Stochastic Processes and their Applications 93 (2001) 109–117

stochastic processes and their applications

www.elsevier.com/locate/spa

On Bernstein-type inequalities for martingales

K. Dzhaparidze *, J.H. van Zanten

CWI, P.O. Box 94079, 1090 GB Amsterdam, Netherlands

Received 23 January 2000; received in revised form 12 October 2000; accepted 13 October 2000

Abstract

Bernstein-type inequalities for local martingales are derived. The results extend a number of well-known exponential inequalities and yield an asymptotic inequality for a sequence of asymptotically continuous martingales. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Locally square integrable martingale; Bernstein inequality; Multiplicative decomposition; Exponential supermartingale; Exponential inequality

1. Introduction

Exponential inequalities for tail probabilities of sums of independent random variables and their extension to martingales have numerous important applications in probability theory and statistics. See for instance, Shorack and Wellner (1986) and van de Geer (1995) for recent applications to statistical theory of empirical process. For another type of statistical applications, see e.g. Liptser and Spokoiny (2000). To mention a few other examples, we refer to the study by Coquet et al. (1994) and Courbot (1998) of the rate of convergence in the functional central limit theorem for martingales and to the theory of decoupling (see e.g. De la Peña and Giné, 1999; De la Peña, 1999).

The classical Bernstein inequality gives a tail bound for sums of independent random variables with a bounded range. If $\xi_1, \xi_2, \ldots$ are zero mean-independent random variables, all bounded in absolute value by some constant $a > 0$ so that $|\xi_i| \leqslant a$ for every i, then the partial sums $M_n = \xi_1 + \cdots + \xi_n$ obey the following inequality:

$$P(|M_n| > z) \leqslant 2\mathrm{e}^{-(1/2)(z^2/L+az/3)}$$

for all $z > 0$ and for all $L > 0$ satisfying $\mathrm{Var}(M_n) = E\xi_1^2 + \cdots + E\xi_n^2 \leqslant L$. See for instance Bennett (1962). Freedman (1975) extended Bernstein's result to the case of discrete-time martingales with bounded jumps. The random variables ξ_i are still assumed bounded by a constant a, but the mutual independence is now unnecessary.

* Corresponding author.

E-mail addresses: kacha@cwi.nl (K. Dzhaparidze), hvz@cwi.nl (J.H. van Zanten).

PII: S0304-4149(00)00086-7

Instead, $\xi_1, \xi_2, \ldots$ is a martingale difference sequence with respect to a certain filtration $\{\mathscr{F}_n\}_{n=0,1,\ldots}$, thus $E(\xi_i \mid \mathscr{F}_{i-1}) = 0$ for every i. The process M defined by $M_n = \xi_1 + \cdots + \xi_n$ is then a martingale with respect to this filtration. Furthermore, the unconditional variance $\mathrm{Var}(M_n)$ is replaced by its conditional counterpart

$$\langle M \rangle_n = \sum_{i=1}^n E(\xi_i^2 \mid \mathscr{F}_{i-1}).$$

Freedman's result then states that if τ is a finite stopping time with respect to the filtration $\{\mathscr{F}_n\}$, then

$$P\left(\max_{n\leqslant\tau} |M_n| > z,\ \langle M \rangle_\tau \leqslant L\right) \leqslant 2\mathrm{e}^{-(1/2)(z^2/L+az/3)}$$

for all $z, L > 0$.

The inequality of Freedman (1975) is further generalized in this paper. We abandon the assumption that the jumps of the martingale in question are bounded but require that they possess finite second moment. Applied to discrete-time martingales, our main result (Theorem 3.3 below) states that if for every $a > 0$ the second-order process H^a is defined by

$$H_n^a = \sum_{i=1}^n \xi_i^2 1_{\{|\xi_i|>a\}} + \sum_{i=1}^n E(\xi_i^2 \mid \mathscr{F}_{i-1}) = \sum_{i=1}^n \xi_i^2 1_{\{|\xi_i|>a\}} + \langle M \rangle_n, \tag{1.1}$$

then at each finite stopping time τ we have

$$P\left(\max_{n\leqslant\tau} |M_n| > z,\ H_\tau^a \leqslant L\right) \leqslant 2\mathrm{e}^{-(1/2)(z^2/L)\psi(az/L)}$$

for every $z, L > 0$. Here ψ is defined by

$$\psi(x) = \frac{2}{x^2}\int_0^x \log(1+y)\,\mathrm{d}y. \tag{1.2}$$

It satisfies

$$\psi(x) \geqslant \frac{1}{1+x/3} \quad \text{if } x \geqslant -1 \tag{1.3}$$

(cf. Shorack and Wellner, 1986, p. 441), therefore the new bound $2\exp(-\frac{1}{2}(z^2/L)\psi(az/L))$ is somewhat sharper then the earlier $2\exp(-\frac{1}{2}z^2/L + az/3)$. Besides, under the earlier condition that $|\xi_i| \leqslant a$ for every i, the first term on the right-hand side of (1.1) vanishes, i.e. $H^a = \langle M \rangle$. Thus, our inequality indeed implies Freedman's inequality, as well as its consequence, the classical Bernstein inequality.

We shall work within the framework of general martingale theory, see for instance Liptser and Shiryayev (1989) or Jacod and Shiryaev (1987). Thus, we consider càdlàg martingales in continuous time (right-continuous martingales with left-hand limits). This class of course includes the discrete-time martingales, but is much larger. Other important subclasses are the martingales with continuous sample paths and the compensated counting processes. The powerful tools of modern martingale theory allow for a unified treatment of all these particular cases. This leads to general results and clarifies the basic principles and ideas in estimating tail probabilities for martingales. Throughout the paper, we will use the standard notations from martingale theory, M for a local martingale,

$\langle M\rangle$ and $[M]$ for its predictable and optional quadratic variations, respectively, ΔM for its jump process and M^* for the process given by $M_t^* = \sup_{s\leqslant t}|M_s|$.

Shorack and Wellner (1986) extended Freedman's result to the setting of general local martingales with bounded jumps. They proved that if $|\Delta M_t|\leqslant a$ for every $t\geqslant 0$, then

$$P(M_\tau^* \geqslant z, \langle M\rangle_\tau \leqslant L) \leqslant 2\mathrm{e}^{-(1/2)(z^2/L)\psi(az/L)} \tag{1.4}$$

for every finite stopping time τ and z, $L>0$. Here ψ is again given by (1.2). The third assertion of Corollary 3.4 of our main Theorem 3.3 shows how easily inequality (1.4) is derived from our inequality (3.7). We also show the relationship with other known results. Assertions (i) and (ii) of Corollary 3.4 provide links to results due to Barlow et al. (1986), Proposition 4.2.1, and Courbot (1998), Theorem 5.2. Assertion (i) tells us that if in (1.4) the predictable quadratic characteristic $\langle M\rangle$ is substituted by the optional second-order process $[M]+\langle M^{\mathrm{d}}\rangle$ (where M^{d} is the purely discontinuous part of M), then so-called sub-Gaussian tail bounds for locally square integrable martingales can be obtained without any additional assumptions on the size of the jumps, exactly as in Barlow et al. (1986), Proposition 4.2.1. In assertion (ii) of Corollary 3.4 the so-called Fuk–Nagaev inequality is derived, cf. Courbot (1998) where the relationship is discussed to the original paper by Fuk and Nagaev (1971), as well as to papers by Haeusler (1984) and Kubilius and Mémin (1994).

At the end of Section 3, we provide another application of Theorem 3.3, concerning a sequence of locally square integrable martingales subject to a certain asymptotic continuity requirement of Lindeberg type. Note that in the case of continuous local martingales, i.e. when $\Delta M=0$, inequality (1.4) gives the sub-Gaussian bound

$$P(M_\tau^* \geqslant z, \langle M\rangle_\tau \leqslant L) \leqslant 2\mathrm{e}^{-(1/2)z^2/L} \tag{1.5}$$

for every finite stopping time τ and z, $L>0$. It is shown that for asymptotically continuous martingales, this inequality is satisfied in the limit (see Corollary 3.5). The last section of the paper is devoted to conditionally symmetric martingales. Theorem 4.1 extends a result by de la Peña (1999) concerning martingales in discrete time.

2. Exponential supermartingales

Throughout, we will use the standard notions of the 'general theory of stochastic processes'. For more details we refer to Liptser and Shiryayev (1989) or Jacod and Shiryaev (1987). The processes under consideration will always be defined on a certain fixed stochastic basis. We assume, for simplicity, that the martingale M in question starts from zero, i.e. $M_0=0$. Then the canonical representation of M reads as $M = M^{\mathrm{c}} + x*(\mu-\nu)$, where M^{c} is the continuous part of M, μ is its jump measure and ν is the predictable compensator of μ. When the latter characteristic ν is so that at each time $t\geqslant 0$

$$\varphi * \nu_t < \infty \quad \text{a.s.}, \quad \varphi(x) = \mathrm{e}^x - 1 - x, \tag{2.1}$$

then we may associate with M, the so-called *cumulant* process

$$G = \tfrac{1}{2}\langle M^{\mathrm{c}}\rangle + \varphi * \nu \tag{2.2}$$

(here and elsewhere below we agree upon the standard notation $\varphi * v_t$ for the pathwise stochastic integral $\int_0^t \int \varphi(x) v(\mathrm{d}s \times \mathrm{d}x)$). Since the exponential function is strictly positive and since $\int v(\{t\} \times \mathrm{d}x) \leqslant 1$, it holds that $\Delta G_t = \int (\exp(x) - 1)\, v(\{t\} \times \mathrm{d}x) > -1$ a.s. for every $t \geqslant 0$. Therefore, the Doléans–Dade exponential $\mathscr{E}(G)$ of G is well defined and satisfies

$$\mathscr{E}_t(G) = \mathrm{e}^{G_t} \prod_{s \leqslant t} (1 + \Delta G_s) \mathrm{e}^{-\Delta G_s}. \tag{2.3}$$

Recall that every positive, special semimartingale can be uniquely written as the product of a local martingale and a predictable, bounded variation process (see Liptser and Shiryayev, 1989, p. 127). Under condition (2.1) the positive semimartingale $\exp(M)$ is special. Its multiplicative decomposition is given by the following well-known theorem (see Liptser and Shiryayev, 1989, Section 4.13, Problem 2).

Theorem 2.1. *Let M be a local martingale whose characteristic v satisfies condition* (2.1). *Then there exists a nonnegative local martingale N such that*

$$\mathrm{e}^M = N \mathscr{E}(G). \tag{2.4}$$

We shall use the following simple corollary of this theorem.

Corollary 2.2. *Let M be a local martingale whose characteristic v satisfies condition* (2.1). *Then the process* $\exp(M - G)$ *is a supermartingale.*

Proof. It follows from Theorem 2.1 that $\exp(M - G) = N\mathscr{E}(G)\exp(-G)$, where N is a nonnegative local martingale. A nonnegative local martingale is also a supermartingale. Therefore, it remains to verify that the process $\mathscr{E}(G)\exp(-G)$ is decreasing. From (2.3) we see that

$$\mathscr{E}(G)_t \mathrm{e}^{-G_t} = \prod_{s \leqslant t} (1 + \Delta G_s) \mathrm{e}^{-\Delta G_s} = \mathrm{e}^{\sum_{s \leqslant t} (\log(1 + \Delta G_s) - \Delta G_s)}.$$

Now, recall that $\Delta G > -1$ and that $\log(1 + x) - x \leqslant 0$ for $x > -1$. Therefore, each term in the sum on the right-hand side is nonpositive and the process is indeed decreasing. The proof is complete. □

It is said that a local martingale M satisfies *Cramér's condition* for $\lambda \in \mathbb{R}$, if at each time $t \geqslant 0$ we have $\varphi(\lambda x) * v_t < \infty$ a.s. where φ is again given by (2.1). Under this condition the cumulant process $G(\lambda)$ of the local martingale λM is well defined and satisfies

$$G(\lambda) = \tfrac{1}{2} \lambda^2 \langle M^c \rangle + \varphi(\lambda x) * v. \tag{2.5}$$

By Corollary 2.2, the process $\exp(\lambda M - G(\lambda))$ is a supermartingale if M satisfies Cramér's condition for $\lambda \in \mathbb{R}$. The following theorem may be regarded as an extension of this fact (to see this, put $f \equiv 0$).

Theorem 2.3. *Let M be a local martingale whose characteristic v is such that for a certain measurable function f with $f(0) = 0$ and a fixed $\lambda \in \mathbb{R}$*

$$|f(\lambda x)| * v < \infty, \quad \varphi(\lambda x - f(\lambda x)) * v < \infty \quad \textit{a.s.} \tag{2.6}$$

with φ *as in* (2.1). *Define the process* $S(\lambda)=M(\lambda)+A(\lambda)$, *where* $M(\lambda)=f(\lambda x)*(\mu-\nu)$ *and* $A(\lambda)=\frac{1}{2}\lambda^2\langle M^c\rangle+\varphi(\lambda x-f(\lambda x))*\nu$. *Then* $\exp(\lambda M-S(\lambda))$ *is a supermartingale.*

Proof. We apply Corollary 2.2 to the local martingale $M'(\lambda)=\lambda M-M(\lambda)$. Since $M(\lambda)$ is purely discontinuous we have $\langle M'(\lambda)^c\rangle=\langle(\lambda M)^c\rangle=\lambda^2\langle M^c\rangle$. For the predictable compensator $\nu'(\lambda)$ of the jump measure of $M'(\lambda)$ it holds that $\varphi*\nu'(\lambda)=\varphi(\lambda x-f(\lambda x))*\nu$. The cumulant process $G'(\lambda)$ of $M'(\lambda)$ is therefore given by

$$\begin{aligned}G'(\lambda)&=\tfrac{1}{2}\langle M'(\lambda)^c\rangle+\varphi*\nu'(\lambda)\\&=\tfrac{1}{2}\lambda^2\langle M^c\rangle+\varphi(\lambda x-f(\lambda x))*\nu=A(\lambda).\end{aligned}$$

By Corollary 2.2 we thus have that $\exp(\lambda M-S(\lambda))=\exp(M'(\lambda)-G'(\lambda))$ is a supermartingale. □

3. Locally square integrable martingales

In this section, we will deal exclusively with locally square integrable martingales. It will be shown that in this case conditions (2.6) are satisfied if we take

$$f(x)=\tfrac{1}{2}x^2 1_{\{|x|>\lambda a\}}, \tag{3.1}$$

where $a,\lambda\geqslant 0$. This will allow us to obtain Corollary 3.1 of Theorem 2.3. The corollary will then provide the principal argument for proving our main Theorem 3.3. For simplicity, let us use the following notation: for all $a,\lambda\geqslant 0$

$$\varphi_a(\lambda)=\frac{\varphi(a\lambda)}{a^2}, \tag{3.2}$$

where φ is as in (2.1).

Corollary 3.1. *Let* M *be a locally square integrable martingale and for* $a\geqslant 0$, *define the process* H^a *by*

$$H_t^a=\sum_{s\leqslant t}(\Delta M_s)^2 1_{\{|\Delta M_s|>a\}}+\langle M\rangle_t. \tag{3.3}$$

Then for all $a,\lambda\geqslant 0$ *the process*

$$e^{\lambda M-\varphi_a(\lambda)H^a} \tag{3.4}$$

is a supermartingale, where $\varphi_a(\lambda)$ *is given by* (3.2).

Proof. Let us first derive some useful properties of the functions φ_a defined by (3.2). Note that $\varphi_0(\lambda)=\lambda^2/2$ and that for each fixed $\lambda\geqslant 0$ the function $\varphi_a(\lambda)$ is increasing in a (look at power series expansions). In particular, $\varphi_0(\lambda)\leqslant\varphi_a(\lambda)$ for $a\geqslant 0$. Moreover, if $|x|\leqslant a$, then

$$\varphi(\lambda x)\leqslant\varphi(\lambda|x|)=\varphi_{|x|}(\lambda)x^2\leqslant\varphi_a(\lambda)x^2. \tag{3.5}$$

Besides, for any $x\in\mathbb{R}$

$$|\varphi(\lambda x-\tfrac{1}{2}\lambda^2x^2)-\tfrac{1}{2}\lambda^2x^2|\leqslant\varphi_0(\lambda)x^2\leqslant\varphi_a(\lambda)x^2. \tag{3.6}$$

To see this check that $|\exp(x-\frac{1}{2}x^2)-1-x|\leqslant\frac{1}{2}x^2$.

With these inequalities at hand, we can now verify conditions (2.6) with f of form (3.1), where $a, \lambda \geqslant 0$ are now fixed. The first inequality is straightforward: $|f(\lambda x)| * v = \frac{1}{2}\lambda^2 x^2 1_{\{|x|>a\}} * v < \infty$ a.s., since M is locally square integrable. So the question really concerns the second one. But in the present case

$$\varphi(\lambda x - f(\lambda x)) = \varphi(\lambda x - \tfrac{1}{2}\lambda^2 x^2) 1_{\{|x|>a\}} + \varphi(\lambda x) 1_{\{|x|\leqslant a\}},$$

hence the second condition of (2.6) is verified by using (3.5), (3.6) and the fact that M is locally square integrable.

Since the conditions of Theorem 2.3 are satisfied with the special choice (3.1) for the function f, we may conclude that the process $\exp(\lambda M - S(\lambda))$ is a supermartingale, where

$$S(\lambda) = f(\lambda x) * (\mu - v) + \tfrac{1}{2}\lambda^2 \langle M^c \rangle + \varphi(\lambda x - f(\lambda x)) * v.$$

Using (3.5) and (3.6) again it is easily verified that the difference $S(\lambda) - \varphi_a(\lambda) H^a$ is a decreasing process. Being the product of a supermartingale and a decreasing process, $\exp(\lambda M - \varphi_a(\lambda) H^a)$ is therefore a supermartingale as well. □

Remark 3.2. The particular case $a = 0$ is of special interest. In this case (3.3) reduces to $H^0 = [M] + \langle M^d \rangle$ where $M^d = M - M^c$ is the purely discontinuous part of M. This is obvious, since by definition $[M] = \sum_{s \leqslant \cdot} (\Delta M_s)^2 + \langle M^c \rangle$.

We now proceed to our main result.

Theorem 3.3. *Let M be a locally square integrable martingale and let H^a be given by* (3.3) *for $a \geqslant 0$. Then for every finite stopping time τ, $a \geqslant 0$ and z, $L > 0$*

$$P(M^*_\tau \geqslant z,\ H^a_\tau \leqslant L) \leqslant 2e^{-(1/2)(z^2/L)\psi(az/L)}, \tag{3.7}$$

where ψ is given by (1.2).

Proof. Since H^a is nondecreasing we have for each $\lambda \geqslant 0$ and z, $L \geqslant 0$ that

$$P\left(\sup_{t\leqslant\tau} M_t \geqslant z, H^a_\tau \leqslant L\right) \leqslant P\left(\sup_{t\leqslant\tau} Z_t(\lambda) \geqslant e^{\lambda z - \varphi_a(\lambda)L}\right), \tag{3.8}$$

where $Z(\lambda) = \exp(\lambda M - \varphi_a(\lambda) H^a)$ is the positive supermartingale of (3.4). Since $Z_0(\lambda) = 1$, the optional sampling theorem yields for any stopping time σ the inequality

$$EZ_\sigma(\lambda) 1_{\{\sigma<\infty\}} \leqslant 1. \tag{3.9}$$

By applying this to a particular stopping time, namely to $\sigma = \inf\{t : Z_t(\lambda) \geqslant \exp(\lambda z - \varphi_a(\lambda)L)\}$, we may extend (3.8) as follows:

$$P\left(\sup_{t\leqslant\tau} Z_t(\lambda) \geqslant e^{\lambda z - \varphi_a(\lambda)L}\right) \leqslant P(\sigma \leqslant \tau) \leqslant P(\sigma < \infty) \leqslant e^{\varphi_a(\lambda)L - \lambda z}.$$

We have first taken into consideration that τ is finite and then applied inequality (3.9). So (3.8) turns into

$$P\left(\sup_{t\leqslant\tau} M_t \geqslant z, H^a_\tau \leqslant L\right) \leqslant e^{\varphi_a(\lambda)L - \lambda z}$$

for each $\lambda \geqslant 0$. Clearly, the same inequality holds with M substituted by $-M$, for the process H^a remains unaltered. Thus,

$$P(M_\tau^* \geqslant z,\ H_\tau^a \leqslant L) \leqslant 2\mathrm{e}^{\varphi_a(\lambda)L-\lambda z} \tag{3.10}$$

for each $\lambda \geqslant 0$. As is well known (see e.g. Courbot, 1998, p. 39), the special choice $\lambda = \log(1 + az/L)/a$ for $a \geqslant 0$ (with $\lambda = z/L$ when $a = 0$) renders the right-hand side as small as possible and yields the desired inequality (3.7). □

The following corollary shows how Theorem 3.3 does provide a number of known special cases mentioned already in the introduction. The inequality in assertion (i) is due to Barlow et al. (1986). The inequality in assertion (ii) is referred to as the *Fuk–Nagaev* inequality, see Courbot (1998), Section II.5. Finally, assertion (iii) concerns the well-known inequality (1.4) in the case of bounded jumps.

Corollary 3.4. *Let M be a locally square integrable martingale and let τ be a finite stopping time.*

(i) *For all $z,\ L > 0$*

$$P(M_\tau^* \geqslant z, [M]_\tau + \langle M^{\mathrm{d}} \rangle_\tau \leqslant L) \leqslant 2\mathrm{e}^{-(1/2)z^2/L},$$

where $M^{\mathrm{d}} = M - M^{\mathrm{c}}$ is the purely discontinuous part of M.

(ii) *For all $a \geqslant 0$ and $z,\ L > 0$*

$$P(M_\tau^* \geqslant z) \leqslant 2\mathrm{e}^{-(1/2)(z^2/L)\psi(az/L)} + P(\langle M \rangle_\tau > L) + P(|\Delta M|_\tau^* > a),$$

where ψ is given by (1.2).

(iii) *If $|\Delta M| \leqslant a$, then for all $z,\ L > 0$ we have* (1.4).

Proof. In view of Remark 3.2, assertion (i) follows directly from Theorem 3.3 upon the substitution $a = 0$. To prove (ii) and (iii), write $H^a = I^a + \langle M \rangle$, where

$$I_t^a = \sum_{s \leqslant t} (\Delta M_s)^2 1_{\{|\Delta M_s| > a\}}. \tag{3.11}$$

Clearly it holds that

$$P(M_\tau^* \geqslant z,\ \langle M \rangle_\tau \leqslant L) \leqslant P(M_\tau^* \geqslant z,\ H_\tau^a \leqslant L) + P(I_\tau^a > 0)$$

for all $a \geqslant 0$. By Theorem 3.3, the first probability on the right-hand side is bounded by $2\exp(-(1/2)(z^2/L)\psi(az/L))$. For the second probability, observe that at any instant $t \geqslant 0$ the sum in (3.11) is positive if and only if at least one jump took place that exceeded a in absolute value, so

$$P(I_\tau^a > 0) = P(|\Delta M|_\tau^* > a).$$

We get

$$P(M_\tau^* \geqslant z,\ \langle M \rangle_\tau \leqslant L) \leqslant 2\mathrm{e}^{-(1/2)(z^2/L)\psi(az/L)} + P(|\Delta M|_\tau^* > a)$$

for all $a \geqslant 0$. This yields inequalities (ii) and (iii). □

The last result in this section, inequality (3.13), may be regarded as an asymptotic Bernstein inequality. It resembles Bernstein's inequality (1.5) for continuous local martingales. In particular, the tail bound is sub-Gaussian. The result requires the asymptotic continuity of a sequence of locally square integrable martingales, expressed as usual in the form of Lindeberg condition (3.12), cf. e.g. Jacod and Shiryaev (1987) or Liptser and Shiryayev (1989). Thus, we are going to deal with a sequence of locally square integrable martingales $\{M^n\}_{n=1,2,\ldots}$, each martingale defined on its own stochastic basis. Therefore, all its characteristics and associated processes will be indexed by n as well. Otherwise, the previous notational conventions will be retained.

Corollary 3.5. *Let $\{M^n\}_{n=1,2,\ldots}$ be a sequence of locally square integrable martingales. Suppose the corresponding sequence of predictable characteristics $\{v^n\}_{n=1,2,\ldots}$ and a sequence of stopping times $\{\tau_n\}_{n=1,2,\ldots}$ are so that*

$$x^2 1_{\{|x|>\varepsilon\}} * v^n_{\tau_n} \xrightarrow{P^n} 0 \tag{3.12}$$

for all $\varepsilon > 0$. Then

$$\limsup_{n\to\infty} P^n(M^{n*}_{\tau_n} \geqslant z, \langle M^n\rangle_{\tau_n} \leqslant L) \leqslant 2\mathrm{e}^{-(1/2)z^2/L} \tag{3.13}$$

for all z, $L > 0$.

Proof. Obviously, we can bound the probability in (3.13) from above by the sum $P^n(M^{n*}_{\tau_n} \geqslant z, H^{n\varepsilon}_{\tau_n} \leqslant K + L) + P^n(I^{n\varepsilon}_{\tau_n} > K)$ for every $\varepsilon, K > 0$ (the processes H^{na} and I^{na} are again given by (3.3) and (3.11)). By Theorem 3.3, the first term does not exceed $2\exp\{-z^2/2((\varepsilon z/3) + L + K)\}$ (to see this substitute a and L in (3.7) by ε and $K + L$, respectively, and then take into consideration inequality (1.3)), while the second term vanishes as $n \to \infty$ in virtue of (3.12) and the Lenglart inequality (see e.g. Jacod and Shiryaev, 1987, Lemma I.3.30). Hence,

$$\limsup_{n\to\infty} P^n(M^{n*}_{\tau_n} \geqslant z, \langle M^n\rangle_{\tau_n} \leqslant L) \leqslant 2\mathrm{e}^{-(1/2)z^2/(\varepsilon z/3)+L+K}$$

for every $\varepsilon, K > 0$. The proof is completed by letting $\varepsilon \downarrow 0$ and $K \downarrow 0$. □

4. Conditionally symmetric martingales

It is said that a local martingale is *conditionally symmetric* if the predictable characteristic of its jumps v is such that for all integrable functions f

$$f(x) * v = f(-x) * v.$$

This notion occurs usually in the discrete time setup, in particular, within the theory of decoupling, cf. e.g. de la Peña and Giné (1999) or de la Peña (1999). In discrete time, when $M_n = \xi_1 + \cdots + \xi_n$ and ξ_n is a martingale difference sequence, conditional symmetry of a martingale simply means that for every n, the conditional distribution of ξ_n given $\mathscr{F}_{n-1}$ is symmetric. Our next result is an extension of a discrete time result of de la Peña (1999), Section 6. Note that under the present symmetry condition, the expression (2.5) for the cumulant $G(\lambda)$ reduces to $G(\lambda) = \frac{1}{2}\lambda^2\langle M^c\rangle + (\cosh(\lambda x) - 1) * v$. Assertion (i) of Corollary 3.4 simplifies to the following result.

Theorem 4.1. *Let M be a locally square integrable and conditionally symmetric martingale. Then at each finite stopping time τ*

$$P(M_\tau^* \geqslant z,\ [M]_\tau \leqslant L) \leqslant 2\mathrm{e}^{-(1/2)z^2/L} \tag{4.1}$$

for all $z, L \geqslant 0$.

Proof. Apply Theorem 2.3 with $f(x) = \frac{1}{2}x^2$. Taking into consideration the conditional symmetry, we get

$$S(\lambda) - \tfrac{1}{2}\lambda^2[M] = (\mathrm{e}^{-(1/2)\lambda^2 x^2}\cosh(\lambda x) - 1) * v$$

which is decreasing since $\cosh x \leqslant \exp(x^2/2)$ for all $x \in \mathbb{R}$. This means, in particular, that $\exp(\lambda M - \frac{1}{2}\lambda^2[M])$ is a supermartingale. Apply now the arguments like in the course of proving Theorem 3.3 but with the latter supermartingale in the place of $Z(\lambda)$. Depart namely from the inequality

$$P\left(\sup_{t\leqslant\tau} M_t \geqslant z,\ [M]_\tau \leqslant L\right) \leqslant P\left(\sup_{t\leqslant\tau} \mathrm{e}^{\lambda M_t - (1/2)\lambda^2[M]_t} \geqslant \mathrm{e}^{\lambda z - (1/2)\lambda^2 L}\right)$$

and obtain

$$P(M_\tau^* \geqslant z,\ [M]_\tau \leqslant L) \leqslant 2\mathrm{e}^{(1/2)\lambda^2 L - \lambda z},$$

cf. (3.8) and (3.10), respectively. To complete the proof, select λ as to render the right-hand side as small as possible. $\square$

References

Barlow, M.T., Jacka, S.D., Yor, M., 1986. Inequalities for a pair of processes stopped at a random time. Proc. London Math. Soc. 3 (52), 142–172.

Bennett, G., 1962. Probability inequalities for the sum of independent random variables. JASA 57, 33–45.

Coquet, F., Mémin, J., Vostrikova, L., 1994. Rate of convergence in the functional central limit theorem for likelihood ratio processes. Math. Methods Statist. 3 (2), 89–113.

Courbot, B., 1998. Vitesses de convergence pour les martingales dans le théorème central limite fonctionnel. Ph.D. Thesis, Université de Rennes 1.

de la Peña, V.H., 1999. A general class of exponential inequalities for martingales and ratios. Ann. Probab. 27, 537–564.

de la Peña, V.H., Giné, E., 1999. Decoupling. Springer, Berlin.

Freedman, D.A., 1975. On tail probabilities for martingales. Ann. Probab. 3, 100–118.

Fuk, D.Kh., Nagaev, S.V., 1971. Probability inequalities for sums of independent random variables. Theory Probab. Appl. 16, 643–660.

Haeusler, E., 1984. An exact rate of convergence in the functional central limit theorem for special martingale difference arrays. Z. Wahr. Verw. Geb. 65, 523–534.

Jacod, J., Shiryaev, A.N., 1987. Limit Theorems for Stochastic Processes. Springer, Berlin.

Kubilius, K., Mémin, J., 1994. Inégalité exponentielle pour les martingales locales. C.R. Acad. Sci. Paris Série I t. 319, 733–738.

Liptser, R.Sh., Shiryayev, A.N., 1989. Theory of Martingales. Kluwer Academic Publishers, Dordrecht.

Liptser, R.Sh., Spokoiny, V., 2000. Deviation probability bound for martingales with applications to statistical estimation. Statist. Probab. Lett. 46, 347–357.

Shorack, G.R., Wellner, J.A., 1986. Empirical Processes with Applications to Statistics. Wiley, New York.

van de Geer, S., 1995. Exponential inequalities for martingales, with applications to maximum likelihood estimation for counting processes. Ann. Statist. 23, 1779–1801.