About Doob's inequality, entropy and Tchebichef

Emmanuel Rio

About Doob's inequality, entropy and Tchebichef

Emmanuel Rio[⋆]

[⋆] Université de Versailles, Laboratoire de mathématiques de Versailles, UMR 8100 CNRS, Bâtiment Fermat, 45 Avenue des Etats-Unis, 78035 Versailles, FRANCE. E-mail: emmanuel.rio@uvsq.fr

Key words and phrases. Doob's inequality; Hardy-Littlewood maximal function; $L \log L$; entropy; binomial rate function; covariance inequalities; Cantelli's inequality; von Bahr-Esseen inequality; sub-Gaussian random variables; bounded differences inequality; McDiarmid's inequality; conditional value at risk.

Mathematical Subject Classification (2010). Primary: 60E15. Secundary: 60G42.

Abstract

In this paper we give upper bounds on the tail or the quantiles of the one-sided maximum of a nonnegative submartingale in the class $L \log L$ or the maximum of a submartingale in L^p. Our upper bounds involve the entropy in the case of nonnegative martingales in the class $L \log L$ and the L^p-norm in the case of submartingales in L^p. Starting from our results on entropy, we also improve the so-called bounded differences inequality. All the results are based on optimal bounds for the conditional value at risk of real-valued random variables.

1 Introduction

This paper is motivated by the question below. Let $(M_k)_{0 \le k \le n}$ be a real-valued submartingale in L^1. Define $M_n^* = \max(M_0, M_1, \dots, M_n)$. How to provide an upper bound on the tail or the quantiles of M_n^* under some additional integrability conditions on the submartingale?

In order to explain our results, we need the definition of the quantile function of a random variable X and some basic properties of this function.

Definition 1.1. *Let X be a real-valued random variable. The tail function H_X is defined by $H_X(t) = \mathbb{P}(X > t)$. The quantile function Q_X is the cadlag inverse of H_X.*

The basic property of Q_X is: $x < Q_X(u)$ if and only if $H_X(x) > u$. This property ensures that $Q_X(U)$ has the same distribution as X for any random variable U with the uniform distribution over $[0, 1]$.

Definition 1.2. *The median $\mu(X)$ of a real-valued random variable X is defined by $\mu(X) = Q_X(1/2)$.*

Let us now recall Doob's maximal inequalities. Below we assume that the random variables $M_0, M_1, \dots, M_n$ are nonnegative. The first inequality is in fact due to Ville (1939, Theorem 1, page 100): for any $x > 0$,

$$\mathbb{P}(M_n^* \ge x) \le x^{-1}\mathbb{E}(M_n) \ \text{ or, equivalently, } Q_{M_n^*}(1/z) \le z\mathbb{E}(M_n) \qquad \text{Ville (1939)}$$

for any $z \geq 1$. From the fact that $Q_{M_n}(U)$ has the same law as M_n, Ville's inequality is equivalent to

$$Q_{M_n^*}(u) \leq u^{-1} \int_0^1 Q_{M_n}(s)ds \quad \text{for any } u \in]0,1]. \tag{1.1}$$

Clearly one cannot derive upper bounds on $\mathbb{E}(M_n^*)$ from this inequality. Doob (1940, Theorem 1.1 or 1953, p. 314) proved the more precise inequality

$$\mathbb{P}(M_n^* \geq x) \leq x^{-1}\mathbb{E}\big(M_n \mathbf{1}_{M_n^* \geq x}\big) \quad \text{for any } x > 0. \qquad \text{Doob (1940)}$$

Assume now that the random variable M_n is in the class $L \log L$ of real-valued random variables X such that $\mathbb{E}(|X| \log^+ |X|) < \infty$. Applying Ville's inequality to the submartingale $(M_k \log^+ M_k)_{0 \leq k \leq n}$, one immediately gets that, for any $x > 1$,

$$\mathbb{P}(M_n^* \geq x) \leq (x \log x)^{-1}\mathbb{E}(M_n \log^+ M_n).$$

This inequality proves that the tail of M_n^* is at most of the order of $(x \log x)^{-1}$ as $x \nearrow \infty$. Nevertheless, first the upper bound tends to ∞ as $x \searrow 1$, even under the normalization condition $\mathbb{E}(M_n) = 1$ and, second, the quantities involved here fail to be homogeneous. Therefore, it seems clear that the above upper bound can be improved.

We now recall the known results for nonnegative submartingales in the class $L \log L$. Up to now, upper bounds on $\mathbb{E}(M_n^*)$ have paid more attention than upper bounds on the tail of M_n^*. Starting from his inequality, Doob (1953, p. 317) obtained the upper bound

$$\mathbb{E}(M_n^*) \leq \tfrac{e}{e-1}\big(\mathbb{E}(M_n \log^+ M_n) + 1\big), \quad \text{where } \log^+ x = \max(0, \log x). \tag{1.2}$$

Define now the integrated quantile function $\tilde{Q}_X$ of a real-valued integrable random variable X by

$$\tilde{Q}_X(u) = u^{-1} \int_0^u Q_X(s)ds \quad \text{for any } u \in]0,1]. \tag{1.3}$$

In mathematical finance, $\tilde{Q}_X$ is called conditional value at risk of X. Clearly $Q_X \leq \tilde{Q}_X$. Blackwell and Dubins (1963) derived from the Doob inequality the upper bound

$$Q_{M_n^*}(u) \leq \tilde{Q}_{M_n}(u) \quad \text{for any } u \in]0,1]. \tag{1.4}$$

Later Dubins and Gilat (1978) proved the optimality of (1.4).

For a nonnegative random variable X, $\tilde{Q}_X$ is known as the Hardy-Littlewood maximal function assoicated with X. Hardy and Littlewood (1930, Theorem 11) proved that

$$\int_0^1 \tilde{Q}_X(u)du \leq c\big(\mathbb{E}(X \log^+ X) + 1\big)$$

for some universal positive constant c, which gives an alternative proof of (1.3), up to the constant. The above inequality is usually called $L \log L$ inequality of Hardy and Littlewood. Gilat (1986, Theorem 3) proved that the two-parameter inequality

$$\int_0^1 \tilde{Q}_X(u)du \leq c\mathbb{E}(X \log X) + d \tag{1.5}$$

holds for any $c > 1$ and any $d \geq e^{-1}c^2(c-1)^{-1}$. In particular, if $c = e/(e-1)$ then (1.5) holds true with $d = e/(e-1)$. Using (1.4), it follows that (1.2) holds true with

$\mathbb{E}(M_n \log M_n)$ instead of $\mathbb{E}(M_n \log^+ M_n)$. The martingale counterpart of (1.5) may be found in Osekowski (2012, Theorem 7.7). Curiously (1.2) and (1.5) fail to be homogeneous, since they are not invariant under the multiplication of the submartingale or the random variable X by a constant factor, so that one can have some doubts about their optimality.

Starting from Doob's inequality and introducing the entropy of M_n, Harremoës (2008) improved Gilat's result. For a nonnegative real-valued random variable X such that $\mathbb{E}(X > 0)$ and $\mathbb{E}(X \log^+ X) < \infty$, define the entropy $\mathcal{H}(X)$ of X by

$$\mathcal{H}(X) = \mathbb{E}\big(X \log X\big) - \mathbb{E}(X) \log \mathbb{E}(X). \tag{1.6}$$

Under the above conditions $\mathcal{H}(X)$ is finite. Furthermore $\mathcal{H}(X) \geq 0$ and $\mathcal{H}(X) = 0$ if and only if X is almost surely constant. Assuming that $M_0 = 1$, as in Ville (1939), Harremoës (2008) derived from Doob's inequality the upper bounds

$$\mathbb{E}(M_n^*) - 1 \leq \mathbb{E}(M_n \log M_n^*) \leq \mathcal{H}(M_n) + \log \mathbb{E}(M_n^*). \tag{1.7}$$

Defining the function $g : [1, \infty[\mapsto [0, \infty[$ by $g(x) = x - 1 - \log x$, (1.7) implies that

$$\mathbb{E}(M_n^*) \leq g^{-1}\big(\mathcal{H}(M_n)\big), \tag{1.8}$$

where $g^{-1} : [0, \infty[\mapsto [1, \infty[$ is the inverse function of g. Harremoës (2008, Theorem 4) also proved that (1.8) is tight. It appears here that the entropy is the adequate quantity for nonnegative submartingales in the class $L \log L$. In the present paper we will obtain estimates for the tails or the quantiles of M_n^* involving entropy. In order to get these estimates, we give a covariance inequality in Section 2. Next, in Section 3 we derive upper bounds on the tail function of M_n^* from (1.4) and this covariance inequality. We also prove that our main inequality is sharp for positive martingales with given entropy and expectation.

Assume now that the random variable M_n fulfills the stronger moment condition $\mathbb{E}|M_n|^p < \infty$ for some $p > 1$. For any real y, let $y_+ = \max(0, y)$. By the Ville inequality applied to the nonnegative submartingale $(M_k - a)_+^p$,

$$\mathbb{E}(M_n^* \geq a + x) \leq x^{-p} \mathbb{E}\big((M_n - a)_+^p\big) \text{ for any } x > 0 \text{ and any real } a. \tag{1.9}$$

Setting $a = \mathbb{E}(M_n)$ in the above inequality, we obtain a deviation inequality for M_n^* around $\mathbb{E}(M_n)$. This inequality proves that the tail of M_n^* is at most of the order of x^{-p} as $x \nearrow \infty$. However, the upper bound tends to ∞ as $x \searrow 0$. Recall now the Tchebichef-Cantelli inequality (*see* Tchebichef (1874)[1] and Cantelli (1932), Inequality (19), p. 53): for any real-valued random variable X in L^2 and any positive x,

$$\mathbb{P}(X \geq \mathbb{E}(X) + x) \leq \sigma^2/(x^2 + \sigma^2), \quad \text{where } \sigma^2 = \operatorname{Var} X. \qquad \text{Tchebichef (1874)}$$

We refer to Savage (1961) for a review of probabilities inequalities of the Tchebichef type with a complete bibliography. This inequality is equivalent to the upper bound

$$Q_X(1/z) \leq \mathbb{E}(X) + \sigma\sqrt{z - 1} \text{ for any } z > 1. \tag{1.10}$$

[1]The left hand side version of this inequality follows immediately from the result of Tchebichef (1874) stated on the last line of p.159 and the first line of p.160, by taking the limit as l tends to ∞.

For instance (1.10) ensures that $\mu(X) \leq \mathbb{E}(X) + \sigma$. In Section 4, we give a maximal version of (1.10) for submartingales in L^p. In the special case $p = 2$ our result yields

$$Q_{M_n^*}(1/z) \leq \mathbb{E}(M_n) + \|M_n - \mathbb{E}(M_n)\|_2 \sqrt{z-1} \quad \text{for any } z > 1, \tag{1.11}$$

which is an extension of the above bound to maxima of submartingales. We then apply our results to martingales in L^p for p in $]1, 2]$ and we compare the so obtained upper bounds on $Q_{M_n^*}$ with the upper bounds that can be derived from the minimization of (1.9) with respect to a. These upper bounds are based on von Bahr-Esseen type inequalities. In particular, in order to make a fair comparison of the results that can be derived from (1.9) with the extension of Tchebichef-Cantelli's inequality to martingales in L^p, we prove a one-sided von Bahr-Esseen type inequality in the Annex.

In Section 5, we consider maxima of martingales $(M_k)_{0 \leq k \leq n}$ in L^p for $p > 2$. The two-sided maximum $|M|_n^*$ is defined by

$$|M|_n^* = \max(|M_0|, |M_1|, \ldots, |M_n|) \tag{1.12}$$

Let

$$\sigma = \|M_n\|_2 \quad \text{and } L_p = \sigma^{-p}\, \mathbb{E}|M_n|^p. \tag{1.13}$$

By Theorem 7.4 in Osekowski (2012),

$$\mathbb{P}(|M|_n^* \geq \sigma x) \leq (L_p/x)^p \quad \text{or, equivalently,} \quad Q_{|M|_n^*}(1/z) \leq \sigma (z L_p)^{1/p}. \tag{1.14}$$

Our aim in Section 5 is to give more precise bounds when $\mathbb{E}(M_0) = 0$. Assume for instance that $p = 4$. Let X be a centered real-valued random variable in L^4. Set $\sigma = \|X\|_2$. Cantelli (1932, p. 56) obtained the inequality below: for any $x > 1$,

$$\mathbb{P}(X \geq \sigma x) \leq \frac{L_4 - 1}{L_4 - 1 + (x^2 - 1)(\max(x^2, L_4) - 1)}, \qquad \text{Cantelli (1932)}$$

where $L_4 = \sigma^{-4}\mathbb{E}(X^4)$. Cantelli's inequality is equivalent to the quantile inequality

$$Q_X(1/z) \leq \sigma \big(1 + \sqrt{(\min(z, L_4) - 1)(z-1)}\ \big)^{1/2} \quad \text{for any } z > 1. \tag{1.15}$$

As shown by Cantelli (1932), this inequality improves the basic inequality

$$Q_X(1/z) \leq \sigma (z L_4)^{1/4} \quad \text{for any } z > 1. \tag{1.16}$$

In Section 5, we obtain extensions of (1.15) to martingales in L^p for some $p > 2$. For instance, if $p = 4$, our result yields

$$Q_{|M|_n^*}(1/z) \leq \sigma \big(1 + \sqrt{(\min(z, L_4) - 1)(z-1)}\ \big)^{1/2} \quad \text{for any } z > 1, \tag{1.17}$$

where σ and L_4 are defined in (1.13), which improves (1.14) in the case $p = 4$.

To conclude this paper, we consider sub-Gaussian martingales. As pointed by Ledoux (1996), entropy methods have interesting applications to concentration inequalities. In Section 6, we apply the results of Section 3 to sub-Gaussian martingales. With this aim in view, we introduce the notion of entropic sub-Gaussian random variable. We then prove that entropic sub-Gaussian random variables satisfy more precise tail inequalities than the usual sub-Gaussian random variables. Finally, in Section 7, we apply the results of Section 6 to the so-called bounded differences inequality.

2 A covariance inequality involving entropy

Throughout this section X is a nonnegative real-valued random variable. We assume that $\mathbb{E}(X \log^+ X) < \infty$ and $\mathbb{E}(X) > 0$. The main result of this section is the covariance inequality below.

Theorem 2.1. *Let X be a nonnegative random variable satisfying the above conditions and η be a real-valued random variable with finite Laplace transform on a right neighborhood of* 0. *Then*

$$\mathbb{E}(X\eta) \le \inf\big\{ b^{-1}\big(\mathbb{E}(X) \log \mathbb{E}\big(e^{b\eta}\big) + \mathcal{H}(X) \big) : b \in]0, \infty[\, \big\}.$$

Proof. A shorter proof can be done using the duality formula for the entropy (*see* Boucheron et al. (2013), Section 4.9 for this formula). However a self-contained proof is more instructive (*see* Remark 2.1 below). Define the two-parameter family of functions $\varphi_{a,b}$ by

$$\varphi_{a,b}(x) = (x/b)\log(x/a) \quad \text{for any } x > 0 \text{ and any positive reals } a \text{ and } b, \tag{2.1}$$

with the convention $0 \log 0 = 0$. Clearly

$$X\eta \le \varphi_{a,b}(X) + \varphi^*_{a,b}(\eta), \quad \text{where } \varphi^*_{a,b}(y) = \sup\{xy - \varphi_{a,b}(x) : x \in [0, \infty[\,\}. \tag{2.2}$$

Next the function $x \mapsto xy - \varphi_{a,b}(x)$ takes its maximum at point $x = ae^{by} - 1$, from which

$$\varphi^*_{a,b}(y) = (a/b)\exp(yb - 1). \tag{2.3}$$

It follows that

$$X\eta \le b^{-1}\big(X \log X - X \log a + a \exp(b\eta - 1)\big). \tag{2.4}$$

Taking the expectation in the above inequality,

$$\mathbb{E}(X\eta) \le b^{-1}\big(\mathbb{E}(X \log X) - \mathbb{E}(X)\log a + (a/e)\mathbb{E}\big(e^{b\eta}\big)\big). \tag{2.5}$$

Let us now minimize the upper bound. Deriving the upper bound with respect to a, we get that the optimal value of a is $a = e\mathbb{E}(X)/\mathbb{E}(e^{b\eta})$. Choosing this value in (2.5), we get that

$$\mathbb{E}(X\eta) \le b^{-1}\big(\mathbb{E}(X \log X) - \mathbb{E}(X)\log \mathbb{E}(X) + \log \mathbb{E}\big(e^{b\eta}\big)\big) \quad \text{for any } b > 0, \tag{2.6}$$

which implies Theorem 2.1.

Remark 2.1. *Notice that the proof of Theorem 3 in Gilat (1986) is based on the inequality $X\eta \le \varphi_{1,b}(X) + \varphi^*_{1,b}(\eta)$, where $b = 1/c$ and $\eta = \log(1/u)$. The minimization with respect to a is omitted, which leads to a suboptimal inequality. The same default appears in the proof of Theorem 7.7 in Osekowski (2012).*

Remark 2.2. *From (2.6) we also have*

$$\mathcal{H}(X) \ge \sup\big\{ b\mathbb{E}(X\eta) - \mathbb{E}(X)\log \mathbb{E}\big(e^{b\eta}\big) : b \in]0, \infty[\, \big\}. \tag{2.7}$$

Recall now the well-known upper bound

$$\mathbb{E}(M_n^*) \le \int_0^1 Q_{M_n}(u)\log(1/u)du, \tag{2.8}$$

which is a direct byproduct of (1.4). If $E(M_n) = 1$, from (2.8) and (2.7) applied with $X = Q_{M_n}(U)$ and $\eta = \log(1/U)$,

$$\mathcal{H}(M_n) \ge \sup\big\{ b\,\mathbb{E}(M_n^*) + \log(1-b) : b \in]0, 1[\, \big\} = g\big(\mathbb{E}(M_n^*)\big), \tag{2.9}$$

which gives a proof of (1.8).

3 Bounds on the tail of M_n^* involving entropy

The main result of this section is the upper bound below on conditional value at risk of X. This upper bound has a variational formulation. From this upper bound we will then derive explicit upper bounds on the tail function of M_n^*.

Theorem 3.1. *Let X be a nonnegative random variable, such that $\mathbb{E}(X) = 1$ and $\mathcal{H}(X) = H$ for some H in $]0, \infty[$. Let $\tilde{Q}_X$ be defined by (1.3). Then, for any $z > 1$*

$$\tilde{Q}_X(1/z) \leq \psi_H(z) \quad \textit{where } \psi_H(z) = \inf_{t>0} t^{-1}\big(H - \log z + \log\big(e^{zt} + z - 1\big)\big). \qquad (a)$$

An other formulation of ψ_H is

$$\psi_H(z) = z \, \inf\big\{\big(H - \log z + \log(c + z - 1)\big)/\log c : c > 1\big\}. \qquad (b)$$

Furthermore

$$\psi_H(z) = z \quad \textit{for any } z \leq e^H \quad \textit{and } \psi_H(z) < z \quad \textit{for any } z > e^H. \qquad (c)$$

Conversely, for any H in $]0, \infty[$ and any $z > 1$, there exist a nonnegative random variable Y such that

$$\mathbb{E}(Y) = 1, \;\; \mathcal{H}(Y) = H \quad \textit{and} \quad \tilde{Q}_Y(1/z) = \psi_H(z). \qquad (d)$$

Proof. We start by the proof of (a). From Theorem 2.1 applied to the random variables $Q_X(U)$ and $B = z\,\mathbf{1}_{zU\leq 1}$,

$$\tilde{Q}_X(1/z) \leq \inf_{t>0} t^{-1}\big(\, H + \log\big(\, z^{-1}e^{zt} + 1 - z^{-1}\,\big)\,\big),$$

which implies (a). To prove (b) it is enough to set $t = z^{-1}\log c$ in the definition of ψ_H. Then $e^{zt} = c$, which gives (b).

To prove (c) and (d), we separate two cases. If $H \geq \log z$,

$$H - \log z + \log(c + z - 1) \geq \log(c + z - 1) \geq \log c.$$

Hence, $\psi_H(z) \geq z$ by Theorem 3.1(b). Now

$$\lim_{c\to\infty} \big(H - \log z + \log(z + c - 1)\big)/\log c = 1, \qquad (3.1)$$

which ensures that $\psi_H(z) = z$. Let $Y = e^H\,\mathbf{1}_{U\leq e^{-H}}$. Then $\mathbb{E}(Y) = 1$, $\mathcal{H}(Y) = H$ and

$$\tilde{Q}_Y(1/z) = z \int_0^{1/z} Q_Y(s)ds = z\int_0^{1/z} e^H\,\mathbf{1}_{s\leq e^{-H}}ds = z = \psi_H(z).$$

which proves (d) in the case $z \leq e^H$. If $H < \log z$, define

$$B = z\,\mathbf{1}_{zU\leq 1}, \;\; Z_t = \exp(tB) \;\; \text{and} \;\; Y_t = Z_t/\mathbb{E}(Z_t). \qquad (3.2)$$

Set

$$R_B(t) = z^{-1}(e^{tz} - 1 + z), \;\; \ell_B(t) = \log R_B(t) \;\; \text{and} \;\; f(t) = t^{-1}(H + \ell_B(t)). \qquad (3.3)$$

(R_B is the Laplace transform of B). By definition, $\psi_H(z)$ is the minimum of f. Now

$$f'(t) = t^{-2}\,\big(\, t\ell_B'(t) - \ell_B(t) - H\,\big).$$

Next ℓ_B is infinitely differentiable, strictly convex and has the asymptotic expansion

$$\ell_B(t) = -\log z + zt + \mathcal{O}(e^{-zt}) \text{ as } t \uparrow \infty.$$

It follows that $g : t \mapsto t\ell'_B(t) - \ell_B(t)$ is continuous, strictly increasing and satisfies $\lim_0 g = 0$ and $\lim_\infty g = \log z > H$. Hence there exists a unique $t_0 > 0$ such that $g(t_0) = H$ and f has a minimum at $t = t_0$. Furthermore, since $f'(t_0) = 0$,

$$\psi_H(z) = \ell'_B(t_0) < z. \tag{3.4}$$

which proves (c) in the case $z > e^H$. Let then $Y = Y_{t_0}$, where Y_t is defined in (3.2): $\mathbb{E}(Y) = 1$ and, with the notations introduced in (3.3),

$$\mathcal{H}(Y) = \mathbb{E}\big((t_0 B - \ell_B(t_0)) \exp(t_0 B)/R_B(t_0) \big) = t_0 \ell'_B(t_0) - \ell_B(t_0) = H.$$

Furthermore, by (3.4),

$$\tilde{Q}_Y(1/z) = e^{t_0 z}/R_B(t_0) = \ell'_B(t_0) = \psi_H(z), \tag{3.5}$$

which gives (d) and completes the proof of Theorem 3.1. $\diamond$

Remark 3.1. *For any nonnegative random variable X and any positive α, $\tilde{Q}_{\alpha X} = \alpha \tilde{Q}_X$ and $\mathcal{H}(\alpha X) = \alpha \mathcal{H}(X)$. Hence Theorem 3.1(a) implies that, for any nonnegative random variable X such that $\mathbb{E}(X) > 0$ and $\mathcal{H}(X) < \infty$,*

$$\tilde{Q}_X(1/z) \le \mathbb{E}(X)\psi_H(z) \ \text{ for any } z > 1, \quad \text{where } H = \mathcal{H}(X)/\mathbb{E}(X). \tag{3.6}$$

Remark 3.2. *From (1.4) and the above Remark, Theorem 3.1 applied to nonnegative submartingale $(M_k)_{0 \le k \le n}$ yields*

$$Q_{M_n^*}(1/z) \le \mathbb{E}(M_n)\psi_H(z) \ \text{ for any } z > 1, \quad \text{where } H = \mathcal{H}(M_n)/\mathbb{E}(M_n). \tag{3.7}$$

By Theorem 3.1(c), $\psi_H(z) < z$ for any $z > e^H$. Consequently, if $\mathbb{E}(M_n)\log z > \mathcal{H}(M_n)$, then $\psi_H(z) < z$ and (3.7) improves Ville's inequality. If $\mathbb{E}(M_n)\log z \le \mathcal{H}(M_n)$, then $\psi_H(z) = z$ and (3.7) does not improve Ville's inequality.

Remark 3.3. *Let μ be any law on $[0, \infty[$ with finite entropy. From Lemma 2 in Dubins and Gilat (1978), there exists a nonnegative continuous time martingale $(M_t)_{t \in [0,1]}$ such that M_1 has the law μ and $M_1^* = \sup\{M_t : t \in [0,1]\}$ has the Hardy-Littlewood maximal distribution associated to μ, which means that $Q_{M_1^*} = \tilde{Q}_{M_1}$. Hence Theorem 3.1 provides an optimal upper bound, at least for continuous time martingales, which shows that Ville's inequality cannot be improved if $z \le e^H$.*

We now give upper bounds on the tail function of M_n^*. For an integrable random variable X, let $\tilde{H}_X$ denote the tail function of the Hardy-Littlewood maximal distribution associated with the law of X. By definition, if U is a random variable with the uniform distribution over $[0, 1]$

$$\tilde{H}_X(t) = \mathbb{P}(\tilde{Q}_X(U) > t) \ \text{ for any real } t, \tag{3.8}$$

where $\tilde{Q}_X$ is given by (1.3). From (1.4),

$$H_{M_n^*} \le \tilde{H}_{M_n}. \tag{3.9}$$

Hence it is enough to bound up $\tilde{H}_{M_n}$. Thus the upper bound below on $\tilde{H}_X$ will be the main ingredient for proving maximal inequalities.

Theorem 3.2. *Let X be a nonnegative random variable, such that $\mathbb{E}(X) = 1$ and $\mathcal{H}(X) = H$ for some H in $]0,\infty[$. For $x > 0$, let $p = \tilde{H}_X(x+1)$ and $v = p/(1-p)$. Define the function L_v^* by*

$$L_v^*(y) = \Big(\frac{v+y}{v+1}\Big)\log\Big(1+\frac{y}{v}\Big) + \Big(\frac{1-y}{v+1}\Big)\log(1-y) \quad \text{if } y \in [0,1] \tag{3.10}$$

and $L_v^*(y) = +\infty$ for $y > 1$. Define also the nonnegative function h by

$$h(x) = (1+x)\log(1+x) - x \quad \text{for } x \geq -1 \quad \text{and} \quad h(x) = +\infty \quad \text{for } x < -1. \tag{3.11}$$

Then

$$L_v^*(vx) \leq H \quad \text{or, equivalently,} \quad ph(x) + (1-p)h(-vx) \leq H.$$

Proof. For any positive v, define the Bernoulli type random variable ξ by

$$\mathbb{P}(\xi = 1) = v/(1+v) \ \text{ and } \mathbb{P}(\xi = -v) = 1/(1+v).$$

Let

$$L_v(t) = \log \mathbb{E}\big(e^{t\xi}\big) = \log\big(ve^t + e^{-vt}\big) - \log(1+v). \tag{3.12}$$

Define the Legendre-Fenchel dual L^* of the convex and increasing function $L : \mathbb{R}^+ \mapsto \mathbb{R}^+$ by

$$L^*(x) = \sup\{xt - L(t) : t > 0\} \quad \text{for any } x \geq 0. \tag{3.13}$$

Then L^* is convex and increasing. From formula (2.55), page 29, in Bercu et al. (2015), the Legendre-Fenchel dual L_v^* of L_v is given by (3.10). Now the inversion formula below holds true (*see* Bercu et al. (2015), page 57):

$$L^{*-1}(x) = \inf\{t^{-1}(L(t) + x) : t > 0\} \quad \text{for any } x \geq 0. \tag{3.14}$$

With the above notations, if ψ_H is the function already defined in Theorem 3.1(a),

$$\psi_H(z) - 1 = v^{-1} \inf_{t>0} t^{-1}\big(H + L_v(t)\big) = v^{-1} L_v^{*-1}(H) \quad \text{if } v = (z-1)^{-1}. \tag{3.15}$$

Applying now Theorem 3.1(a) with $z = 1/p$ and noticing that $\tilde{Q}_X(p) = x+1$ (thanks to the continuity of $\tilde{Q}_X$), we get that $x \leq \psi_H(1/p) - 1$. From (3.15), this inequality is equivalent to $vx \leq L_v^{*-1}(H)$. Since L_v^* is nondecreasing and $L_v^*(L_v^{*-1}(H)) \leq H$, this inequality implies the first part of Theorem 3.2.

Now, by identity (2.66) in Bercu et al. (2015),

$$L_v^*(vx) = v(1+v)^{-1}h(x) + (1+v)^{-1}h(-vx) = ph(x) + (1-p)h(-vx), \tag{3.16}$$

which concludes the proof of Theorem 3.2. ⋄

We now derive explicit upper bounds on $\tilde{H}_X$ from Theorem 3.2.

Theorem 3.3. *For any $x > 0$, let $h_0 = x^2/2$ and $h_1 = (1+x)\log(1+x) - x$. Under the assumptions of Theorem 3.2,*

$$\tilde{H}_X(x+1) \leq 2H\Big/\Big(2H + h_1 + \sqrt{h_1^2 + 4H(h_0 - h_1)}\Big) < H/(H+h_1). \tag{a}$$

Hence, if $x \leq \sqrt{2H}$, then $\tilde{H}_X(x+1) \leq H/(H+h_0)$.

From Theorem 3.3 and (3.9), we immediately get the corollary below.

Corollary 3.1. *Let $(M_k)_{0\le k\le n}$ be a nonnegative submartingale such that $\mathbb{E}(M_n)>0$ and $\mathcal{H}(M_n)<\infty$. Set $H=\mathcal{H}(M_n)/\mathbb{E}(M_n)$. With the same notations as in Theorem 3.3, for any $x>0$,*

$$\mathbb{P}\big(M_n^*\ge \mathbb{E}(M_n)(1+x)\big)\le 2H\Big/\Big(2H+h_1+\sqrt{h_1^2+4H(h_0-h_1)}\Big)<H/(H+h_1). \quad (a)$$

Consequently

$$\mathbb{P}\big(M_n^*\ge \mathbb{E}(M_n)(1+x)\big)\le 2H/(2H+x^2) \ \ \textit{for any } x \ \textit{ in } [0,\sqrt{2H}\,]. \quad (b)$$

Remark 3.4. *The above bounds own the same structure as the Tchebichef-Cantelli inequality. Note that the first upper bound in (a) is equivalent $H/(H+h_1)$ as $x\nearrow\infty$.*

Remark 3.5. *Setting $x=\sqrt{2H}$ in Corollary 3.1(b), $\mu(M_n^*)\le\mathbb{E}(M_n)(1+\sqrt{2H}\,)$. This result improves the trivial upper bound $\mu(M_n^*)\le 2\mathbb{E}(M_n)$ iff $H<1/2$.*

Proof of Theorem 3.3. Let $x>0$. We start by proving that

$$\varphi_v(x):=L_v^*(vx)\ge v(1-v)h(x)+v^2(x^2/2) \ \text{ for any } v>0 \text{ and any } \ x>0. \quad (3.17)$$

To prove (3.17), we derive φ_v twice:

$$\varphi_v''(x)=v(1+x)^{-1}(1-vx)^{-1}=p\big((1+x)^{-1}+v(1-vx)^{-1}\big), \quad (3.18)$$

where $p=v/(1+v)$ (*see* Bercu et al. (2015), page 34). Deriving again,

$$\varphi_v^{(3)}(x)=p\big(-(1+x)^{-2}+v^2(1-vx)^{-2}\big)\ge p(v^2-1)(1+x)^{-2}. \quad (3.19)$$

Next, integrating this inequality and using the initial condition $\varphi_v''(0)=p(1+v)$,

$$\varphi_v''(x)\ge p(1-v^2)(1+x)^{-1}+pv(1+v)=v(1-v)h''(x)+v^2, \quad (3.20)$$

since $p(1{+}v)=v$. Finally, integrating twice this inequality and using the initial conditions $\varphi(0)=\varphi'(0)=0$, we get (3.17).

Let $p=\tilde H_X(x+1)$, $v=p/(1-p)$, $h_1=h(x)$ and $h_0=x^2/2$. From Theorem 3.2 and (3.17),

$$v^2(h_0-h_1)+vh_1-H\le 0. \quad (3.21)$$

In the above inequation $h_0-h_1>0$. Solving this inequation of order two with respect to v, we obtain that

$$v\le(-h_1+\sqrt{\Delta}\,)/(2h_0-2h_1), \quad \text{where } \ \Delta=h_1^2+4H(h_0-h_1).$$

Now $(-h_1+\sqrt{\Delta}\,)/(2h_0-2h_1)=2H/(h_1+\sqrt{\Delta}\,)$. Consequently

$$v\le\frac{2H}{h_1+\sqrt{\Delta}}:=\frac{A}{B}. \quad (3.22)$$

Since $p=v/(1+v)$, it follows that $p\le A/(A+B)$, which gives the first part of (a). The second part of (a) follows by noting that $4H(h_0-h_1)>0$.

If furthermore $x \le \sqrt{2H}$, then $h_0 \le H$, which implies that

$$h_1^2 + 4H(h_0 - h_1) \ge h_1^2 + 4h_0(h_0 - h_1) = (2h_0 - h_1)^2.$$

From the above inequality and the fact that $2h_0 - h_1 > 0$,

$$2H \Big/ \Big(2H + h_1 + \sqrt{h_1^2 + 4H(h_0 - h_1)}\,\Big) \le 2H/(2H + 2h_0),$$

which ends up the proof. $\diamond$

Numerical comparisons. To conclude this section, we compare Corollary 3.3 with usual tail inequalities for maxima of martingales. Here we assume that $(M_k)_{0 \le k \le n}$ is a positive martingale such that $\mathbb{E}(M_n) = 1$. Then, by the Ville inequality

$$\mathbb{P}(M_n^* \ge 1 + x) \le 1/(1 + x). \tag{3.23}$$

Next, let h be defined by (3.11): by the Ville inequality applied to the nonnegative submartingale $(h(M_k - 1)\,)_{0 \le k \le n}$,

$$\mathbb{P}(M_n^* \ge 1 + x) \le H/h_1 \ \text{ where } H = \mathcal{H}(M_n) \ \text{ and } h_1 = h(x). \tag{3.24}$$

Finally, by Corollary 3.1, if $h_0 = x^2/2$ and $h_1 = h(x)$,

$$\mathbb{P}(M_n^* \ge 1 + x) \le 2H \Big/ \Big(2H + h_1 + \sqrt{h_1^2 + 4H(h_0 - h_1)}\,\Big) \tag{3.25}$$

which implies the weaker inequality

$$\mathbb{P}(M_n^* \ge 1 + x) \le H/(H + h_1). \tag{3.26}$$

Below I give the numerical values of the above upper bounds for $H = 1/2$ and $x = \sqrt{e} - 1 = 0.649$, $x = 1$, $x = 2$, $x = 4$, $x = 8$, $x = 24$ and $x = 99$.

Ineq.	x= 0.649	x=1	x=2	x=4	x=8	x=24	x=99
(3.23)	0.607	0,500	0,333	0,200	0,1111	0,04000	0,01000
(3.24)	2.847	1.294	0,386	0,123	0,0425	0,00885	0,00138
(3.25)	0.670	0,500	0,247	0,100	0,0382	0,00848	0,00135
(3.26)	0.822	0,564	0,278	0,110	0,0407	0,00878	0,00138

From Remarks 3.2 and 3.3, Inequality (3.23) is optimal for $x \le e^H - 1$, which motivates the choice $x = \sqrt{e} - 1$ for the first column. One can see that (3.25) gives better estimates for $x = 2$ and $x = 4$ than (3.24) and (3.26). For $x = 99$ or $x = 24$, (3.24) and (3.26) are nearly equivalent and (3.25) remains more efficient than (3.26).

4 Tchebichef type inequalities

At the present time the Tchebichef-Cantelli inequality has not yet been extended to random variables in L^p, for arbitrary $p > 1$. In this section we give an extension of this inequality to the Hardy-Littlewood maximal distribution associated with the law of a random variable X in L^p. Next we apply this result to submartingales in L^p. So, let $(M_k)_{k \in [0,n]}$ is a submartingale in L^p. From Gilat and Meilijson (1988), the nonnegativity assumption can be dropped in (1.4). Hence, in order to bound $Q_{M_n^*}$, it is enough to bound up $\tilde{Q}_X$ for a random variable X in L^p.

Theorem 4.1. *Let p be any real in $[1, \infty[$ and X be a real-valued random variable in L^p. Let $\tilde{Q}_X$ be defined by (1.3). Then*

$$\tilde{Q}_X(1/z) \leq \mathbb{E}(X) + z^{1/p}(1 + (z-1)^{1-p})^{-1/p}\|X - \mathbb{E}(X)\|_p \quad \textit{for any } z > 1. \qquad (a)$$

Conversely, for any $p > 1$ and any $z > 1$, there exists a random variable X in L^p such that

$$\mathbb{E}(X) = 0, \ \|X\|_p = 1 \quad \textit{and} \quad \tilde{Q}_X(1/z) = z^{1/p}(1 + (z-1)^{1-p})^{-1/p}. \qquad (b)$$

If furthermore X has a symmetric law, then

$$\tilde{Q}_X(1/z) \leq \|X\|_1 + (z/2)^{1/p}(1 + (z/2-1)^{1-p})^{-1/p}\| \, |X| - \|X\|_1 \, \|_p \quad \textit{for any } z > 2. \quad (c)$$

Remark 4.1. *If $p = 2$, the upper bound is equal to $\mathbb{E}(X) + \sigma\sqrt{z-1}$, where σ is the standard deviation of X. Since $Q_X \leq \tilde{Q}_X$, it implies (1.10). For $p > 1$, the upper bound tends to $\mathbb{E}(X)$ as $z \searrow 1$, which proves that Theorem 4.1 is efficient for any value of z.*

Remark 4.2. *Assume that $\mathbb{E}(X) = 0$. If $p = 2$, from Remark 4.1, Theorem 4.1(a) is equivalent to the tail inequality $\tilde{H}_X(x) \leq \sigma^2/(\sigma^2 + x^2)$. For p in $\{3, 4, 3/2, 4/3\}$, one can also derive explicit tail inequalities from Theorem 4.1(a).*

Applying (1.4), we immediately get the corollary below.

Corollary 4.1. *Let p be any real in $[1, \infty[$ and $(M_k)_{k \in [0,n]}$ be a submartingale in L^p:*

$$Q_{M_n^*}(1/z) \leq \mathbb{E}(M_n) + z^{1/p}(1 + (z-1)^{1-p})^{-1/p}\|M_n - \mathbb{E}(M_n)\|_p \quad \textit{for any } z > 1. \qquad (a)$$

If furthermore M_n has a symmetric law, then

$$Q_{M_n^*}(1/z) \leq \|M_n\|_1 + (z/2)^{1/p}(1 + (z/2-1)^{1-p})^{-1/p}\| \, |M_n| - \|M_n\|_1 \, \|_p \quad \textit{for any } z > 2. \quad (b)$$

Remark 4.3. *For $z = 2$, Corollary 4.1(a) yields*

$$\mu(M_n^*) \leq \mathbb{E}(M_n) + \|M_n - \mathbb{E}(M_n)\|_p. \qquad (4.1)$$

In particular (4.1) ensures that the median of M_n^* is less than $\mathbb{E}(M_n) + \|M_n - \mathbb{E}(M_n)\|_1$. Note however that (4.1) is an immediate consequence of Ville's inequality applied to the submartingale $\big((M_k - \mathbb{E}(M_n))_+\big)_{0 \leq k \leq n}$.

Remark 4.4. *For $p = 2$, in the symmetric case, one can prove that Corollary 4.1(b) is strictly more efficient than Corollary 4.1(a) for any $z > 2$.*

Proof of Theorem 4.1. Clearly it suffices to prove the result in the case $\mathbb{E}(X) = 0$. Then $\tilde{Q}_X(1) = 0$. Set $u = 1/z$. For any b in $[0, 1]$,

$$u\,\tilde{Q}_X(u) = u\,\tilde{Q}_X(u) - b\,\tilde{Q}_X(1) = \int_0^1 Q_X(s)\big(\mathbf{1}_{s \leq u} - b\big)ds. \qquad (4.2)$$

For $p = 1$, choosing $b = 1/2$ in (4.2), we get that

$$u\,\tilde{Q}_X(u) \leq \tfrac{1}{2}\int_0^1 |Q_X(s)|ds = \tfrac{1}{2}\mathbb{E}|X|,$$

which implies Theorem 4.1(a) in the case $p = 1$. For $p > 1$, applying the Hölder inequality on $[0, 1]$ with exponents p and $q = p/(p-1)$ to the functions Q_X and $\mathbf{1}_{[0,u]} - b$, we get that

$$u\,\tilde{Q}_X(u) \le \sigma_p \left(u(1-b)^q + (1-u)b^q\right)^{1/q}$$

or, equivalently,

$$\tilde{Q}_X(u) \le \sigma_p z^{1/p} \left((1-b)^q + (z-1)b^q\right)^{1/q}. \tag{4.3}$$

We now minimize the upper bound with respect to b. Let $f(b) = (1-b)^q + (z-1)b^q$. Then f is strictly convex and

$$q^{-1} f'(b) = -(1-b)^{q-1} + (z-1)b^{q-1} = 0 \ \text{ iff } \ z - 1 = (1-b)^{q-1}/b^{q-1}.$$

Next $1/(q-1) = p-1$. Consequently the critical point b_0 exists and

$$1 - b_0 = b_0(z-1)^{p-1}, \ \text{ whence } b_0 = 1/\left(1 + (z-1)^{p-1}\right).$$

Setting $b = b_0$ in (4.3), we then get that

$$\tilde{Q}_X(u) \le \sigma_p z^{1/p}(z-1)^{1/q}\left((z-1)^{p-1} + 1\right)^{-1/p},$$

which gives Theorem 4.1(a).

We now prove Theorem 4.1(b). Let X be the Bernoulli random variable defined by

$$\mathbb{P}\left(X = z^{\frac{1}{p}}(1+(z-1)^{1-p})^{-\frac{1}{p}}\right) = 1/z = 1 - \mathbb{P}\left(X = -(z-1)^{-1} z^{\frac{1}{p}}(1+(z-1)^{1-p})^{-\frac{1}{p}}\right). \tag{4.4}$$

Then $\mathbb{E}(X) = 0$ and $\tilde{Q}_X(1/z) = z^{1/p}(1 + (z-1)^{1-p})^{-1/p})$. Furthermore

$$\mathbb{E}|X|^p = \left((1/z)z + (1 - 1/z)(z-1)^{-p} z\right)(1 + (z-1)^{1-p})^{-1} = 1,$$

which ends up the proof of Theorem 4.1(b).

To prove (c), it suffices to prove that, for any real-valued random variable X in L^1 with a symmetric law,

$$\tilde{Q}_X(1/z) = \tilde{Q}_{|X|}(2/z) \ \text{ for any } z > 2, \tag{4.5}$$

and next to apply (a) to the random variable $|X|$. Now, for any symmetric random variable X and any positive x, $H_{|X|}(x) = 2H_X(x)$, which implies that $Q_X(s) = Q_{|X|}(2s)$ for any $s < 1/2$. Hence, for any $z > 2$,

$$\tilde{Q}_X(1/z) = z \int_0^{1/z} Q_{|X|}(2s)ds = (z/2) \int_0^{2/z} Q_{|X|}(u)du,$$

which proves (4.5). Hence Theorem 4.1(c) holds true. $\diamond$

We now apply Corollary 4.1 to martingales in L^p for p in $]1, 2]$ and we compare the so obtained upper bound with the upper bound that can be derived from (1.9). So, let $(M_k)_{k\in[0,n]}$ be a martingale in L^p. Let

$$X_k = M_k - M_{k-1} \ \text{ and } \Delta_p = \mathbb{E}|X_1]^p + \cdots + \mathbb{E}|X_n|^p. \tag{4.6}$$

By Proposition 1.8 in Pinelis (2015),

$$\mathbb{E}|M_n|^p \le \mathbb{E}|M_0|^p + K_p\Delta_p \ \text{ where } K_p = \sup_{x\in[0,1]} \left(px^{p-1} + (1-x)^p - x^p\right). \tag{4.7}$$

As shown by Pinelis (2015), for $p < 2$ the constant K_p is strictly larger than 1. Furthermore this constant is decreasing with respect to p and tends to 2 as $p \searrow 1$.

From (4.7) and Corollary 4.1, we get the corollary below for martingales.

Corollary 4.2. *Let p be any real in $]1,2]$ and $(M_k)_{k\in[0,n]}$ be a martingale in L^p such that $M_0 = 0$. Then*

$$Q_{M_n^*}(1/z) \le (K_p\Delta_p)^{1/p} z^{1/p}(1+(z-1)^{1-p})^{-1/p} \quad \textit{for any } z > 1.$$

Now, from Proposition 8.1(b) in the Annex, under the conditions of Corollary 4.2,

$$\mathbb{E}\big((M_n + t)_+^p \big) \le t^p + \Delta_p \ \text{ for any } t \ge 0. \tag{4.8}$$

Starting from (1.4) and applying the above inequality, one obtains the theorem below.

Theorem 4.2. *Let p be any real in $]1,2]$ and $(M_k)_{k\in[0,n]}$ be a martingale in L^p such that $M_0 = 0$. Set $q = p/(p-1)$. Then*

$$Q_{M_n^*}(1/z) \le \Delta_p^{1/p} z^{1/p}\big(1 - z^{1-q}\big)^{1/q} \quad \textit{for any } z > 1.$$

Proof. We prove Theorem 4.2 in the case $\Delta_p = 1$. The general case follows by dividing the random variables X_k by $\Delta_p^{1/p}$. From (1.4)

$$Q_{M_n^*}(1/z) \le z \int_0^{1/z} Q_{M_n}(s)ds \le -t + z\int_0^{1/z} (Q_{M_n}(s)+t)_+ ds \ \text{ for any } z > 1.$$

Let U be a random variable with uniform law over $[0,1]$. Since $Q_{(M_n+t)_+}(U)$ has the same law as $(M_n + t)_+$ and $(Q_{M_n}(s) + t)_+ = Q_{(M_n+t)_+}(s)$,

$$\int_0^{1/z} (Q_{M_n}(s)+t)_+ ds = \mathbb{E}(Q_{(M_n+t)_+}(U)\, \mathbf{1}_{U\le 1/z}) \le z^{-1/q} \|(M_n+t)_+\|_p$$

by the Hölder inequality. Noticing that $1-1/q = 1/p$, the two above inequalities together with (4.8) imply that

$$Q_{M_n^*}(1/z) \le -t + z^{1/p}(t^p+1)^{1/p} \text{ for any } t \ge 0.$$

Hence, minimizing with respect to t

$$Q_{M_n^*}(1/z) \le Q_p(1/z), \quad \text{where } \ Q_p(1/z) = \inf\{-t + z^{1/p}\big(1+t^p\big)^{1/p} : t \in \mathbb{R}^+\}. \tag{4.9}$$

In view of the above inequality, it only remains to prove that

$$Q_p(1/z) = \big(z^{q-1} - 1\big)^{1/q}, \quad \text{where } q = p/(p-1). \tag{4.10}$$

Now the function $f : t \mapsto -t + z^{1/p}\big(1+t^p\big)^{1/p}$ is convex and positive on $\mathbb{R}^+$. If $z = 1$, $\lim_{t\uparrow\infty} f(t) = 0$, which implies that $Q_p(1) = 0$. Otherwise the function f has a unique minimum at point $t = t_z = (z^{1/(p-1)} - 1)^{-1/p}$ and $f(t_z) = \big(z^{q-1}-1\big)^{1/q}$, which completes the proof of (4.10). ⋄

Remark 4.5. *Since $K_2 = 1$, Theorem 4.2 and Corollary 4.2 are equivalent for $p = 2$.*

To conclude this section, we compare Theorem 4.2 and Corollary 4.2 for p in $]1,2[$ and $\Delta_p = 1$. As $z \nearrow \infty$, the upper bound in Corollary 4.2 is equivalent to $K_p^{1/p} z^{1/p}$, while Theorem 4.2 gives an upper bound equivalent to $z^{1/p}$ as $z \nearrow \infty$. Since $K_p > 1$, it appears that Theorem 4.2 provides better bounds for large values of z. However Corollary 4.2 provides better estimates for small values of z, as shown by the numerical table below.

In the table below, I give the values of the upper bounds corresponding to Corollary 4.2 and Theorem 4.2 in the special case $p = 3/2$. Then $K_p = 1.306...$ as shown in Pinelis (2015). It appears that Corollary 4.2 provides better bounds for small values of z, including $z = 10$.

Inequality	z=2	z=4	z=6	z=8	z=10	z=12	z=14	z=16	z=20
Corol. 4.2.	1.20	2.22	3.08	3.86	4.58	5.26	5.90	6.51	7.67
Theor. 4.2.	1.44	2.47	3.27	3.98	4.63	5.23	5.80	6.34	7.36

5 Cantelli type inequalities

Let p be any real strictly more than 1 and X be a centered random variable in L^{2p}. In this section we give an extension of the Cantelli inequality to the Hardy-Littlewood maximal distribution associated with $|X|$. Next we apply this result to martingales in L^{2p}. Let us start by our extension of Cantelli's inequality.

Theorem 5.1. *Let p be any real in $]1,\infty[$ and X be a real-valued random variable in L^{2p}, such that $\mathbb{E}(X) = 0$. Set $\sigma^2 = \mathbb{E}(X^2)$. Then, for any $z > 1$,*

$$\tilde{Q}_{|X|}(1/z) \le \big(\sigma^2 + z^{1/p}(1 + (z-1)^{1-p})^{-1/p}\|X^2 - \sigma^2\|_p\big)^{1/2}. \qquad (a)$$

Let a be any positive real and let $z_p > 1$ be the unique solution of the equation

$$z - 1 + (z-1)^p = za^p. \qquad (5.1)$$

Then, for any $z \ge z_p$, there exists a symmetric random variable X in L^{2p} such that

$$\|X\|_2 = 1, \ \|X^2 - 1\|_p = a \ \text{ and } \ \tilde{Q}_{|X|}(1/z) = \big(1 + z^{1/p}(1 + (z-1)^{1-p})^{-1/p}a\big)^{1/2}. \qquad (b)$$

If furthermore X has a symmetric law, then

$$\tilde{Q}_X(1/z) \le \big(\sigma^2 + (z/2)^{1/p}(1 + (z/2-1)^{1-p})^{-1/p}\|X^2 - \sigma^2\|_p\big)^{1/2} \ \text{ for any } z > 2. \qquad (c)$$

Now, recall that, if $(M_k)_{k\in[0,n]}$ is a martingale in L^1, then $(|M_k|)_{k\in[0,n]}$ is a submartingale in L^1. Hence, from Theorem 5.1 and (1.4) we immediately get the corollary below.

Corollary 5.1. *Let p be any real in $]1,\infty[$ and $(M_k)_{k\in[0,n]}$ be a martingale in L^{2p} such that $\mathbb{E}(M_0) = 0$. Let $|M|_n^*$ be defined by (1.12). Set $V_n = \mathbb{E}(M_n^2)$. Then, for any $z > 1$,*

$$Q_{|M|_n^*}(1/z) \le \big(V_n + z^{1/p}(1 + (z-1)^{1-p})^{-1/p}\|M_n^2 - V_n\|_p\big)^{1/2}. \qquad (a)$$

If furthermore M_n has a symmetric law, then, for any $z > 2$,

$$Q_{M_n^*}(1/z) \le \big(V_n + (z/2)^{1/p}(1 + (z/2-1)^{1-p})^{-1/p}\|M_n^2 - V_n\|_p\big)^{1/2}. \qquad (b)$$

Remark 5.1. *From (1.14) applied with $p = 2$*

$$Q_{|M|_n^*}(1/z) \le \sigma\sqrt{z} \ \text{ for any } z > 1, \quad \text{where } \ \sigma = \sqrt{V_n}. \qquad (5.2)$$

Let then $a_p = \|V_n^{-1}M_n^2 - 1\|_p$ and let z_p' be the solution of equation (5.1) with $a = a_p$. One can easily prove that the upper bound in (a) is strictly less that $\sqrt{V_n z}$ if and only if $z > z_p'$. Then, from Theorem 5.1(b) and Remark 3.3, Corollary 5.1 cannot be further improved.

Remark 5.2. *When $p = 2$, Corollary 5.1(a) and (5.2) yield (1.17). Under the same assumptions (1.14) gives the upper bound*

$$Q_{|M|_n^*}(1/z) \leq \sigma(L_4 z)^{1/4}. \tag{5.3}$$

Let us now compare (1.17), (5.2) and (5.3). Clearly (1.17) and (5.2) are equivalent for $z \leq L_4$. Next (5.3) and (1.17) are strictly more efficient than (5.2) iff $z > L_4$. Now the upper bound in (5.3) is less than the upper bound in (1.17) iff

$$zL_4 \leq \left(1 + \sqrt{(L_4 - 1)(z-1)}\right)^2, \quad \textit{i.e.} \quad (z-1) + (L_4 - 1) \leq 2\sqrt{(L_4-1)(z-1)}.$$

The above inequality holds true if and only if $z = L_4$. Hence (1.17) is strictly more efficient than (5.3) for $z \neq L_4$. Consequently (1.17) is more efficient than (5.2) and (5.3) for any value of z.

Proof of Theorem 5.1. By the Jensen inequality,

$$\tilde{Q}_{|X|}(1/z) \leq \sqrt{\tilde{Q}_{X^2}(1/z)}.$$

Theorem 5.1(a) follows from Theorem 4.1(a) applied to the random variable X^2 via the above inequality. Now Theorem 5.1(c) follows from Theorem 5.1(a) and (4.5). It remains to prove (b). Let $u = 1/z$. Define the random variable Y by

$$\mathbb{P}\big(Y = az^{\frac{1}{p}}(1 + (z-1)^{1-p})^{-\frac{1}{p}}\big) = u = 1 - \mathbb{P}\big(Y = -a(z-1)^{-1}z^{\frac{1}{p}}(1 + (z-1)^{1-p})^{-\frac{1}{p}}\big).$$

If $z \geq z_p$, then $Y + 1 \geq 0$. Let then ε be a symmetric sign, independent of Y. Define X by $X = \varepsilon\sqrt{Y+1}$. Then $\mathbb{E}(X) = 0$, $\mathbb{E}(X^2) = 1 + \mathbb{E}(Y) = 1$ and $\|X^2 - 1\|_p = \|Y\|_p = a$. Now

$$\tilde{Q}_{|X|}(1/z) = \Big(1 + az^{\frac{1}{p}}(1 + (z-1)^{1-p})^{-\frac{1}{p}}\Big)^{1/2},$$

which completes the proof of Theorem 5.1(b). ⋄

We now apply Corollary 5.1 to sums of independent random variables. Here it will be convenient to introduce a condition of fourth order on the random variables.

Definition 5.1. *A real-valued random variable X in L^4 is said to be sub-Gaussian at order 4 if X satisfies $\|X - \mathbb{E}(X)\|_4^4 \leq 3\|X - \mathbb{E}(X)\|_2^4$.*

Let $X_1, X_2, \ldots$ be a sequence of independent centered random variables in L^4. Suppose furthermore that these random variables are sub-Gaussian at order 4. Let

$$M_0 = 0 \ \text{ and } M_k = X_1 + X_2 + \cdots + X_k \ \text{ for } k > 0. \tag{5.4}$$

Then $(M_k)_{k\in[0,n]}$ satisfies the assumtions of Corollary 5.1 with $p = 2$. Now, by the usual inequality for moments of order 4 of sums of independent random variables,

$$L_4 := V_n^{-2}\mathbb{E}(M_n^4) = 3 + V_n^{-2}\sum_{k=1}^n \Big(\mathbb{E}(X_k^4) - 3(\mathbb{E}(X_k^2))^2\Big), \tag{5.5}$$

which shows that $L_4 \leq 3$ if the the random variables X_k are sub-Gaussian at order 4. Hence Corollary 5.1(a) and (5.2) imply the proposition below.

Proposition 5.1. *Let $X_1, X_2, \dots$ be a sequence of independent centered random variables in L^4. Suppose furthermore that these random variables are sub-Gaussian at order 4. Let the martingale $(M_k)_{0\le k\le n}$ be defined by (5.4). Then*

$$Q_{|M|_n^*}(1/z) \le V_n^{1/2}\big(1 + \sqrt{(\min(z,3)-1)(z-1)}\,\big)^{1/2} \quad \textit{for any } z > 1. \tag{5.6}$$

Remark 5.3. *Inequality (5.6) is equivalent to the tail inequality*

$$\mathbb{P}(|M|_n^* \ge \sqrt{V_n}\, x) \le 2\big(2 + (x^2-1)\max(x^2-1, 2)\,\big)^{-1}.$$

The above upper bound is equivalent to $(2/x^4)$ as $x \nearrow \infty$. Under the same conditions, (5.3) yields the less efficient upper bound $(3/x^4)$.

Assume now that the random variables $X_1, X_2, \dots$ are symmetric. By (1.4) and (4.5),

$$Q_{M_n^*}(1/z) \le \tilde{Q}_{M_n}(1/z) \le \tilde{Q}_{|M_n|}(2/z).$$

It follows that

$$Q_{M_n^*}(1/z) \le \sigma\sqrt{z/2} \ \text{ for any } z > 2, \ \text{ where } \sigma = \sqrt{V_n}. \tag{5.7}$$

Then the above inequality and Corollary 5.1(b) imply the proposition below.

Proposition 5.2. *Let $X_1, X_2, \dots$ be a sequence of independent symmetric random variables in L^4. Suppose furthermore that these random variables are sub-Gaussian at order 4. Let the martingale $(M_k)_{0\le k\le n}$ be defined by (5.4). Then*

$$Q_{M_n^*}(1/z) \le V_n^{1/2}\big(1 + \sqrt{(\min(\,(z-2)/4, 1)\,(z-2)}\,\big)^{1/2} \quad \textit{for any } z \ge 2. \tag{5.8}$$

Example 5.1. Here we compare Proposition 5.1 with the so-called Kearns-Saul inequality (*see* Bercu et al. (2015), Section 2.5) in the case of weighted sums of Bernoulli type random variables. Let $\eta_1, \eta_2, \dots$ be a sequence of Bernoulli random variables with law $b(p)$ for some $p < 1/2$. Let $a_1, a_2, \dots, a_n$ be a finite sequence of real numbers. Set

$$X_k = a_k(\eta_k - p) \ \text{ for any } k > 0. \tag{5.9}$$

From the elementary inequality

$$\|\eta_k - p\|_4^4 - 3\|\eta_k - p\|_2^4 = 1 - 6p(1-p), \tag{5.10}$$

the random variables η_k are sub-Gaussian at order 4 if and only if $p \ge (1 - 3^{-1/2})/2$. Under this condition, Proposition 5.1 gives

$$V_n^{-1/2} Q_{|M|_n^*}(1/z) \le \big(1 + \sqrt{2(z-1)}\,\big)^{1/2} \ \text{ for any } z \ge 3. \tag{5.11}$$

Now, let ℓ_p denote the log-Laplace transform of $\eta_1 - p$. Hoeffding (1963, Section 4) proved that

$$\inf_{x>0} x^{-2}\ell_p^*(x) = \frac{\log(1/p-1)}{1-2p} \ \text{ and } \ \inf_{x<0} x^{-2}\ell_p^*(x) = \frac{1}{2p(1-p)}.$$

Using the fact that $(\ell_p^*)^* = \ell_p$, it appears immediately that the above inequality is equivalent to [2]

[2]The first constant has been rediscovered independently by Bobkov (1998) and Kearns and Saul (1998)

$$\sup_{t>0} t^{-2}\ell_p(t) = \frac{1-2p}{4\log(1/p-1)} \quad \text{and} \quad \sup_{t<0} t^{-2}\ell_p(t) = \frac{p(1-p)}{2}.$$

Since $(1-2p)/\log(1/p-1) \leq 2p(1-p)$, it follows that, for any real t,

$$\log \mathbb{E}(e^{tM_n}) \leq \frac{(1-2p)|a|_2^2 t^2}{4\log(1/p-1)}, \quad \text{where} \quad |a|_2^2 = \sum_{k=1}^{n} a_k^2.$$

From this upper bound and usual arguments on exponential martingales,

$$V_n^{-1/2} Q_{|M|_n^*}(1/z) \leq \sqrt{c_p \log(2z)}, \quad \text{where} \quad c_p = \frac{1-2p}{p(1-p)\log(1/p-1)}. \tag{5.12}$$

Below I give the numerical values of the upper bounds (5.2), (5.11) and (5.12) when $p(1-p) = 1/6$, for some integer values of z.

Ineq.	z=5	z=10	z=20	z= 30	z= 40	z=50	z=75	z=100
(5.11)	1.96	2.29	2.68	2.94	3.14	3.30	3.63	3.88
(5.12)	2.46	2.81	3.11	3.28	3.40	3.48	3.63	3.73
(5.2)	2.24	3.16	4.47	5.48	6.32	7.07	8.66	10.00

It appears here that the Cantelli type inequality (5.11) is more efficient for $z \leq 75$, which includes values of statistical interest. For $z > 75$, (5.12) provides better bounds. The Kolmogorov inequality (5.2), efficient for $z \leq 3$, is of poor quality for $z > 3$.

Example 5.2. Let $X_1, X_2, \ldots$ be a sequence of independent and symmetric random variables in L^4 with variance 1. Assume furthermore that $\mathbb{E}(X_k^4) = 3$ for any positive k. Then, by Proposition 5.2,

$$n^{-1/2} Q_{M_n^*}(1/z) \leq \left(1 + \sqrt{z-2}\right)^{1/2} \quad \text{for any } z \geq 2. \tag{5.13}$$

We now compare (5.13) with the Fuk-Nagaev type inequality established in Rio (2017). Under the above conditions, Theorem 3.1(b) in Rio (2017) applied with $q = 4$ yields

$$n^{-1/2} Q_{M_n^*}(1/z) \leq \sqrt{2\log z} + 1.50\, n^{-1/4}(3z/2)^{1/4} \quad \text{for any } z > 1. \tag{5.14}$$

For $n \geq 8$, (5.14) is asymptotically more efficient than (5.13).

Below I give the numerical values of the upper bounds (5.13) and (5.14) when $n = 5^4 = 625$, for some integer values of z.

Ineq.	z=20	z=50	z=100	z= 200	z= 500	z=820	z=1000	z=10000
(5.13)	2.29	2.82	3.30	3.88	4.83	5.44	5.71	10.05
(5.14)	3.15	3.68	4.08	4.50	5.10	5.44	5.58	7.68

It appears here that the Cantelli type inequality (5.13) is more efficient for $z \leq 820$. However (5.14) is much more efficient for large values of z.

6 Entropic sub-Gaussian random variables

In this section, we are interested in sub-Gaussian random variables. For any real-valued random variable X with a finite Laplace transform on $\mathbb{R}$, define

$$\ell_X(t) = \log \mathbb{E}\left(e^{tX}\right) \quad \text{for any real } t. \tag{6.1}$$

Let b be any positive real. The random variable X is said to be sub-Gaussian with parameter b iff X has a finite Laplace transform on $\mathbb{R}$ and

$$\ell_X(t) \leq t\,\mathbb{E}(X) + b^2(t^2/2) \quad \text{for any} \quad t > 0. \tag{6.2}$$

This property implies that the variance of X is bounded by b^2. Our aim in this section is to improve the well-known equivalent inequalities

$$H_X(bx) \leq e^{-x^2/2} \ \text{ for any } x > 0 \ \text{ or } \ Q_X(p) \leq b\sqrt{2|\log p|} \ \text{ for any } p \in]0,1[, \tag{6.3}$$

valid for any centered sub-Gaussian random variable X with parameter b. We refer to Boucheron et al. (2013, Section 2.3) for an introduction to sub-Gaussian random variables with a proof of (6.3) and to Bobkov et al. (2006) for estimates of the sub-Gaussian constant.

In order to improve (6.3), we consider here a slightly stronger condition on the moment-generating function.

Definition 6.1. *Let b be any positive real. A real-valued random variable X is said to be entropic sub-Gaussian with parameter b if X has a finite moment-generating function on $\mathbb{R}$ and*

$$t\ell'_X(t) - \ell_X(t) \leq b^2(t^2/2) \quad \text{for any} \quad t > 0.$$

We denote the collection of such random variables by $\mathcal{G}_\mathcal{E}(b)$.

Remark 6.1. *X belongs to $\mathcal{G}_\mathcal{E}(b)$ if and only if (X/b) belongs to $\mathcal{G}_\mathcal{E}(1)$.*

If X belongs to $\mathcal{G}_\mathcal{E}(b)$, then $X - \mathbb{E}(X)$ satisfies (6.2) with the same parameter b (*see* Ledoux (1996), pages 69-70). However the class $\mathcal{G}_\mathcal{E}(b)$ does not contain all the sub-Gaussian random variables with parameter b, and thus, there is some hope to improve (6.2) for entropic sub-Gaussian random variables with parameter b. Theorem 6.1 below is a progress in this direction.

Theorem 6.1. *Let p be any real in $]0,1[$. Set $v = p/(1-p)$. Then*

$$\sup_{X \in \mathcal{G}_\mathcal{E}(1)} \big(\tilde{Q}_X(p) - \mathbb{E}(X)\big) = \sqrt{1/v} = \sqrt{(1/p) - 1} \quad \textit{for any } p \geq 1/2. \tag{a}$$

Let L_v^* be defined by (3.10). Then, for any $p < 1/2$,

$$\sup_{X \in \mathcal{G}_\mathcal{E}(1)} \big(\tilde{Q}_X(p) - \mathbb{E}(X)\big) \leq \inf_{x \in]0,1[} \frac{L_v^*(x) + \log(1 + x/v)}{\sqrt{2L_v^*(x)}} \leq \Big(\frac{2|\log v|}{1 - v^2}\Big)^{1/2}. \tag{b}$$

Furthermore the above upper bound is strictly less than $\sqrt{\min(1/v, 2|\log p|)}$.

Applying (1.4), we immediately derive from Theorem 6.1 the corollary below for sub-Gaussian martingales.

Corollary 6.1. *Let M_n be a submartingale in L^1. Suppose that the random variable M_n is entropic sub-Gaussian with parameter b. Let p be any real in $]0,1/2[$. Set $v = p/(1-p)$. Then*

$$b^{-1}Q_{M_n^*}(p) \leq \inf_{x \in]0,1[} \frac{L_v^*(x) + \log(1 + x/v)}{\sqrt{2L_v^*(x)}} \leq \Big(\frac{2|\log v|}{1 - v^2}\Big)^{1/2} < \sqrt{min(1/v, 2|\log p|)}\,.$$

Proof of Theorem 6.1. We start by proving (b). Let X be any random variable in the class $\mathcal{G}_{\mathcal{E}}(1)$ and λ be any positive real. Define the random variable Y_λ from X by

$$Y_\lambda = \exp(\lambda X - \ell_X(\lambda)). \tag{6.4}$$

By the Jensen inequality applied to the convex function $x \mapsto e^{\lambda x}$,

$$\exp\big(\lambda \tilde{Q}_X(p)\big) \leq p^{-1} \int_0^p \exp\big(\lambda Q_X(s)\big) ds,$$

which is equivalent to

$$\tilde{Q}_X(p) \leq \lambda^{-1}\big(\ell_X(\lambda) + \log \tilde{Q}_{Y_\lambda}(p)\big). \tag{6.5}$$

By definition $\mathbb{E}(Y_\lambda) = 1$. Hence, we may apply Theorem 3.1(a) applied with $z = 1/p$ to Y_λ. Using also (3.15), we then get that

$$\tilde{Q}_{Y_\lambda}(p) \leq 1 + v^{-1} L_v^{*-1}(H_\lambda) \ \text{ where } H_\lambda = \mathcal{H}(Y_\lambda) \ \text{ and } \ v = p/(1-p). \tag{6.6}$$

Now

$$H_\lambda = \mathbb{E}\big((\lambda X - \ell_X(\lambda)) e^{\lambda X - \ell_X(\lambda)}\big) = \lambda \ell_X'(\lambda) - \ell(\lambda). \tag{6.7}$$

Since X is entropic sub-Gaussian with parameter 1, it follows that $H_\lambda \leq \lambda^2/2$. Hence, from (6.6) and the monotonicity of L_v^{*-1},

$$\tilde{Q}_{Y_\lambda}(p) \leq 1 + v^{-1} L_v^{*-1}(\lambda^2/2). \tag{6.8}$$

Combining the above inequality, (6.5) and the fact that an entropic sub-Gaussian random variable is sub-Gaussian with the same parameter, we get that, for any positive λ,

$$\tilde{Q}_X(p) \leq \lambda^{-1}\big(\lambda^2/2 + \log\big(1 + v^{-1} L_v^{*-1}(\lambda^2/2)\big)\big), \tag{6.9}$$

Let x be any real in $]0, 1[$. Taking $\lambda = \sqrt{2L_v^*(x)}$ in the above inequality, we obtain

$$\tilde{Q}_X(p) \leq (2L_v^*(x))^{-1/2}\big(L_v^*(x) + \log(1 + x/v)\big) := \varphi(x). \tag{6.10}$$

Since this upper bound is valid for any x in $]0, 1[$, it implies the first part of (b). Now, if $p < 1/2$, $v = p/(1-p) < 1$. Therefore, we can choose $x = 1 - v$ in (6.10). For this choice of x,

$$\log(1 + x/v) = |\log v| = -\log v = -\log(1 - x).$$

Therefrom $L_v^*(1 - v) = (1 + v)^{-1}(1 - v)|\log v|$ and

$$\varphi(1 - v) = \sqrt{2|\log v|/(1 - v^2)}, \tag{6.11}$$

which gives the second part of (b).

We now prove that

$$2|\log v|/(1 - v^2) < \min(1/v, 2\log(1 + 1/v)), \tag{6.12}$$

which implies the last statement of Theorem 6.1, since $(1/p) = 1 + (1/v)$. First

$$|\log v| = \sum_{k>0} \frac{(1-v)^k}{k} < (1 - v) + \frac{(1-v)^2}{2} \sum_{j\geq 0} (1 - v)^j = \frac{1 - v^2}{2v}.$$

This inequality ensures that $2|\log v|/(1-v^2) < 1/v$. And second, starting from the inequality $(1-v)h(v) + vh(v-1) > 0$, we get that

$$(1-v^2)\log(1+v) + v^2 \log v > 0, \quad \text{or, equivalently } (1-v^2)\log(1+1/v) + \log v > 0.$$

Therefrom $|\log v|/(1-v^2) < \log(1+1/v)$, which ends up the proof of (6.12).

We now prove (a). If X is entropic sub-Gaussian with parameter 1, then the variance of X is less than 1. Consequently, by Theorem 4.1 applied with $p=2$, $\sigma=1$ and $z=1/p$,

$$\tilde{Q}_X(p) - \mathbb{E}(X) \le \sqrt{(1/p)-1} = \sqrt{1/v}.$$

It remains to prove that there exist some random variable X, entropic sub-Gaussian with parameter 1 and fulfilling the equality in (a) of Theorem 6.1. To prove this fact, we will use the lemma below.

Lemma 6.1. *For any $p \ge 1/2$, the Bernoulli law $b(p)$ is entropic sub-Gaussian with parameter $\sqrt{p(1-p)}$.*

Proof of Lemma 6.1. We start by noticing that, for any random variable X with finite Laplace transform $(t\ell_X' - \ell_X)'(t) = t\ell_X''(t)$. Therefrom, if $\ell_X''(t) \le b^2$ for any positive t, then X is entropic sub-Gaussian with parameter b.

Now let X be a random variable with law $b(p)$. Then $\ell_X(t) = \log(1-p+pe^t)$ and

$$\ell_X''(t) = p(1-p)e^t(1-p+pe^t)^{-2} = p(1-p)\big((1-p)e^{-t/2} + pe^{t/2}\big)^{-2}.$$

Next $(1-p)e^{-t/2} + pe^{t/2} = \cosh(t/2) + (2p-1)\sinh(t/2) \ge 1$ for any $p \ge 1/2$ and any positive t. Hence $\ell_X''(t) \le p(1-p)$ for any positive t, which implies Lemma 6.1. $\diamond$

We now complete the proof of Theorem 6.1(a). Let U be a random variable with uniform law over $[0,1]$. Let $p \ge 1/2$. Set $X = (\,p(1-p)\,)^{-1/2}\,\mathbf{1}_{U\le p}$. From Lemma 6.1, the random variable X is entropic sub-Gaussian with parameter 1. Now

$$\tilde{Q}_X(p) = (\,p(1-p)\,)^{-1/2} \quad \text{and} \quad \mathbb{E}(X) = (\,p/(1-p)\,)^{1/2},$$

whence

$$\tilde{Q}_X(p) - \mathbb{E}(X) = (\,p(1-p)\,)^{-1/2}(1-p) = \sqrt{(1/p)-1}\,. \qquad \diamond$$

Numerical comparisons. Here we compare Proposition 5.2 and Corollary 6.1 with the usual inequalities in the case of weighted sums of symmetric random variables. Let $\eta_1, \eta_2, \dots$ be a sequence of independent and symmetric random variables with variance 1. Assume furthermore that the random variables $\eta_1, \eta_2, \dots$ are entropic sub-Gaussian with parameter 1, and sub-Gaussian at order 4. For example, it can easily be proven that this condition holds true if $|\eta_k| \le \sqrt{3}$ almost surely for any positive k. Let $a_1, a_2, \dots$ be a sequence of positive reals. Set

$$M_k = a_1\eta_1 + a_2\eta_2 + \cdots + a_k\eta_k \quad \text{for any } k > 0. \tag{6.13}$$

Then M_n is entropic sub Gaussian with parameter $b_n = \sqrt{V_n} = \sqrt{a_1^2 + \cdots + a_n^2}$. Define the function $\psi :]0,\infty[\mapsto]0,\infty[$ by

$$\psi(v) = \inf_{x\in]0,1[} \big(2L_v^*(x)\big)^{-1/2}\big(L_v^*(x) + \log(1+x/v)\big) \quad \text{for } v \le 1. \tag{6.14}$$

Let z be any real in $]1, \infty[$. Under the above conditions, Corollary 6.1 yields

$$b_n^{-1} Q_{M_n^*}(1/z) \leq \psi(1/(z-1)) \quad \text{for any } z \geq 2. \tag{6.15}$$

and Proposition 5.2 gives

$$b_n^{-1} Q_{M_n^*}(1/z) \leq \big(1 + \sqrt{(\min((z-2)/4, 1)(z-2)}\,\big)^{1/2} \quad \text{for any } z \geq 2. \tag{6.16}$$

Now the usual bound (6.3) yields

$$b_n^{-1} Q_{M_n^*}(1/z) \leq \sqrt{2 \log z}, \quad \text{for any } z \geq 1. \tag{6.17}$$

Below I give the numerical values of these upper bounds for some integer values of z.

Ineq.	z=2	z= 4	z=6	z= 10	z= 20	z=30	z=40	z=50
(6.15)	1.00	1.55	1.80	2.07	2.39	2.56	2.67	2.75
(6.16)	1.00	1.41	1.73	1.96	2.29	2.51	2.68	2.82
(6.17)	1.18	1.67	1.89	2.15	2.45	2.61	2.72	2.80

It appears here that (6.15) is more efficient as soon as $z \geq 40$. However (6.16) provides better estimates for $z < 40$.

7 A more efficient bounded differences inequality

Ths section is devoted to the bounded differences inequality, sometimes called McDiarmid inequality (*see* McDiarmid (1989), Corollary 6.10). Let $E^n = E_1 \times E_2 \cdots \times E_n$ and let $X = (X_1, \dots X_n)$ be a random vector in E^n with independent components. Let $f : E^n \mapsto \mathbb{R}$ be a bounded measurable function. For all $1 \leq k \leq n$, denote by $\mathcal{F}^{(k)}$ the σ-algebra generated by $X_1, \dots, X_n$ except X_k,

$$\mathcal{F}^{(k)} = \sigma(X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n).$$

Assume that for each $1 \leq k \leq n$, there exist two $\mathcal{F}^{(k)}$-measurable bounded random variables A_k and B_k such that

$$A_k \leq f(X) \leq B_k \qquad \text{almost surely.} \tag{7.1}$$

Then, for any positive x,

$$\mathbb{P}(f(X) \geq \mathbb{E}(f(X)) + x) \leq \exp\Big(-\frac{2x^2}{C_n}\Big), \quad \text{where } C_n = \sum_{k=1}^n \|B_k - A_k\|_\infty^2. \tag{7.2}$$

This inequality is often called bounded differences inequality.

We now recall an improvement of this inequality, due to Bercu et al. (2015): instead of assuming a uniform bound on each oscillation, they only assume a bound on the sum of squares. If $Z = f(X)$, by Theorem 2.62 in Bercu et al. (2015), for any positive x,

$$\mathbb{P}(Z \geq \mathbb{E}(Z) + x) \leq \exp\big(-2x^2/D_n\big), \quad \text{where } D_n = \Big\| \sum_{k=1}^n (B_k - A_k)^2 \Big\|_\infty. \tag{7.3}$$

Of course, this inequality is equivalent to the quantile inequality

$$Q_Z(p) \leq \mathbb{E}(Z) + \sqrt{D_n |\log p|/2} \quad \text{for any } p \in]0, 1[. \tag{7.4}$$

The proof of the above inequality is based on the entropy method, which has been widely developed by Ledoux (1996). In particular Bercu et al. (2015, page 56) prove that the random variable Z is entropic sub-Gaussian with parameter $\sqrt{D_n/4}$. Consequently Theorem 6.1 yields the new more efficient concentration inequality below.

Theorem 7.1. *Let Z and D_n be defined as in (7.3) and (7.4). Let p be any real in $]0,1[$. Set $v = p/(1-p)$. Then, under the conditions of Inequality (7.4),*

$$\tilde{Q}_Z(p) - \mathbb{E}(Z) \le \sqrt{D_n/4v}, \quad \textit{for any } p \ge 1/2, \tag{a}$$

and, for any $p < 1/2$,

$$\tilde{Q}_Z(p) - \mathbb{E}(Z) \le \frac{\sqrt{D_n}}{2} \inf_{x \in]0,1[} \frac{L_v^*(x) + \log(1 + x/v)}{\sqrt{2L_v^*(x)}} \le \left(\frac{D_n |\log v|}{2(1-v^2)} \right)^{1/2}. \tag{b}$$

Numerical comparisons. Let $n \ge 2$. Suppose that $\mathbb{E}(Z) = 0$ and $C_n = 4$. Define the function $\psi :]0,\infty[\mapsto]0,\infty[$ by $\psi(v) = \sqrt{1/v}$ for $v \ge 1$ and

$$\psi(v) = \inf_{x \in]0,1[} \big(2L_v^*(x)\big)^{-1/2} \big(L_v^*(x) + \log(1 + x/v)\big) \text{ for } v < 1. \tag{7.5}$$

Let z be any real in $]1,\infty[$. Since $D_n \le C_n$, Theorem 7.1 implies that

$$Q_Z(1/z) \le \tilde{Q}_Z(1/z) \le \psi(1/(z-1)). \tag{7.6}$$

By the usual bounded differences inequality (7.2),

$$Q_Z(1/z) \le \sqrt{2|\log z|}\,. \tag{7.7}$$

Now, let Y be a standard normal. By Inequality (2.8) in Pinelis (2006),

$$Q_Z(1/z) \le Q_Y(1/(c_{5,0} z)) \text{ where } c_{5,0} = 5!\,(e/5)^5 = 5.699... \tag{7.8}$$

From Remark 2.4 in Pinelis (2006), (7.8) is more efficient than (7.7) for $z \ge 5.96...$

Let us now recall some known lower bounds. Let $\Delta_k = \|B_k - A_k\|_\infty$. For any m in $[1,n]$, From Proposition 5.7 in Owhadi et al. (2013) applied with $\Delta_1 = \cdots = \Delta_m = 2m^{-1/2}$ and $\Delta_k = 0$ for $k > m$, there exists a centered random variable Z_m satisfying the conditions of Theorem 7.1, such that

$$Q_{Z_m}(1/z) \ge 2\sqrt{m}\big(1 - z^{-1/m}\big) \text{ for any } z > 1.$$

Consequently, for any $z > 1$, there exists a random variable Z satisfying the conditions of Theorem 7.1, such that

$$Q_Z(1/z) \ge \max\big(2(1 - 1/z), 2\sqrt{2}(1 - z^{-1/2})\big). \tag{7.9}$$

Below I give the numerical values of the above upper and lower bounds for integer values of z, including the median, the quartile and the decile.

Ineq.	z=2	z=4	z=6	z=8	z=10	z=12	z=14	z=16	z=20
(7.6)	1.00	1.55	1.80	1.96	2.07	2.16	2.23	2.29	2.39
(7.7)	1.18	1.67	1.89	2.04	2.15	2.23	2.30	2.35	2.45
(7.8)	1.35	1.71	1.89	2.02	2.11	2.18	2.24	2.29	2.37
(7.9)	1.00	1.50	1.67	1.83	1.93	2.01	2.07	2.12	2.20

One can see that (7.8) is better than (7.6) for $z = 20$ and almost equivalent for $z = 16$. However D_n is often strictly less than C_n.

8 Annex: a one-sided von Bahr-Esseen inequality

As in Pinelis (2015), we introduce a class of generalized moment functions, including the power functions. However the functions we consider vanish on $]-\infty, 0]$. So, let

$$\mathcal{F} = \{ f \in C^1(\mathbb{R}) : f_{|\mathbb{R}^-} = 0, \ f' \text{ increasing and concave on } \mathbb{R}^+, \ \lim_{+\infty} f' = \infty \}. \tag{8.1}$$

Let $x_+ = \max(0, x)$ for any real x. The one-sided power functions $x \mapsto x_+^p$ belong to $\mathcal{F}$ for p in $]1, 2]$.

Proposition 8.1. *Let $(M_k)_{0 \le k \le n}$ be a real-valued martingale in L^1. Set $X_k = M_k - M_{k-1}$. Then, for any f in $\mathcal{F}$,*

$$\mathbb{E}\big(f(M_n)\big) \le \mathbb{E}\big(f(M_0) + f(|X_1|) + \cdots + f(|X_n|)\big). \tag{a}$$

In particular, for any p in $]1, 2]$, any $t \ge 0$ and any martingale $(M_k)_{0 \le k \le n}$ in L^p such that $M_0 = 0$,

$$\mathbb{E}\big((M_n + t)_+^p\big) \le t^p + \mathbb{E}\big(|X_1|^p + \cdots + |X_n|^p\big). \tag{b}$$

Proof. (b) follows from (a) applied to the martingale $(M_k + t)_{0 \le k \le n}$ and the function $f(x) = x_+^p$. We now prove (a). For $n > 0$, let

$$\Delta_n = f(M_n) - f(M_{n-1}) - f'(M_{n-1}) X_n. \tag{8.2}$$

From the martingale property,

$$\mathbb{E}(\Delta_n) = \mathbb{E}\big(f(M_n)\big) - \mathbb{E}\big(f(M_{n-1})\big). \tag{8.3}$$

Hence (a) will follow by induction on n if we prove that

$$\mathbb{E}(\Delta_n) \le \mathbb{E}\big(f(|X_n|)\big). \tag{8.4}$$

By the Taylor formula.

$$\Delta_n = \int_0^{X_n} (f'(M_{n-1} + s) - f'(M_{n-1})) ds \le \int_0^{|X_n|} \sup_{a \in \mathbb{R}} |f'(a+s) - f'(a)| \, ds.$$

Now f' is nondecreasing. Hence $|f'(a+s) - f'(a)| = f'(a+s) - f'(a)$ for any $s \ge 0$. Since f' is concave on $\mathbb{R}^+$, it follows that, if $s \ge 0$,

$$|f'(a+s) - f'(a)| \le f'(s) \ \text{ for any } a \ge 0.$$

For $a < 0$, $f'(a) = 0$, whence $|f'(a+s) - f'(a)| = f'(a+s) \le f'(s)$, using the monotonicity of f' again. In any cases $|f'(a+s) - f'(a)| \le f'(s)$ for $s \ge 0$. Consequently

$$\Delta_n \le \int_0^{|X_n|} f'(s) ds = f(|X_n|). \tag{8.5}$$

Taking then the expectation in (8.5), we get (8.4), which ends the proof.$\diamond$

Remark 8.1. *This result cannot be derived from the von Bahr-Esseen inequality for absolute moments of Pinelis (2015), since the constant for absolute moments is strictly larger than* 1.

References

[1] Bercu, B., Delyon, B. and Rio, E. Concentration inequalities for sums and martingales. SpringerBriefs in Mathematics. Springer, Cham, x+120 pp. (2015).

[2] Blackwell, D. and Dubins, L. A converse to the dominated convergence theorem. *Illinois J. Math.* **7**, 508-514 (1963).

[3] Bobkov, S. On deviations from medians, Manuscript (1998).

[4] Bobkov, S., Houdré, C. and Tetali, P. The subgaussian constant and concentration inequalities. *Israel J. Math.* **156**, 255-283 (2006).

[5] Boucheron, S., Lugosi, G. and Massart, P. Concentration inequalities. A nonasymptotic theory of independence. Oxford University Press, Oxford, x+481 pp. (2013).

[6] Cantelli, F. P. Sui confini della probabilità. *Atti Congresso Bologna 1928* **6**, 47-59 (1932).

[7] Doob, J. L. Regularity properties of certain families of chance variables. *Trans. Amer. Math. Soc.* **47**, 455-486 (1940).

[8] Doob, J. L. Stochastic processes. John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, viii+654 pp. (1953)

[9] Dubins, L. and Gilat, D. On the distribution of maxima of martingales. *Proc. Amer. Math. Soc.* **68**, no. 3, 337-338 (1978).

[10] Gilat, D. The best bound in the $L \log L$ inequality of Hardy and Littlewood and its martingale counterpart. *Proc. Amer. Math. Soc.* **97**, no. 3, 429-436 (1986).

[11] Gilat, D. and Meilijson, I. A simple proof of a theorem of Blackwell & Dubins on the maximum of a uniformly integrable martingale. *Séminaire de Probabilités* XXII, 214-216, Lecture Notes in Math. **1321**, Springer, Berlin (1988).

[12] Hardy, G. and Littlewood, J. A maximal theorem with function-theoretic applications. *Acta Math.* **54**, no. 1, 81-116 (1930).

[13] Harremoës, P. Some new maximal inequalities. *Statist. Probab. Lett.* **78**, no. 16, 2776-2780 (2008).

[14] Hoeffding, W. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13-30 (1963).

[15] Kearns, M. J. and Saul, L. K. Large deviation methods for approximate probabilistic inference. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, San-Francisco, 311-319 (1998)

[16] Ledoux, M. On Talagrand's deviation inequalities for product measures. *ESAIM Probab. Statist.* **1**, 63-87 (1996).

[17] McDiarmid, C. On the method of bounded differences. Surveys in combinatorics, 1989 (Norwich, 1989), *London Math. Soc. Lecture Note Ser.* **141**, 148-188 (1989).

[18] Owhadi, H., Scovel, C., Sullivan, T., McKerns, M., and Ortiz, M. Optimal uncertainty quantification. *SIAM Review, 2013* **55**, no 2, 271-345 (2013).

[19] Osekowski, A. Sharp martingale and semimartingale inequalities. *Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series)* **72**. Birkhäuser/Springer (2012).

[20] Pinelis, I. On normal domination of (super)martingales. *Electron. J. Probab.* **11**, 1049-1070 (2006).

[21] Pinelis, I. Best possible bounds of the von Bahr-Esseen type. *Ann. Funct. Anal.* **6**, no 4, 1-29 (2015).

[22] Rio, E. About the constants in the Fuk-Nagaev inequalities. *Electron. Commun. Probab.* **22**, (28), 1-12 (2017).

[23] Savage, I. R. Probability inequalities of the Tchebycheff type. *Journal of Research of the National Bureau of Standards-B. Mathematics and Mathematical Physics B*, **65**, no 3, 211-222 (1961).

[24] Tchebichef, P. Sur les valeurs limites des intégrales. *Journal de mathématiques pures et appliquées, 2ème série,* **19**, 157-160 (1874).

[25] Ville, J. Etude critique de la notion de collectif. *Monographies des probabilités, publiées par E. Borel,* fascicule no. 3. Gauthier-Villars, Paris, 144 pp. (1939).