i=1}^{n} p_i = d\right\}.$$

Let $\bar{X}_1(p_1), \ldots, \bar{X}_n(p_n)$ be independent random variables with distributions $P(\bar{X}_i(p_i) = 1) = p_i$, $P(\bar{X}_i(p_i) = 0) = 1 - p_i$, $i = 1, \ldots, n$ Then for all $(p_1, p_2, \ldots, p_n) \in H(d)$ and all $B \in \Im$ $\sum_{i=1}^{n} P(\bar{X}_i(p_i) \in B\backslash\{0\}) = \lambda(B)$, where λ is such that $\lambda(\{1\}) = \lambda(\mathbb{R}_+) = d$. The distribution of the random variable $T(\lambda)$ is the same as the distribution of the random variable $Z(d)$. Using Theorem 3, we obtain

$$\sup_{(p_1,\ldots,p_n)\in H(d)} E\left(\sum_{i=1}^{n} \bar{X}_i(p_i)\right)^t = \sup_n E\left(\sum_{i=1}^{n} \bar{X}_i(d/n)\right)^t = E(Z(d))^t \tag{11}$$

for $t \geqslant 1$. Let $\Lambda_1(A_t, D) = \{\lambda \in \Lambda\colon \int_0^{\infty} x\,\mathrm{d}\lambda(x) = D, \int_0^{\infty} x^t\,\mathrm{d}\lambda(x) = A_t\}$,

$$\Lambda_2(A_t, D) = \left\{\lambda \in \Lambda\colon \int_0^{\infty} x\,\mathrm{d}\lambda(x) \leqslant D, \int_0^{\infty} x^t\,\mathrm{d}\lambda(x) \leqslant A_t\right\}.$$

It is evident that

$$\sup_{(X,n)\in U_k(A_t,D)} E\left(\sum_{i=1}^{n} X_i\right)^t = \sup_{\lambda\in\Lambda_k(A_t,D)} \sup_{(X,n)\in W_1(\lambda)} E\left(\sum_{i=1}^{n} X_i\right)^t,$$

$$\sup_{(X,n)\in U_{k+2}(A_t,D)} E\left(\sum_{i=1}^{n} X_i\right)^t = \sup_{\lambda\in\Lambda_k(A_t,D)} \sup_{(X,n)\in W_2(\lambda)} E\left(\sum_{i=1}^{n} X_i\right)^t, \quad k=1,2.$$

Using these relations and Theorem 3, we obtain the following lemma.

Lemma 3. *If* $t \geqslant 1$, *then*

$$\sup_{(X,n)\in U_k(A_t,D)} E\left(\sum_{i=1}^{n} X_i\right)^t = \sup_{(X,n)\in U_{k+2}(A_t,D)} E\left(\sum_{i=1}^{n} X_i\right)^t = \sup_{\lambda\in\Lambda_k(A_t,D)} E(T(\lambda))^t, \quad k=1,2.$$

Lemma 4. *If* $t \geqslant 2$, *then the function* $f(t,z,v) = v^{-1/(t-1)}((v^{1/(t-1)}+z)^t - z^t)$ *is concave in* $v>0$ *for* $z>0$. *If* $1<t<2$, *then the function* $g(t,z,v)=(v+z)^t - v^t - z^t$ *is concave in* $v>0$ *for* $z>0$.

Proof. It is not difficult to see that

$$\partial^2 f(t,z,v)/\partial v^2 = (t/(t-1)^2)v^{-1/(t-1)-2}z^t(((1+u_1)^{t-1}-1)-(t-1)u_1(1+u_1)^{t-2}),$$

$$\partial^2 g(t,z,v)/\partial v^2 = t(t-1)v^{t-2}((1+u_2)^{t-2}-1),$$

where $u_1 = v^{1/(t-1)}/z$, $u_2 = z/v$. Since $1+(2-t)u \leqslant (1+u)^{2-t}$, and, therefore, $(t-1)u(1+u)^{t-2} \geqslant (1+u)^{t-1}-1$ for $t\geqslant 2$, $u>0$, and $(1+u)^{t-2}\leqslant 1$ for $1<t<2$, $u>0$, we have $\partial^2 f(t,z,v)/\partial v^2 \leqslant 0$, if $t\geqslant 2$, and $\partial^2 g(t,z,v)/\partial v^2\leqslant 0$, if $1<t<2$. The proof is complete. □

Let $g_0, g_1, \ldots, g_k$ be arbitrary continuous and linearly independent functions on $\mathbb{R}$, $\mathbb{C}=\mathbb{C}(a_1,\ldots,a_k)$ be a set of random variables X satisfying k conditions $Eg_i(X)=a_i$, $i=1,\ldots,k$, $\mathbb{C}_k$ be a subset of $\mathbb{C}$ consisting of random variables X which assume at most $k+1$ value. The following lemma is an evident consequence of the results obtained in Hoeffding (1955) and Karr (1983).

Lemma 5 (Zolotarev, 1986). *If* $\sup_{X\in\mathbb{C}} Eg_0(X)<\infty$, *then* $\sup_{X\in\mathbb{C}} Eg_0(X) = \sup_{X\in\mathbb{C}_k} Eg_0(X)$.

Remark 1. Utev (1985) noted that in the problem of determining extrema of $Eg_0(X)$ over the set of symmetric random variables X with fixed $Eg_i(X)=a_i$, $i=1,\ldots,k$, it suffices to consider only random variables X which assume at most $2k+1$ value. This can be easily proved with the help of some modification of the argument in Hoeffding (1955). It is also not difficult to show that under the condition of nonnegativity of random variables X it suffices to consider only random variables assuming at most $k+1$ value, one of which is zero.

Lemma 6. *Let* X *be a nonnegative random variable*, $EX_i=a$, $EX_i^t=b$, $a,b>0$, $z\geqslant 0$. *If* $t\geqslant 2$, *then*

$$E(X+z)^t \leqslant E(V_1(t,a,b)+z)^t. \quad (12)$$

Proof. By Remark 1, it suffices to consider random variables X taking three values $0, x, y$, where $x, y > 0$. Let $P(X = x) = p$, $P(X = y) = q$, $0 < x < y$, $p + q \leqslant 1$, $xp + yq = a$, $x^t p + y^t q = b$. Since $EX^t = EV^t(t, a, b) = b$, one can assume that $z > 0$. Let $p(x, y) = (y^{t-1} - b/a)/(y^{t-1} - x^{t-1})$, $q(x, y) = (b/a - x^{t-1})/(y^{t-1} - x^{t-1})$. Since $p(x, y) \geqslant 0$, $q(x, y) \geqslant 0$, $p(x, y) + q(x, y) = 1$, $x^{t-1} p(x, y) + y^{t-1} q(x, y) = b/a$, from Lemma 4 it follows that $f(t, z, x^{t-1}) p(x, y) + f(t, z, y^{t-1}) q(x, y) \leqslant f(t, z, b/a)$ for $t \geqslant 2$. It is easy to see that this inequality is equivalent to (12). The proof is complete. □

Lemma 7. *Let* $0 < a_1 \leqslant a_2$, $0 < b_1 < b_2$, $a_i^t \leqslant b_i$, $i = 1, 2$, $z \geqslant 0$. *If* $t \geqslant 2$, *then*

$$E(V(t, a_1, b_1) + z)^t \leqslant E(V(t, a_2, b_2) + z)^t. \tag{13}$$

Proof. It is not difficult to see that (13) is equivalent to the inequality

$$f(t, z, y_1) r \leqslant f(t, z, y_2), \tag{14}$$

where $r = a_1/a_2$, $y_j = b_j/a_j$, $j = 1, 2$. It is evident that $0 < r \leqslant 1$, $y_1 r < y_2$, $y_1, y_2 > 0$. Let $r < 1$. Set $x = (y_2 - r y_1)/(1 - r)$. Under the conditions of the lemma the function $f(v)$ is nonnegative and concave. Consequently,

$$f(t, z, y_1) r \leqslant f(t, z, y_1) r + f(t, z, x)(1 - r) \leqslant f(t, z, y_1 r + x(1 - r)) = f(t, z, y_2).$$

It is easy to see that under the conditions of the lemma the function $f(t, z, v)$ is nondecreasing in $v > 0$. This implies inequality (14) for $r = 1$. The proof is complete. □

Lemmas 6 and 7 imply the following:

Lemma 8. *Let* $X, Y, V(t, a, b)$ *be independent nonnegative random variables*, $EY^t < \infty$, $a, b > 0$, $a^t \leqslant b$. *If* $EX \leqslant a$, $EX^t \leqslant b$, *then* $E(X + Y)^t \leqslant E(V(t, a, b) + Y)^t$ *for* $t \geqslant 2$.

Lemma 9. *If* X, Y *are independent nonnegative random variables*, $EX = a$, $EX^t < \infty$, $EY^t < \infty$, *then* $E(X + Y)^t - EX^t \leqslant E(a + Y)^t - a^t$ *for* $1 < t < 2$.

Proof. It is clear that it suffices to consider the case $Y = z > 0$. Let us show that

$$E(X + z)^t - EX^t \leqslant (a + z)^t - a^t \tag{15}$$

for $1 < t < 2$. It suffices to consider random variables X taking three values $0, x, y$, where $x, y > 0$. Let $P(X = x) = p$, $P(X = y) = q$, $0 < x < y$, $p + q \leqslant 1$, $xp + yq = a$. From Lemma 4 it follows that if $t \geqslant 2$, then

$$\begin{aligned} g(t, z, x) p + g(t, z, y) q &= (g(t, z, x) p/(p + q) + g(t, z, y) q/(p + q))(p + q) \\ &\leqslant g(t, z, a/(p + q))(p + q) = g(t, z, a/(p + q))(p + q) + g(t, z, 0)(1 - p - q) \leqslant g(t, z, a). \end{aligned}$$

It is easy to check that the latter inequality is equivalent to (15). □

Lemma 10. *Let* $a, b > 0$, $a^t \leqslant b$, Y *be a nonnegative random variable with* $EY^t < \infty$. *Let* J *be a set of nonnegative random variables* X *which are independent of* Y *and satisfy the conditions* $EX = a$, $EX^t = b$. *If* $1 < t < 2$, *then*

$$\sup_{X \in J} E(X + Y)^t = b + E(a + Y)^t - a^t. \tag{16}$$

Proof. From Lemma 9 it follows that it suffices to find a sequence of nonnegative random variables X_n which are independent of Y and satisfy the conditions $EX_n = a$, $EX_n^t = b$, $\lim_{n\to\infty} E(X_n + Y)^t = b + E(a + Y)^t - a^t$. If $b = a^t$ then it suffices to take $X_n = a$. Let $a^t < b$. Similarly to the proof of Lemma 7.6 in Utev (1985) set $\delta_n = 1/n$, $P(X_n = a) = 1 - \delta_n$, $P(X_n = b_n) = \delta_n^*$, $P(X_n = 0) = \delta_n - \delta_n^*$, where $\delta_n^* = a\delta_n/b_n$, $b_n = ((b - a^t(1 - \delta_n))/(a\delta_n))^{1/(t-1)}$. It is evident that $b_n \geqslant a$, $0 \leqslant \delta_n^* \leqslant \delta_n$, $\delta_n \to 0$, $b_n \to \infty$, $b_n^t\delta_n^* \to b - a^t$. We have $E(X_n + Y)^t = E(a + Y)^t(1 - \delta_n) + EY^t(\delta_n - \delta_n^*) + (E(b_n + Y)^t - b_n^t)\delta_n^* + b_n^t\delta_n^*$. It suffices to check that $(E(b_n + Y)^t - b_n^t)\delta_n^* \to 0$. Since (see Lemma 7.5 in Utev, 1985) $||1 + x|^t - 1| \leqslant 2^t t(|x| + |x|^t)$ for $t \geqslant 1$, $x \in \mathbb{R}$, we obtain $(E(b_n + Y)^t - b_n^t)\delta_n^* \leqslant b_n^t\delta_n^* 2^t t(EY/b_n + EY^t/b_n^t) \to 0$. The proof is complete. □

Since the function $(a + z)^t - a^t$, $t > 1$, is nondecreasing in $a > 0$ for $z > 0$, we obtain the following:

Lemma 11. *If* $0 \leqslant a_1 \leqslant a_2$, $0 \leqslant b_1 \leqslant b_2$, Y *is a nonnegative random variable*, $EY^t < \infty$, *then* $b_1 + E(a_1 + Y)^t - a_1^t \leqslant b_2 + E(a_2 + Y)^t - a_2^t$, $t > 1$.

Proof of Theorem 2. Subsequently using Lemmas 8, 10 and 11, we obtain (3) and (5). If $t \geqslant 2$, then Lemma 3 and relation (3) imply

$$\sup_{(X,n)\in U_1(A_t,D)} E\left(\sum_{i=1}^n X_i\right)^t = \sup_{(X,n)\in U_3(A_t,D)} E\left(\sum_{i=1}^n X_i\right)^t = \sup_n E\left(\sum_{i=1}^n V_i(t, Dn^{-1}, A_t n^{-1})\right)^t$$
$$= (A_t/D)^{t/(t-1)} \sup_n E\left(\sum_{i=1}^n \bar{X}_i(d/n)\right)^t, \tag{17}$$

where $d = (D^t/A_t)^{1/(t-1)}$, and, in addition to that,

$$\sup_{(X,n)\in U_2(A_t,\ D)} E\left(\sum_{i=1}^n X_i\right)^t = \sup_{(X,n)\in U_4(A_t,\ D)} E\left(\sum_{i=1}^n X_i\right)^t = \sup_{\substack{0<A_t'\leqslant A_t\\ 0<D'\leqslant D}} \sup_{(X,n)\in U_3(A_t',\ D')} E\left(\sum_{i=1}^n X_i\right)^t$$
$$= \sup_n \sup_{\substack{0<A_t'\leqslant A_t\\ 0<D'\leqslant D}} E\left(\sum_{i=1}^n V_i(t, D'n^{-1}, A_t'n^{-1})\right)^t$$
$$= \sup_n E\left(\sum_{i=1}^n V_i(t, Dn^{-1}, A_t n^{-1})\right)^t. \tag{18}$$

From relation (11) it follows that

$$\sup_n E\left(\sum_{i=1}^n \bar{X}_i(d/n)\right)^t = E\bar{Z}^t(A_t, D). \tag{19}$$

Using (17)–(19), we get (4).

If $1 < t \leqslant 2$, then, using Lemma 3 and relation (5), we obtain that

$$\sup_{(X,n)\in U_1(A_t,D)} E\left(\sum_{i=1}^n X_i\right)^t = \sup_{(X,n)\in U_3(A_t,D)} E\left(\sum_{i=1}^n X_i\right)^t = \sup_n (A_t + D^t - D^t/n^{t-1}) = A_t + D^t, \tag{20}$$

$$\sup_{(X,n)\in U_2(A_t,D)} E\left(\sum_{i=1}^{n} X_i\right)^t = \sup_{(X,n)\in U_4(A_t,D)} E\left(\sum_{i=1}^{n} X_i\right)^t = \sup_{\substack{0<A_t'\leqslant A_t \\ 0<D'\leqslant D}} (A_t' + D'^t) = A_t + D^t. \quad (21)$$

(20) and (21) imply relation (6). □

Proof of Theorem 1. Set

$$F(M_t) = \sup_{(X,n)\in U(M_t)} E\left(\sum_{i=1}^{n} X_i\right)^t.$$

Using the evident inequalities

$$\sup_{(X,n)\in U_1(M_t,M_t^{1/t})} E\left(\sum_{i=1}^{n} X_i\right)^t \leqslant F(M_t) \leqslant \sup_{(X,n)\in U_2(M_t,M_t^{1/t})} E\left(\sum_{i=1}^{n} X_i\right)^t$$

and relations (4) and (6), we get that $F(M_t) = 2M_t$ for $1 < t < 2$ and $F(M_t) = EZ^t(1)M_t$ for $t \geqslant 2$. Since $A^*(t) = \sup_{M_t>0} F(M_t)/M_t$, we obtain that $A^*(t) = 2$, $1 < t < 2$; $A^*(t) = EZ^t(1)$, $t \geqslant 2$. The proof is complete. □

Acknowledgements

The authors are grateful to the referee for useful suggestions having led to improvement of the paper.

References

Figiel, T., Hitczenko, P., Johnson, W.B., Schechtman, G., Zinn, J., 1997. Extremal properties of Rademacher functions with applications to the Khintchine and Rosenthal inequalities. Trans. Amer. Math. Soc. 349, 997–1027.

Hoeffding, W., 1955. The extrema of the expected value of a function of independent random variables. Ann. Math. Statist. 26, 268–276.

Ibragimov, R., Sharakhmetov, Sh., 1995. On the best constant in Rosenthal's inequality. Theses of reports of the conference on probability theory and mathematical statistics dedicated to the 75th anniversary of Academician S. Kh. Sirajdinov (Fergana, Uzbekistan), Tashkent, pp. 43–44 (in Russian).

Ibragimov, R., Sharakhmetov, Sh., 1997. On an exact constant for the Rosenthal inequality. Teor. Veroyatnost. i Primenen. 42, 341–350 (translation in Theory Probab. Appl. 42 (1997). 294–302 (1998)).

Ibragimov, R., Sharakhmetov, Sh., 1998a. Exact bounds on the moments of symmetric statistics. Seventh Vilnius Conference on Probability Theory and Mathematical Statistics. 22nd European Meeting of Statisticians. Abstracts of communications. Vilnius, Lithuania, pp. 243–244.

Ibragimov, R., Sharakhmetov, Sh., 1998b. The best constant in Rosenthal's inequality for random variables with zero mean. Theory Probab. Appl., to appear.

Jonhson, W.B., Schechtman, G., Zinn, J., 1985. Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. Ann. Probab. 13, 234–253.

Karr, A.A., 1983. Extreme points of certain sets of probability measures, with applications. Math. Oper. Res. 8, 74–85.

Marshall, A., Olkin, I., 1979. Inequalities: Theory of Majorization and its Applications. Academic Press, New York, 569 pp.

Nagaev, S.V., 1990. On a new approach to the study of the distribution of a norm of a random element in Hilbert space. Probability Theory and Mathematical Statistics. Proceeding of the Fifth Vilnius Conference, Moklas Vilnius, Lithuania, pp. 214–226.

Nagaev, S.V., 1997. Some refinements of probabilistic and moment inequalities. Theory Probab. Appl. 42, 707–713.

Nagaev, S.V., Pinelis, I.F., 1977. Some inequalities for the distributions of sums of independent random variables. Theory of Probab. Appl. 22, 248–256.

Pinelis, I.F., 1980. Estimates for moments of infinite-dimensional martingales. Math. Notes 27, 459–462.

Pinelis, I.F., 1994. Optimum bounds for the distributions of martingales in Banach spaces. Ann. Probab. 22, 1679–1706.

Pinelis, I.F., Utev, S.A., 1984. Estimates of moments of sums of independent random variables. Theory Probab. Appl. 29, 574–577.

Prokhorov, Yu.V., 1962. Extremal problems in limit theorems. Proceedings of the Sixth All-Union Conference on Probability Theory and Mathematical Statistics, Vilnius, pp. 77–84.

Rosenthal, H.P., 1970. On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. Israel J. Math. 8, 273–303.

Sazonov, V.V., 1974. On the estimation of moments of sums of independent random variables. Theory Probab. Appl. 19, 371–374.

Utev, S.A., 1985. Extremal problems in moment inequalities. In Proceedings of the Mathematical Institute of the Siberian Branch of the USSR Academy of Sciences, Moscow, Vol. 5, pp. 56–75 (in Russian).

Zolotarev, V.M., 1997. Modern theory of summation of random variables. VSP, Utrecht, 412 pp.