Liet. matem. rink, **46***, spec. nr.*, 2006, 433–438

An exact bound for tail probabilities for a class of conditionally symmetric bounded martingales

Dainius DZINDZALIETA (MII)
e-mail: dainiusda@yahoo.com

Abstract. We consider the class, say $\mathcal{M}_{n,sym}$, of martingales $M_n = X_1 + \cdots + X_n$ with conditionally symmetric bounded differences X_k such that $|X_k| \leqslant 1$. We find explicitly a solution, say $D_n(x)$, of the variational problem $D_n(x) \stackrel{def}{=} \sup_{M_n \in \mathcal{M}_{n,sym}} \mathbb{P}\{M_n \geqslant x\}$. We show that this problem is equivalent to one when you want to find out the symmetric random walk with bounded length of steps which maximizes the probability to visit an interval $[x; \infty]$. The function $x \mapsto D_n(x)$ allows a simple description and is closely related to the binomial tail probabilities. We can interpret the result as a *final* and *optimal* upper bound $\mathbb{P}\{M_n \geqslant x\} \leqslant D_n(x)$, $x \in \mathbb{R}$, for the tail probability $\mathbb{P}\{M_n \geqslant x\}$.

Keywords: tail probabilities, martingales, random walk, isoperimetric.

1. Introduction and results

In this paper we solve the variational problem

$$D_n(x) \stackrel{def}{=} \sup_{M_n \in \mathcal{M}_{n,sym}} \mathbb{P}\{M_n \geqslant x\}, \tag{1.1}$$

for $x \in \mathbb{R}$, where $\mathcal{M}_{n,sym}$ stands for the class of martingales $M_n = X_1 + \cdots + X_n$ with conditionally symmetric bounded differences $X_k = M_k - M_{k-1}$ (we set $M_0 = 0$). Recall that the differences X_k of a martingale M_n satisfy $\mathbb{E}(X_k \mid \mathcal{F}_{k-1}) = 0$ with respect to an increasing family $\emptyset = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}n \subset \mathcal{F}$ of σ-algebras of a measurable space $(\Omega, \mathcal{F})$. We assume that the differences are bounded (that is, $|X_k| \leqslant 1$) and conditionally symmetric, that is, the conditional distribution of X_k, given M_{k-1}, is symmetric.

Our methods are similar in spirit to a method used by Bentkus (2001) to obtain $\sup_{M_n \in \mathcal{M}_n} \mathbb{P}\{M_n \geqslant x\}$, for $x \in \mathbb{Z}$, without the symmetry assumption. we would like to note that in the paper we expose some results obtained in a joint research project with V. Bentkus.

The result can be applied to describe random walks maximizing the probability to visit an interval, some dominant models in measure concentration phenomena, random graphs theory and etc.

We can interpret $W_n = \{0, X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_n\} = \{0, M_1, \ldots, M_n\}$ as an n step random walk starting at 0. Let $P_x(W_n)$ be the probability to visit an interval $[x, \infty]$ at most in n steps, that is, $P_x(W_n) = \mathbb{P}\{\max_{0 \leqslant k \leqslant n} M_k \geqslant x\}$.

For integer x, we can reformulate 1.1 as the isoperimetric problem

$$D_n(x) \equiv P_x(R_n^x) = \sup_{W_n, sym} \mathbb{P}_x(W_n), \tag{1.2}$$

where $R_n^x = \{0, \varepsilon_1, \varepsilon_1 + \varepsilon_2, \ldots, \varepsilon_1 + \cdots + \varepsilon_n\}$ is the symmetric simple random walk starting at zero ($\varepsilon, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are i.i.d. Rademacher's r.v., that is, $\mathbb{P}\{\varepsilon = 1\} = \mathbb{P}\{\varepsilon = -1\} = \frac{1}{2}$). In other words, among symmetric random walks with bounded steps, the symmetric simple random walk maximizes the probability to visit an interval $[x, \infty]$, $x \in \mathbb{Z}$.

For all $x \in \mathbb{R}$ we can reformulate 1.1 as an isoperimetric problem related to random walks as well, We change some lengths of a simple symmetric random walk R_n^x, so that 1.2 still holds for all $x \in \mathbb{R}$.

We get that D_n is constant on intervals $(m2^{-n+1}; (m+1)2^{-n+1}]$, so we construct random walks for $x = m2^{-n+1}$, $m \in \mathbb{Z}$. For $x \in (m2^{-n+1}; (m+1)2^{-n+1})$ the random walk is the same as for $x = (m+1)2^{-n+1}$, $m \in \mathbb{Z}$. For a simplicity we construct a random walk $W_n^x = R_n^x - x = \{-x, \varepsilon_1 - x, \varepsilon_1 + \varepsilon_2 - x, \ldots, \varepsilon_1 + \cdots + \varepsilon_n - x\}$, that is, we start our walk not from zero, but from $-x$.

We call a set $E \subset \mathbb{R}$ absorbing, if random walk stays in E after its visit. Let's introduce an increasing family of sets $Z = E_1 \subset E_2 \subset \cdots \subset E_{n-1} \subset E_n$, as E_1 – the set of integer numbers, and $E_m = e_m$: $\{2e_m \in E_{m-1}\}$.

We get that this is a sequence of absorbing sets for a random walk W_n^x. Let $E_{m(y)}$ be the absorbing set that $y \in E_{m(y)}$, but $x \notin E_{m(y)-1}$. If the walker is on the absorbing set E_1 then it continue the walk in the same manner as for integer x. If the position y of the walker after k steps is not on the absorbing set E_1 then it goes to the nearest even number if $n - k$ is odd and to the nearest odd number otherwise or due to the symmetry makes a step of the same length to the other side, that is, it gets on a set E_1 or $E_{m(y)-1}$ with equal probabilities 1/2, for all $k \in \mathbb{Z}$ and $y \in \mathbb{R}$, except the special case when $n - k$ is even and $y \in (0; 1)$. In this case the walker goes to the nearest integer number or makes a step of the same length to the other side, that is, it gets on a set E_1 or $E_{m(y)-1}$, with equal probabilities 1/2. In other words, all the time walker tries to get on the absorbing set E_1 by going to the nearest even or odd number (in special case to the nearest integer number), but due to the symmetry condition it has to go to the other side, on the set $E_{m(y)-1}$ with equal probabilities.

Define the stopping time $\tau(x)$ of a sequence $\{\varepsilon_i\}$ as the number of steps after which random walk R_n^x first time visits an interval $[x; \infty)$. We get that

$$M_n(x) = R^x_{n \wedge \tau(x)}.$$

We give an explicit description of $D_n(x)$ for $x \in \mathbb{R}$ using a discrete random variable, say Y_n, which assumes 2^n values, say

$$0 \equiv y_0^{(n)} < y_1^{(n)} < \cdots < y_{2^n-1}^{(n)} \equiv n,$$

with equal probabilities $2^{-n} = \mathbb{P}\{Y_n = y_k^{(n)}\}$, for all $k = 0, 1, \ldots, 2^n - 1$. Then D_n can be represented as the survival function $\mathbb{P}\{Y_n \geqslant x\} = D_n(x)$ of Y_n. To define the values of Y_n, we use the Cartesian product $U \times V$ of vectors $U = (u_1, \dots, u_p) \in \mathbb{R}^p$ and $V = (v_1, \dots, v_q) \in \mathbb{R}^q$ representing it as

$$U \times V = (u_1, \dots, u_{p,} v_1, \dots, v_q).$$

To define the values $y_k^{(n)}$ of Y_n, it is convenient to use $(2n-1)$-dimensional vectors, say

$$Z^{(n)} = (z_1^{(n)}, \dots, z_{2^n-1}^{(n)}),$$

setting

$$\begin{aligned}
y_0^{(n)} &= 0, \\
y_1^{(n)} &= 2^{-n}\big|(z_1^{(n)})\big|_1, \\
y_2^{(n)} &= 2^{-n}\big|(z_1^{(n)}, z_2^{(n)})\big|_1, \\
&\vdots \\
y_{2^n-1}^{(n)} &= 2^{-n}\big|(z_1^{(n)}, \dots, z_{2^n-1}^{(n)})\big|_1,
\end{aligned}$$

where $|S|_1 = |s_1| + \cdots + |s_m|$ stands for the l_1-norm of the vector $S = (s_1, \dots, s_m)$.

For example we have

$$\begin{aligned}
Z^{(1)} &= \big(\{2\}\big), \\
Z^{(2)} &= \big(\{2, 2\}, \{4\}\big), \\
Z^{(3)} &= \big(\{2, 2, 4\}, \{2, 2, 4\}, \{8\}\big), \\
Z^{(4)} &= \big(\{2, 2, 4, 2, 2, 4\}, \{2, 2, 4, 8\}, \{2, 2, 4, 8\}, \{16\}\big), \\
Z^{(5)} &= \big(\{2, 2, 4, 2, 2, 4, 2, 2, 4, 8\}, \{2, 2, 4, 2, 2, 4, 2, 2, 4, 8\}, \\
&\qquad \{2, 2, 4, 8, 16\}, \{2, 2, 4, 8, 16\}, \{32\}\big).
\end{aligned}$$

We use auxiliary braces $\{\dots\}$ to emphasize the block structure of the vector $Z^{(n)}$.

To complete the definition of the $y_k^{(n)}$, we have to define the vectors $Z^{(n)}$. The vector $Z^{(n)} = (S_1^{(n)}, \dots, S_n^{(n)})$ has the block structure. Each $S_k^{(n)}$ is a $\binom{n}{s(n,k)}$-dimensional vector, which each coordinate being the powers of 2. Here $s(n, k) = \big[\frac{n-k+1}{2}\big]$. We set $Z^{(1)} = S_1^{(1)} = (2)$ and define $Z^{(n+1)}$ inductively using $Z^{(n)}$. The definition depends on whether n is even or odd.

If n is odd we define

$$\begin{aligned}
S_1^{(n+1)} &= S_1^{(n)} \times S_1^{(n)}, \\
S_2^{(n+1)} \equiv S_3^{(n+1)} &= S_2^{(n)} \times S_3^{(n)},
\end{aligned}$$

$$\vdots$$

$$S_{n-1}^{(n+1)} \equiv S_n^{(n+1)} = S_{n-1}^{(n)} \times S_n^{(n)},$$

$$S_{n+1}^{(n+1)} = (2^{n+1}).$$

If n is even we define

$$S_1^{(n+1)} \equiv S_2^{(n+1)} = S_1^{(n)} \times S_2^{(n)},$$

$$\vdots$$

$$S_{n-1}^{(n+1)} \equiv S_n^{(n+1)} = S_{n-1}^{(n)} \times S_n^{(n)},$$

$$S_{n+1}^{(n+1)} = (2^{n+1}).$$

It is clear from the definition of $Z^{(n)}$ that $|Z^{(n)}|_1 = \sum_{k=1}^n |S_k^{(n)}|_1 = n \cdot 2^n$ and $|S_k^{(n)}|_1 = 2^n$.

We have

$$D_n(x) = \mathbb{P}\{Y_n \geqslant x\}, \quad \text{for all } \ x \in \mathbb{R}. \tag{1.3}$$

2. Proof

Let's prove the theorem first for $x \leqslant 0$ or $x > n$. If $x \leqslant 0$ then $D_n(x) = 1$. Taking $X_i \equiv 0$ we get $D_n(x) \geqslant 1$, but it is clear that $D_n(x) \leqslant 1$. If $x > n$ then $D_n(x) = 0$, and $D_n(x) \leqslant 0$,since $X_i \leqslant 1$ but we always have $D_n(x) \geqslant 0$.

For $0 < x \leqslant n$ we use induction in n.

The case $n = 1$. In this case $D_1(x) = \frac{1}{2}$. It is obvious since $\sup_{M_1} \mathbb{P}\{X_1 \geqslant x\} = \frac{1}{2} \sup_{M_1} \mathbb{P}\{|X_1| \geqslant x\} \leqslant \frac{1}{2}$(here and later for a simplicity instead of $M_n \in \mathcal{M}_{n,sym}$ we write only M_n). Taking X_1 such that $\mathbb{P}\{X_1 = \pm 1\} = \frac{1}{2}$, we also have that $\sup_{M_1} \mathbb{P}\{X_1 \geqslant x\} \geqslant \frac{1}{2}$.This concludes the proof for $n = 1$.

The case $n > 1$. First we prove that $D_n(x) \leqslant \mathbb{P}\{Y_n \geqslant x\}$. By the induction assumption the equality (1) holds for $1, \dots, n-1$. In particular we have

$$\sup_{M_{n-1}} \mathbb{P}\{M_{n-1} \geqslant x\} = \mathbb{P}\{Y_{n-1} \geqslant x\}, \tag{2.1}$$

for all $x \in \mathbb{R}$.

Using (2.1) and conditioning on X_1, we have

$$\sup_{M_n} \mathbb{P}\{M_n \geqslant x\} = \sup_{M_{n-1}} \mathbb{E}\, \mathbb{P}\left\{M_{n-1} \geqslant x - X_1 \mid X_1\right\} \leqslant \sup_{M_1} \mathbb{E}\, D_{n-1}(x - X_1).$$

Due to the block structure of $Z^{(n)}$ and the definition of Y_n we have

$$\mathbb{P}\{Y_n \geqslant k\} - \mathbb{P}\{Y_n \geqslant x\} \geqslant \mathbb{P}\{Y_n \geqslant 2k - x + 1\} - \mathbb{P}\{Y_n \geqslant k + 1\}, \tag{2.2}$$

for $k \leqslant x \leqslant k+1$ and $k \in \mathbb{N}$.

$$\mathbb{P}\{Y_n \geqslant x\} - \mathbb{P}\{Y_n \geqslant y\} \geqslant \mathbb{P}\{Y_n \geqslant x+k\} - \mathbb{P}\{Y_n \geqslant y+k\}, \tag{2.3}$$

for all $x, y \in \mathbb{R}$, such that $x \leqslant y$ and $k \in \mathbb{N}$.

Applying (2.2) and (2.3) to $\sup_{t\in[0;1]}[D_{n-1}(x-t) + D_{n-1}(x+t)]$ we get that if n is even then

$$\begin{aligned}
&\sup_{t\in[0;1]}\Big[\mathbb{P}\big\{Y_{n-1}(x-t)\big\} + \mathbb{P}\big\{Y_{n-1}(x+t)\big\}\Big] \\
&= \left\{\begin{array}{l}
\mathbb{P}\big\{Y_{n-1} \geqslant x+\{x\}\big\} + \mathbb{P}\big\{Y_{n-1} \geqslant x-\{x\}\big\} \\
\qquad\qquad \text{if } x\in[0;1/2) \text{ or } x\in[2k-1;2k), \quad k\in\mathbb{N} \\
\mathbb{P}\big\{Y_{n-1} \geqslant x-(1-\{x\})\big\} + \mathbb{P}\big\{Y_{n-1} \geqslant x+(1-\{x\})\big\} \\
\qquad\qquad \text{if } x\in[1/2;1) \text{ or } x\in[2k;2k+1), \quad k\in\mathbb{N}
\end{array}\right\} \\
&= 2\,\mathbb{P}\{Y_n \geqslant x\},
\end{aligned}$$

if n is odd then

$$\begin{aligned}
&\sup_{t\in[0;1]}\Big[\mathbb{P}\big\{Y_{n-1}(x-t)\big\} + \mathbb{P}\big\{Y_{n-1}(x+t)\big\}\Big] \\
&= \left\{\begin{array}{l}
\mathbb{P}\big\{Y_{n-1} \geqslant x+\{x\}\big\} + \mathbb{P}\big\{Y_{n-1} \geqslant x-\{x\}\big\} \\
\qquad\qquad \text{if } x\in[2k;2k+1), \quad k\in\mathbb{N} \\
\mathbb{P}\big\{Y_{n-1} \geqslant x-(1-\{x\})\big\} + \mathbb{P}\big\{Y_{n-1} \geqslant x+(1-\{x\})\big\} \\
\qquad\qquad \text{if } x\in[2k-1;2k), \quad k\in\mathbb{N}
\end{array}\right\} \\
&= 2\,\mathbb{P}\{Y_n \geqslant x\}.
\end{aligned}$$

By Lemma 1 we get:

$$\begin{aligned}
\mathbb{E}\,D_{n-1}(x-X_1) &= \int_{-1}^{1} D_{n-1}(x-t)L(X_1)(\mathrm{d}t) \\
&\leqslant \frac{1}{2}\sup_{t\in[0;1]}\big[D_{n-1}(x-t) + D_{n-1}(x+t)\big].
\end{aligned}$$

So we get

$$D_n(x) \leqslant \mathbb{E}\,D_{n-1}(x-X_1) \leqslant \mathbb{P}\{Y_n \geqslant x\}.$$

To prove that $D_n(x) \geqslant \mathbb{P}\{Y_n \geqslant x\}$ we use a martingale $M_n(x)$ defined in the introduction such that $\mathbb{P}\{M_n \geqslant x\} = \mathbb{P}\{Y_n \geqslant x\}$.

3. Auxiliary result

LEMMA 1. *Let μ be a symmetric probability measure, such that $\mu([-1;1]) = 1$. For every measurable function $f(t)$ we have:*

$$\int_{-1}^{1} f(t)\mu(\mathrm{d}t) \leqslant \frac{1}{2} \sup_{t\in[0;1]} \big[f(t) + f(-t)\big].$$

Proof.

$$\begin{aligned}\int_{-1}^{1} f(t)\mu(\mathrm{d}t) &= \frac{1}{2}\int_{-1}^{1} \big(f(t) + f(-t)\big)\mu(\mathrm{d}t)\\ &\leqslant \frac{1}{2}\int_{-1}^{1} \sup_{t\in(0;1]} \big[f(t) + f(-t)\big]\mu(\mathrm{d}t)\\ &= \frac{1}{2}\sup_{t\in[0;1]} \big[f(t) + f(-t)\big]\int_{-1}^{1}\mu(\mathrm{d}t) = \frac{1}{2}\sup_{t\in[0;1]} \big[f(t) + f(-t)\big].\end{aligned}$$

References

1. V. Bentkus, An inequality for large devation probabilities of sums of bounded i.i.d. random variables, *Lith. Math. J.*, **41**, 144–153 (2001).
2. V. Bentkus, On Hoeffding's inequalities, *Ann. Probab.*, **32**, 1650–1673 (2004).
3. W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.*, **58**, 13–30 (1963).
4. S. Karlin, W.J. Studden, Tchebycheff systems: With applications in analysis and statistics, in: *Pure and Applied Mathematics*, vol. XV, Interscience Publishers John Wiley & Sons, New York-London-Sydney (1966).

REZIUMĖ

D. Dzindzalieta. Tikslus rėžis sąlyginai aprėžtų martingalų klasės uodegų tikimybėms

Mes nagrinėjame martingalų $M_n = X_1 + \ldots + X_n$ klasę $\mathcal{M}_{n,sym}$, kurių skirtumai X_k yra sąlyginai simetriniai ir aprėžti, kaip kad $|X_k| \leqslant 1$. Mes išreikštine forma gauname variacinio uždavinio $D_n(x) \stackrel{def}{=} \sup_{M_n\in\mathcal{M}_{n,sym}} \mathbb{P}\{M_n \geqslant x\}$ sprendinį $D_n(x)$. Mes parodome, kad šis uždavinys yra ekvivalentus uždaviniui, kai norima rasti simetrinį atsitiktinį klaidžiojimą su aprėžtais žingsnių ilgiais, kuris maksimizuoja tikimybę patekti į intervalą $[x;\infty]$. Mes galime interpretuoti rezultatą, kai *galutinį* ir *optimalų* viršutinį rėžį $\mathbb{P}\{M_n \geqslant x\} \leqslant D_n(x)$, $x \in \mathbb{R}$, uodegos tikimybei $\mathbb{P}\,\{M_n \geqslant x\}$.