Acta Applicandae Mathematicae **79:** 35–46, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*


# On Probability and Moment Inequalities for Supermartingales and Martingales *

S. V. NAGAEV
*Sobolev Institute of Mathematics, Novosibirsk, Russia. e-mail: nagaev@math.usc.ru*

(Received: 2 September 2002)

**Abstract.** The probability inequality for sum $S_n = \sum_{j=1}^n X_j$ is proved under the assumption that the sequence $S_k$, $k = \overline{1,n}$, forms a supermartingale. This inequality is stated in terms of the tail probabilities $\mathbf{P}(X_j > y)$ and conditional variances of the random variables $X_j$, $j = \overline{1,n}$. The well-known Burkholder moment inequality is deduced as a simple consequence.

**Mathematics Subject Classification (2000):** 60G42.

**Key words:** filtered probability space, expectation, martingale, supermartingale, Burkholder inequality, Bernstein and Bennet–Hoeffding inequalities, Rosenthal inequality, Fuk inequality, separable Banach space.

## 1. Introduction

Let a sequence of the random variables $S_k$, $k \geqslant 1$, form a supermartingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_k)_{k\geqslant 0}, \mathbf{P})$ with $S_0 = 0$, $\mathcal{F}_0 = \{\oslash, \Omega\}$, i.e.,

$$\mathbf{E}\{S_k / \mathcal{F}_{k-1}\} \leqslant S_{k-1}.$$

Put $X_k = S_k - S_{k-1}$, $k \geqslant 1$. Define the random variables $\sigma_k^2$ by the equalities

$$\sigma_k^2 = \mathbf{E}\{X_k^2 / \mathcal{F}_{k-1}\}.$$

Denote

$$B_k^2 = \sum_1^k \sigma_j^2, \qquad \overline{S}_n = \max_{1\leqslant k\leqslant n} S_k, \qquad \overline{X}_n = \max_{1<k\leqslant n} X_k.$$

Define

$$Q(x) = \mathbf{P}(\overline{X}_n > x) + \mathbf{P}(B_n > x).$$

The main purpose of the present paper is to obtain upper bounds on the probabilities $\mathbf{P}(\overline{S}_n > y)$ in terms of $Q(x)$ generalizing the inequality of Theorem 4 in [7] (see also [13], Theorem 1.10).

* This work was suppored by RFBR (grant 96-01-01529).

THEOREM 1. *Let* $0 < \gamma \leqslant 1$ *and* $t \geqslant \max(e^6, e^4/\gamma^2)$. *Then, for every* $y > 0$,

$$\mathbf{P}(\overline{S}_n > y) < c(t,\gamma)y^{-t}\int_0^y Q(\varepsilon_t u)u^{t-1}\,du, \tag{1}$$

*where*

$$\varepsilon_t = \frac{\ln t - 2\ln\ln t}{2t}, \qquad c(t,\gamma) = \frac{2e^{6\gamma t}}{\gamma}.$$

*If* $\varepsilon_t = \eta/t$, $\eta > 0$, *then inequality* (1) *holds for every* $t > 0$, *with* $c(t,\gamma)$ *replaced by* $te^{3\eta\alpha(\eta)}/\eta\alpha(\eta)$, *where* $\alpha(\eta) = e^{\eta+1}$.

The proof of theorem will be given below, in Section 2. Inequality (1) is close, in form to the main inequality of [15]. The method of the proof is similar to that used in the papers [14, 15]. There are points of contiguity with [18] (see Corollary 3 and Lemma 3 in Section 2).

The upper bounds for $\mathbf{P}(\overline{S}_n > y)$ which extend the corresponding inequalities from [7] to martingales were obtained in 1973 by Fuk [6]. The restrictions imposed by Fuk may seem too strong. It turned out, however, that they are fulfilled, in particular, for the martingale

$$\mathbf{E}\left\{\left\|\sum_1^n X_j\right\| \Big/ \mathcal{F}_k\right\},$$

where $X_j$ are independent random variables taking values in a separable Banach space, $\mathcal{F}_k$ is the $\sigma$-algebra generated by random variables $X_1, X_2, \dots, X_k$, provided

$$\mathbf{E}\|X_j\|^t < \infty, \quad j \in \overline{1,n}$$

(see, in this connection, [12, 17, 23]).

Haeusler [8] generalized one of Fuk's inequalities as follows: *for any* $x$, $u$, $v > 0$,

$$\mathbf{P}(\overline{S}_n > x) < \sum_{i=1}^n \mathbf{P}(X_i > u) + \mathbf{P}(B_n > v) + P(x,u,v), \tag{2}$$

*where*

$$P(x,u,v) = \exp\left(\frac{x}{u}\left(1 - \ln\left(\frac{xu}{v^2}\right)\right)\right).$$

In [11], this result is extended to continuous-time martingales. In [2], $P$ is substituted for

$$P(x,u,v) = \exp\left(\frac{x}{u} - \left(\frac{v^2}{u^2} + \frac{x}{u}\right)\ln\left(\frac{xu}{v^2} + 1\right)\right)$$

(see also [5]).

Let us show that (1) does not follow from (2). Assume for simplicity that, for $x > x_0$,

$$\sum_{i=1}^{n} \mathbf{P}(X_i > x) = \frac{1}{x^\alpha}, \qquad \mathbf{P}(B_n > x) = \frac{1}{x^\alpha}, \quad \alpha > \mathrm{e}^6 - 1.$$

Then $Q(x) < 2/x^\alpha$ for $x > x_0$. Put now $t = \alpha + 1$ into (1). As a result, we obtain the bound

$$\mathbf{P}(\overline{S}_n > x) < cx^{-\alpha}, \tag{3}$$

where the constant $c$ depends on $\alpha$.

Assume now that we want to obtain the bound (3) using inequality (2). Then $v$ must satisfy the condition

$$\mathbf{P}(B_n > v) < \frac{c}{x^\alpha}.$$

The latter will hold for $v > \varepsilon x$, where $\varepsilon = c^{-1/\alpha}$. But then

$$\ln \frac{v^2}{xu} > \ln \frac{\varepsilon^2 x}{u}.$$

As regards $u$, the latter must satisfy the same condition as $v$, i.e., $u > \varepsilon x$. Therefore,

$$\frac{x}{u} \ln \frac{v^2}{xu} \geqslant \varepsilon^{-2} \min_y y \ln y.$$

Hence,

$$P(x, u, v) > \exp\bigl(c^{2/\alpha}\mathrm{e}^{-1}\bigr),$$

i.e., for $x$ large enough, $P(x, u, v) > c/x^\alpha$. Thus, inequality (2) does not allow us to deduce the bound (3) which is true for all $x > x_0$.

Similar arguments show that it is impossible to deduce Burkholder inequality (4) by means of that of Haesler (4). Pinelis [19] has recently extended the Bernstein and Bennet–Hoeffding inequalities to martingales in Banach spaces. In conformity to ordinary martingales, the conditions of Pinelis involve the restriction $B_n^2 < c < \infty$. In [4], one of Fuk's inequalities containing a normal component is extended to a Banach space under the assumption that $\mathbf{E}\|X_j\|^3 < \infty$, $j \in \overline{1, n}$. In addition, the conditions of the same type as those of Fuk are laid on conditional second moments. Generalizations of the Bernstein and Bennet–Hoeffding inequalities are obtained in [21] and [3] as well. Large deviations of $S_n$ are studied in [22] under the condition $\max_{1 \leqslant k \leqslant n} \mathbf{E}|X_k|^t < \infty$. However, main efforts has up to now been concentrated on obtaining moment inequalities (see, in this connection, the survey paper [18]). We turn our attention on one such inequality due to Burkholder [1]:

$$\mathbf{E}^{1/t}|\hat{S}_n|^t < c_t\bigl(D_t^{1/t} + \mathbf{E}^{1/t} B_n^t\bigr). \tag{4}$$

Here $S_k$ is a martingale, $\hat{S}_n = \max_{1\leqslant k\leqslant n}|S_k|$, $D_t = \mathbf{E}(\max_{1\leqslant k\leqslant n}|X_k|^t)$, $t > 2$, $c_t$ is a constant which depends only on $t$. We shall demonstrate now that one can easily deduce inequality (4) by using (1). By multiplying both sides of (1) for $t+1$ by $ty^{t-1}$ and integrating with respect to $y$ from 0 to $\infty$, we obtain

COROLLARY 1. *For any $t$ and $\gamma$ such that $t > \max(\mathrm{e}^6, \mathrm{e}^2/\gamma^2) - 1$ and $0 < \gamma \leqslant 1$,*

$$\mathbf{E}\{\overline{S}_n^t;\ \overline{S}_n \geqslant 0\} < c_0(t,\gamma)\varepsilon_{t+1}^{-t}(\overline{D}_t + \mathbf{E}B_n^t), \tag{5}$$

*where*

$$\overline{D}_t = \mathbf{E}\{\overline{X}_n^t;\ \overline{X}_n \geqslant 0\}, \qquad c_0(t,\gamma) = \frac{2\mathrm{e}^{6\gamma(t+1)}}{\gamma},$$

$$\varepsilon_t = \frac{\ln t - 2\ln\ln t}{2t}.$$

*If $\varepsilon_t = \eta/t$, then inequality (5) holds for all $t > 0$ with $c_0(t,\gamma)$ replaced by $c_1(t,\gamma) = (\eta\alpha(\eta))^{-1}(t+1)\mathrm{e}^{3\eta\alpha(\eta)}$.*

If $(S_k)_{k\geqslant 1}$ is a martingale, then the inequalities in Corollary 1 remain valid for $\mathbf{E}\{|\tilde{S}_n|^t;\ \tilde{S}_n \leqslant 0\}$, where $\tilde{S}_n = \min_{1\leqslant k\leqslant n}\tilde{S}_k$, $\overline{D}_t$ are replaced by

$$\tilde{D}_t = \mathbf{E}\Big\{|X_k|^t;\ \min_{1\leqslant k\leqslant n} X_k \leqslant 0\Big\}.$$

Summing the bounds on

$$\mathbf{E}\{\overline{S}_n^t;\ \overline{S}_n \geqslant 0\}$$

and

$$\mathbf{E}\{|\tilde{S}_n|^t;\ \tilde{S}_n \leqslant 0\},$$

we arrive at inequality (4). From (5) it follows that

$$\overline{\lim}_{t\to\infty} c_t \frac{\ln t}{t} \leqslant 1.$$

Note for comparison that

$$\overline{\lim}_{t\to\infty} c_t \frac{\ln t}{t} = \frac{1}{\mathrm{e}},$$

if $X_k$ is symmetrical and independent [9].

COROLLARY 2. *Let $X_j$ be independent with $\mathbf{E}X_j = 0$. Then, for every $t > 0$, $\eta > 0$, $y \geqslant B_n\varepsilon_t^{-1}$, we have*

$$\mathbf{P}(\overline{S}_n > y) < c(t,\eta)y^{-t}\left(\int_0^y u^{t-1}\sum_1^n \mathbf{P}(X_i > \varepsilon_t u)\,\mathrm{d}u + t^{-1}\varepsilon_t^{-1}B_n^t\right),$$

*where*

$$c(t,\eta) = t\frac{e^{3\eta\alpha(\eta)}}{\eta\alpha(\eta)}, \quad \alpha(\eta) = e^{\eta+1}, \ \varepsilon_t = \frac{\eta}{t}.$$

*For $y < B_n$, the term $t^{-1}\varepsilon_t^{-1}B_n^t$ may be omitted.*

*Proof.* Obviously, $B_n^2 = \sum_1^n \mathbf{E}X_j^2$. Hence,

$$\mathbf{P}(B_n > x) = \begin{cases} 1, & x < \sqrt{\sum_1^n \mathbf{E}X_1^2}, \\ 0, & x \geqslant \sqrt{\sum_1^n \mathbf{E}X_j^2}, \end{cases}$$

and

$$\int_0^y \mathbf{P}(B_n > \varepsilon_t u)u^{t-1}\,du = \begin{cases} 0, & y < B_n, \\ t^{-1}\varepsilon_t^{-t}B_n^t, & y \geqslant B_n. \end{cases}$$

It remains to apply inequality (1) with $\varepsilon_t = \eta/t$. □

*Remark.* Since $\mathbf{P}(\overline{X}_n > x) \leqslant \sum_1^n \mathbf{P}(X_j > x)$, one can replace $\overline{D}_t$ in inequality (5) by $A_t^+ = \sum_1^n \mathbf{E}\{X_j^t;\ X_j \geqslant 0\}$. Respectively, in inequality (4) one can replace $D_t$ by $A_t = \sum_1^n \mathbf{E}|X_j|^t$, making the latter similar to the Rosenthal inequality [20].

## 2. Proof of the Main Result

First, we shall prove several lemmas. We need the following notation:

$$S_{kj} = \sum_{k+1}^{j} X_i, \quad j > k, \qquad B_{kj}^2 = \sum_{k+1}^{j}\sigma_i^2, \qquad a_k = \mathbf{E}(X_k|\mathcal{F}_{k-1}),$$

$$X_k^* = X_k - a_k, \qquad S_{kn}^* = \sum_{i=k+1}^{n} X_i^*.$$

LEMMA 1. *For every $k \geqslant 0$,*

$$\mathbf{P}\Big(\max_{k<j\leqslant n} S_{kj} \geqslant x \;/\; \mathcal{F}_k\Big) \leqslant e^{-hx}\mathbf{E}\big(e^{hS_{kn}^*} \;/\; \mathcal{F}_k\big)$$

*with probability* 1.

*Proof.* Put

$$\tau = \inf\{j : S_{kj}^* \geqslant x,\ k < j \leqslant n\}.$$

Let $\chi_j$ be the the indicator of the event $\{\tau = j\}$. It is easily seen that

$$\mathbf{E}\big(\chi_j e^{hS_{kn}^*} \;/\; \mathcal{F}_j\big) \geqslant \chi_j e^{hS_{kj}^*} \geqslant \chi_j e^{hx}.$$

Taking the conditional expectation with respect to $\mathcal{F}_k$, we have

$$\mathbf{E}\big(\chi_j e^{hS^*_{kn}} / \mathcal{F}_k\big) \geqslant e^{hx}\mathbf{E}(\chi_j / \mathcal{F}_k) = e^{hx}\mathbf{P}(\tau = j / \mathcal{F}_k).$$

Applying this bound, we arrive at the inequality

$$\begin{aligned}\mathbf{E}\big\{e^{hS^*_{kn}} / \mathcal{F}_k\big\} &\geqslant \sum_{j=k+1}^{n} \mathbf{E}\big(\chi_j e^{hS^*_{kn}} / \mathcal{F}_k\big) \geqslant e^{hx} \sum_{j=k+1}^{n} \mathbf{P}(\tau = j / \mathcal{F}_k) \\ &= e^{hx}\mathbf{P}\Big(\max_{k<j\leqslant n} S^*_{kj} \geqslant x / \mathcal{F}_k\Big) \\ &\geqslant e^{hx}\mathbf{P}\Big(\max_{k<j\leqslant n} S_{kj} \geqslant x / \mathcal{F}_k\Big),\end{aligned}$$

and this is equivalent to the conclusion of the lemma. □

LEMMA 2. *Let* $X_j \leqslant y$, $j = \overline{k+1, n}$, $y > 0$, *and* $B_{kn} \leqslant C$. *Then*

$$\inf_h e^{-hx}\mathbf{E}\big(e^{hS^*_{kn}} / \mathcal{F}_k\big) < \exp\left\{\frac{x}{y}\left(1 - \left(1 + \frac{C^2}{y^2}\right)\ln\left(\frac{xy}{C^2} + 1\right)\right)\right\}$$

*with probability* 1.

*Proof.* It is easily seen that

$$\mathbf{E}\big(e^{hS^*_{kn}} / \mathcal{F}_k\big) = \mathbf{E}\left(\prod_{k+1}^{n} f_j(h) Z_{kn}(h) / \mathcal{F}_k\right), \tag{6}$$

where

$$f_j(h) = \mathbf{E}\big\{e^{hX^*_j} / \mathcal{F}_{j-1}\big\}, \qquad Z_{kn}(h) = \prod_{k+1}^{n} \frac{e^{hX^*_j}}{f_j(h)}.$$

Clearly,

$$f_j(h) = e^{-ha_j}\mathbf{E}e^{hX_j} = e^{-ha_j}\big(1 + ha_j + \mathbf{E}\big\{\big(e^{hX_j} - 1 - hX_j\big) / \mathcal{F}_{j-1}\big\}\big).$$

Note that the function $(e^{hx} - 1 - hx)/x^2$ increases with increasing $x$. Consequently,

$$\mathbf{E}\big\{\big(e^{hX_j} - 1 - hX_j\big) / \mathcal{F}_{j-1}\big\} \leqslant \frac{e^{hy} - 1 - hy}{y^2}\sigma_j^2.$$

Hence,

$$\prod_{k+1}^{n} f_j(h) < \exp\left\{\frac{e^{hy} - 1 - hy}{y^2}B^2_{kn}\right\} < \exp\left\{\frac{e^{hy} - 1 - hy}{y^2}C^2\right\}. \tag{7}$$

For all $k < j \leqslant n$,

$$\mathbf{E}\left(\frac{e^{hX^*_j}}{f_j(h)} / \mathcal{F}_{j-1}\right) \equiv 1$$

with probability 1. Therefore,

$$\mathbf{E}(Z_{kn}(h) \,/\, \mathcal{F}_k) = 1. \tag{8}$$

From (6)–(8) it follows that

$$\mathbf{E}(e^{hS^*_{kn}} / \mathcal{F}_k) \leqslant \exp\left\{\frac{e^{hy} - 1 - hy}{y^2} C^2\right\}. \tag{9}$$

Minimizing the function

$$\exp\left\{\frac{e^{hy} - 1 - hy}{y^2} C^2 - hx\right\}$$

with respect to $h$, we arrive at the conclusion of the lemma. □

COROLLARY 3. *Under conditions of Lemma* 2

$$\mathbf{P}\Big(\max_{k<j\leqslant n} S_{kj} \geqslant x \,/\, \mathcal{F}_k\Big) < \exp\left\{\frac{x}{y} - \left(\frac{x}{y} + \frac{C^2}{xy}\right)\ln\left(\frac{xy}{C^2} + 1\right)\right\} := P_0(x; y, C) \tag{10}$$

*a.s.* (*cf.* [19], *Theorem* 8.2 *and* [18], *inequality* (11.27)).

LEMMA 3. *If* $X_j \leqslant y$, $j = \overline{k+1, n}$, *then, for every* $C > 0$,

$$\mathbf{P}\Big(\max_{k<j\leqslant n} S_{kj} \geqslant x \,/\, \mathcal{F}_k\Big) < P_0(x; y, C) + \mathbf{P}(B_{kn} > C \,/\, \mathcal{F}_k) \quad a.s. \tag{11}$$

*Proof.* Let the stopping time $\tau$ of the supermartingale $S_k$ be defined by the equality $\tau = \min\{k : B_k > C\}$. Denote the stopped martingale by $S'_k$. Put $X'_k = S'_k - S'_{k-1}$, $S'_{kj} = \sum_{k+1}^{j} X'_l$. The supermartingale $S'_k$ satisfies the conditions of Lemma 2. Therefore, by (10) we have

$$\mathbf{P}\Big(\max_{k<j\leqslant n} S'_{kj} > x \,/\, \mathcal{F}_k\Big) < P_0(x; y, C).$$

If $B^2_{kn} \leqslant C^2$, then $B^2_{kj} \leqslant C^2$ for all $k < j < n$, i.e., $X'_j = X_j$, $k \leqslant j \leqslant n$, and, therefore,

$$\max_{k<j\leqslant n} S'_{kj} = \max_{k<j\leqslant n} S_{kj}.$$

This means that

$$\mathbf{P}\Big(\max_{k<j\leqslant n} S'_{kj} > x \,/\, \mathcal{F}_k\Big) < P_0(x; y, C) + \mathbf{P}(B_{kn} > C \,/\, \mathcal{F}_k).$$ □

LEMMA 4. *For any positive $t, s, y, C$, we have*

$$\begin{aligned}\mathbf{P}(\overline{S}_n > t+s+y) &< \mathbf{P}(\overline{S}_n > t)P_0(s; y, C) + \\ &\quad + \mathbf{P}(B_n > C) + \mathbf{P}(\overline{X}_n > y).\end{aligned} \tag{12}$$

*Proof.* It is easy to see that

$$\begin{aligned}\{\overline{S}_n > t+s+y\} \subset &\left\{\overline{S}_n > t+s+u,\ \max_{1\leqslant j\leqslant n} X_j \leqslant y\right\} \\ &\cup \left\{\max_{1\leqslant j\leqslant n} X_j > y\right\}.\end{aligned} \tag{13}$$

Next,

$$\begin{aligned}&\left\{\overline{S}_n > t+s+y,\ \max_{1\leqslant j\leqslant n} X_j \leqslant y\right\} \\ &\subset \bigcup_1^n \left\{\tau = k,\ \max_{k<j\leqslant n} S_{kj} \geqslant s,\ \max X_j \leqslant y\right\},\end{aligned} \tag{14}$$

where $\tau = \inf\{k : S_k > t\}$. By Lemma 3, for every $C > 0$,

$$\begin{aligned}&\mathbf{P}\left(\tau = k,\ \max_{k<j\leqslant n} S_{kj} \geqslant s,\ \max_{1\leqslant j\leqslant n} X_j \leqslant y\right) \\ &< P_0(s; y, C)\mathbf{P}(\tau = k) + \mathbf{P}(B_{kn} > C,\ \tau = k).\end{aligned} \tag{15}$$

The conclusion of the lemma follows from (13)–(15). □

Let $\beta$ be any positive number. Consider the sequence

$$y_m = (1+\alpha\varepsilon)^m \beta, \quad m \geqslant 1. \tag{16}$$

Note that

$$y_m - y_{m-1} = \alpha\varepsilon y_{m-1}. \tag{17}$$

Putting $t = y_{m-1}$, $s = (\alpha-1)\varepsilon y_{m-1}$, $C = y = \varepsilon y_{m-1}$ into (13), we have, for $m > 1$,

$$\mathbf{P}(\overline{S}_n > y_m) < \mathbf{P}(\overline{S}_n > y_{m-1})p + Q(\varepsilon y_{m-1}),$$

where $p = p(\alpha) = P_0((\alpha-1)\varepsilon y_{m-1};\ \varepsilon y_{m-1},\ \varepsilon y_{m-1})$. By (10) we have $p(\alpha) < \rho(\alpha) := \exp\{\alpha(\ln\alpha - 1)\}$. Therefore, for $m \geqslant 1$,

$$\mathbf{P}(\overline{S}_n > y_m) < \sum_1^{m-1} Q(\varepsilon y_k)\rho^{m-k-1} + \rho^{m-1}. \tag{18}$$

In what follows, we suppose that $\rho(\alpha) < 1$. By (16) we have

$$k = \frac{\ln(y_k/\beta)}{\ln(1+\alpha\varepsilon)}.$$

Consequently,

$$\rho^k = \left(\frac{y_k}{\beta}\right)^{\ln\rho/\ln(1+\alpha\varepsilon)}. \tag{19}$$

Substituting this expression into (18), we get the bound

$$\mathbf{P}(\overline{S}_n > y_m) < \sum_{1}^{m-1} Q(\varepsilon y_k)\left(\frac{\beta}{y_{m-k-1}}\right)^{s(\alpha,\varepsilon)} + \left(\frac{\beta}{y_{m-1}}\right)^{s(\alpha,\varepsilon)}, \tag{20}$$

where

$$s(\alpha, \varepsilon) = -\frac{\ln\rho(\alpha)}{\ln(1+\alpha\varepsilon)}.$$

By (16) we have

$$y_{m-k-1} = \frac{y_{m+1}}{y_{k+2}}\beta = \frac{y_{m+1}}{y_{k-1}}\left(\frac{y_{k-1}}{y_{k+2}}\beta\right) = (1+\alpha\varepsilon)^{-3}\beta\frac{y_{m+1}}{y_{k-1}}.$$

Next, in view of (17),

$$\frac{1}{y_k - y_{k-1}} = \frac{1}{\alpha\varepsilon y_{k-1}}.$$

Consequently,

$$\begin{aligned} &Q(\varepsilon y_k)\left(\frac{\beta}{y_{m-k-1}}\right)^{s(\alpha,\varepsilon)} \\ &\quad = (1+\alpha\varepsilon)^{3s(\alpha\varepsilon)}(\alpha\varepsilon)^{-1}Q(\varepsilon y_k)y_{k-1}^{s(\alpha,\varepsilon)-1}y_{m+1}^{-s(\alpha,\varepsilon)}(y_k - y_{k-1}) \\ &\quad < \frac{\omega(\alpha,\varepsilon)}{y_{m+1}^{s(\alpha,\varepsilon)}}\int_{y_{k-1}}^{y_k} Q(\varepsilon u)u^{s(\alpha,\varepsilon)-1}\,\mathrm{d}u, \end{aligned} \tag{21}$$

where

$$\omega(\alpha, \varepsilon) = \frac{(1+\alpha\varepsilon)^{3s(\alpha,\varepsilon)}}{\alpha\varepsilon} = \frac{1}{\alpha\varepsilon\rho^3}. \tag{22}$$

From (20) and (21) it follows that

$$\mathbf{P}(\overline{S}_n > y_m) < \frac{\omega(\alpha,\varepsilon)}{y_{m+1}^{s(\alpha,\varepsilon)}}\int_{y_{k-1}}^{y_k} Q(\varepsilon u)u^{s(\alpha,\varepsilon)-1}\,\mathrm{d}u + \left(\frac{\beta}{y_{m-1}}\right)^{s(\alpha,\varepsilon)}. \tag{23}$$

Let $y_m \leqslant y < y_{m+1}$. Then, in view of (23),

$$\begin{aligned}\mathbf{P}(\overline{S}_n > y) < \mathbf{P}(\overline{S}_n > y_m) &< \frac{\omega(\alpha,\varepsilon)}{y^{s(\alpha,\varepsilon)}}\int_0^y Q(\varepsilon u)u^{s(\alpha,\varepsilon)-1}\,\mathrm{d}u + \\ &\quad + \left(\frac{(1+\alpha\varepsilon)\beta}{y}\right)^{s(\alpha,\varepsilon)}.\end{aligned}$$

Hence, since $\beta$ is arbitrary,

$$\mathbf{P}(\overline{S}_n > y) < \frac{\omega(\alpha,\varepsilon)}{y^{s(\alpha,\varepsilon)}} \int_0^y Q(\varepsilon u) u^{s(\alpha,\varepsilon)-1}\,du. \tag{24}$$

Let us now consider the quantities $\omega(\alpha,\varepsilon)$ and $s(\alpha,\varepsilon)$. Put $\alpha = \gamma t/(\ln t - 2\ln\ln t)$, where $\gamma$ is any positive number satisfying the only restriction $\gamma \geqslant 1$. Clearly,

$$\ln \frac{t}{\ln t - 2\ln\ln t} = \ln t - \ln\ln t + \ln\left(1 - \frac{2\ln\ln t}{\ln t}\right). \tag{25}$$

For $t > e^6$,

$$\frac{\ln\ln t}{\ln t} < \frac{\ln 6}{6} < 0.3. \tag{26}$$

Hence,

$$\ln\left(1 - \frac{2\ln\ln t}{\ln t}\right) > -1. \tag{27}$$

If

$$t > \frac{e^4}{\gamma^2}, \tag{28}$$

then

$$\ln\gamma > 2 - \frac{1}{2}\ln t. \tag{29}$$

In view of (25), (27), and (29), $\ln\alpha - 1 > \frac{1}{2}(\ln t - 2\ln\ln t)$, that is $-\ln\rho := \alpha(\ln\alpha - 1) > (\alpha/2)(\ln t - \ln\ln t)$. Hence,

$$s(\alpha,\varepsilon) := -\frac{\ln\rho}{\ln(1+\alpha\varepsilon)} > t \tag{30}$$

if $\varepsilon = (1/2)(\ln t - 2\ln\ln t)/t$, and the condition (28) holds. By (25) we have

$$0 < \ln\alpha < \ln t - \ln\ln t.$$

Hence, by (25) and (26) we have

$$-\ln\rho < \alpha\ln\alpha < \gamma t \frac{\ln t - \ln\ln t}{\ln t - 2\ln\ln t} < 2\gamma t.$$

Therefore, for $\varepsilon = (1/2)(\ln t - 2\ln\ln t)/t$,

$$\frac{1}{\rho^3\alpha\varepsilon} < \frac{2e^{6\gamma t}}{\gamma}. \tag{31}$$

The desired inequality (1) follows from (22), (24), and (31).

Turn now to the case $\varepsilon_t = \eta/t$. It is easily seen that

$$s(\alpha(\eta), \eta/t) = \frac{\alpha(\eta)\eta}{\ln(1 + \alpha(\eta)\eta/t)} > t. \tag{32}$$

On the other hand, in this case, we have

$$\omega(\alpha, \varepsilon_t) = \frac{t e^{3\alpha(\eta)\eta}}{\alpha(\eta)\eta}. \tag{33}$$

Combining (24), (32), and (33), we get the second conclusion of the theorem. □

## References


1. Burkholder, D. L.: Distribution function inequalities for martingales, *Ann. Probab.* **1**(1) (1973), 19–42.
2. Courbot, B.: Rates of convergence in the functional CLT for martingales, *C. R. Acad. Sci. Paris* **328** (1999), 509–513.
3. De la Peña, V. H.: A general class of exponential inequalities for martingales and ratios, *Ann. Probab.* **27**(1) (1999), 537–554.
4. Dehling, H. and Utev, S. A.: An exponential inequality for martingales, *Sib. Adv. Math. J.* **3**(3) (1993), 197–203.
5. Dzhaparidze, H. and van Zanten, J. H.: On Bernstein-type inequalities for martingales, *Stochastic Process. Appl.* **93**(1) (2001), 109–118.
6. Fuk, D. Kh.: Some probabilistic inequalities for martingales, *Siberian Math. J.* **14**(1) (1973), 131–137.
7. Fuk, D. Kh. and Nagaev, S. V.: Some probability inequalities for sums of independent random variables, *Theory Probab. Appl.* **16**(4) (1971), 643–660.
8. Haeusler, E.: An exact rate of convergence in the functional central limit theorem for special martingale difference arrays, *Z. Warsch. verw. Gebiete* **65**(4) (1984), 523–534.
9. Ibragimov, R. and Sharakhmetov, Sh.: On the sharp constant in Rosenthal inequality, *Theor. Probab. Appl.* **42**(2) (1997), 294–302.
10. Johnson, W. B., Schechtman, G. and Zinn, J.: Best constants in moment inequalities for linear combination of independent and exchangeable random variables, *Ann. Probab.* **13** (1985), 234–253.
11. Kubilius, K. and Mémin, J.: Inégalité exponentielle pour les martingales locales, *C. R. Acad. Sci. Paris* **319** (1994), 733–737.
12. Lesign, E. and Volny, D.: Large deviations for martingales, *Stochastic Process. Appl.* **66**(1) (2001), 143–159.
13. Nagaev, S. V.: Large deviations of sums independent random variables, *Ann. Probab.* **7**(5) (1979), 745–789.
14. Nagaev, S. V.: Probability inequalities for sums of independent random variables taking values in a Banach space (in Russian), In: *Limit Theorems of the Theory of Probabilities and Related Topics*, Proc. Inst. Math. Siberian Branch USSR Acad. Sci. 1, 1982, pp. 159–167.
15. Nagaev, S. V.: Probability inequalities for sums of independent random variables with values in a Banach space, *Siberian. Math. J.* **28**(4) (1987), 652–664.
16. Nagaev, S. V.: Some refinements of probabilistic and moment inequalities, *Theory Prob. Appl.* **42**(4) (1997), 707–713.
17. Nagaev, S. V. and Pinelis, I. F.: On large deviations for sums of independent random variables taking values in a Banach space, In: *Abstr. Comm. Second Vilnius Conf. Probab. Theory and Math. Stat.*, Vol. 2, Vilnius, 1977, pp. 66–67.

18. Peshkir, G. and Shiryaev, A. N.: The Khinchine inequalities and martingale extending of sphere of their action, *Russian Math. Surveys* **50**(5) (1995), 849–1099.
19. Pinelis, I. F.: Optimum bounds for the distributions of martingales in Banach spaces, *Ann. Probab.* **22** (1994), 1679–1706.
20. Rosenthal, H. P.: On the subspaces of $L^p (p > 2)$ spanned by sequences of independent random variables, *Israel J. Math.* **8**(3) (1970), 273–303.
21. Van de Geer, S.: Exponential inequalities for martingales with applications to maximum likelihood estimation for counting process, *Ann. Statist.* **23**(5) (1995), 1779–1801.
22. Volodin, N. A. and Morozova, L. N.: Some estimates of probabilities of large deviations for martingales and sums of random vectors (in Russian), In: *Stochastic Processes and Mathematical Statistics*, FAN, Tashkent, 1978, pp. 35–43.
23. Yurinskii, V. V.: Exponential inequalities for large deviations, *Theor. Probab. Appl.* **19**(1) (1974), 152–153.