Electron. Commun. Probab. **23** (2018), no. 78, 1–12.
https://doi.org/10.1214/18-ECP178
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

About Doob's inequality, entropy and Tchebichef

Emmanuel Rio*

Abstract

In this note we give upper bounds on 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.

Keywords: Doob's inequality; Hardy-Littlewood maximal function; $L\log L$; entropy; binomial rate function; covariance inequalities; Cantelli's inequality; subGaussian random variables; bounded differences inequality; McDiarmid's inequality; conditional value at risk.
AMS MSC 2010: Primary 60E15, Secondary 60G42.
Submitted to ECP on December 6, 2017, final version accepted on October 7, 2018.

1 Introduction

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

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 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 $[0,1]$.

Let us now recall Doob's maximal inequalities. Below we assume that the random variables $M_0, M_1, \ldots, M_n$ are nonnegative. The first inequality is in fact due to Ville ([19], Theorem 1, page 100): for any $x > 0$,

$$\mathbb{P}(M_n^* > x) \le x^{-1}\mathbb{E}(M_n) \text{ or, equivalently, } Q_{M_n^*}(1/z) \le z\mathbb{E}(M_n) \text{ for any } z > 1. \tag{1.1}$$

Doob ([6], Theorem 1.1) proved the more precise inequality

$$\mathbb{P}(M_n^* \ge x) \le x^{-1}\mathbb{E}\big(M_n \mathbf{1}_{M_n^*\ge x}\big) \text{ for any } x > 0. \tag{1.2}$$

*UMR 8100 CNRS, Laboratoire de Mathématiques de Versailles, France. E-mail: emmanuel.rio@uvsq.fr

Define now the nonincreasing function $\tilde{Q}_X$ by

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

Clearly $Q_X \leq \tilde{Q}_X$. Notice that, since $Q_X(U)$ has the same law as X, $\tilde{Q}_X(1) = \mathbb{E}(X)$. For a nonnegative random variable X, $\tilde{Q}_X$ is the maximal function associated with X introduced by Hardy and Littlewood [10]. In mathematical finance, $\tilde{Q}_X$ is called conditional value at risk (CVaR) of X. Blackwell and Dubins [2] derived from (1.2) the upper bound

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

Dubins and Gilat [7] proved the optimality of (1.4). Later Gilat and Meilijson [9] proved that the nonnegativity assumption on $(M_k)_{0 \leq k \leq n}$ can be dropped in (1.4). As said by Gilat [8], “Thus a complete equivalence is established between Doob’s maximal martingale inequalities and the corresponding results of Hardy and Littlewood (*on the CVaR of* X)”. Hence, from now on, the main focus will be on optimal upper bounds on the CVaR of a random variable X with given expectation, satisfying an additional moment condition.

Let us start by considering the class $L \log L$ of nonnegative random variables X such that $\mathbb{E}(X \log^+ X) < \infty$, where $\log^+ x = \max(0, \log x)$. Up to now, upper bounds on $\mathbb{E}(\tilde{Q}_X(U))$ have received more attention than upper bounds on $\tilde{Q}_X(u)$. Gilat ([8], Theorem 3) proved that the inequality

$$\mathbb{E}(\tilde{Q}_X(U)) \leq c\mathbb{E}(X \log X) + d \tag{1.5}$$

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 $d = e/(e-1)$. The martingale counterpart of (1.5) may be found in Osekowski ([14], Theorem 7.7).

Starting from (1.2) and introducing the entropy of M_n, Harremoës [11] improved the martingale counterpart of (1.5). 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 $(M_k)_{0 \leq k \leq n}$ is a nonnegative martingale such that $M_0 = 1$, Harremoës [11] derived from (1.2) the upper bound

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

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

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

Harremoës ([11], Theorem 4) also proved that (1.8) is tight. It appears here that the entropy is the adequate quantity in the class $L \log L$. Therefore, in Section 2, we give an elementary covariance inequality involving entropy. This covariance inequality is applied in Section 3. Particularly, Theorem 3.1, which is the main result of this note, provides a sharp upper bound on the CVaR of a nonnegative random variable in the class $L \log L$. An interesting application is that (1.1) can be improved only if $\log z > \mathcal{H}(M_n)/\mathbb{E}(M_n)$ (*see* Remark 3.2 in Section 3). In Section 5, we apply Theorem 3.1 to the class of subGaussian random variables introduced (in spirit) by Ledoux [13]: we say that a centered real-valued random variable Y with finite Laplace transform on $\mathbf{R}_+$ is entropic subGaussian on the right with parameter 1 if $\mathcal{H}(e^{tY})/\mathbb{E}(e^{tY}) \le t^2/2$ for any positive t. Using Theorem 3.1, we obtain sharper bounds on $Q_Y(u)$ than the usual bound

$$Q_X(u) \le \min\bigl(\sqrt{2|\log u|}, \sqrt{(1/u) - 1}\,\bigr), \tag{1.9}$$

valid for any centered random variable X such that $\log \mathbb{E}(e^{tX}) \le t^2/2$ for any positive t. However (1.9) is suboptimal and has not yet been improved. So, in Section 5, we also give the exact upper bound on $Q_X(u)$ for a random variable X fulfilling the above condition and we prove that our bound on $Q_Y(u)$ remains sharper than the exact bound on $Q_X(u)$. In Sections 6 and 7, we apply these results to martingales with bounded increments and functions of independent random variables.

Assume now that the submartingale $(M_k)_{0\le k\le n}$ is in L^p for some $p > 1$. For any real y, let $y_+ = \max(0, y)$. By (1.1) applied to the nonnegative submartingale $(M_k - \mathbb{E}(M_n)\,)_+^p$,

$$\mathbb{P}(M_n^* \ge \mathbb{E}(M_n) + x) \le x^{-p}\mathbb{E}\bigl((M_n - \mathbb{E}(M_n)\,)_+^p\,\bigr) \quad \text{for any } x > 0. \tag{1.10}$$

However this upper bound tends to ∞ as $x \searrow 0$. Recall now that, for any real-valued random variable X in L^2 and any positive x,

$$\mathbb{P}(X \ge \mathbb{E}(X) + x) \le \sigma^2/(x^2 + \sigma^2), \quad \text{where } \sigma^2 = \operatorname{Var} X.$$

The above inequality, often called Tchebichef-Cantelli inequality (*see* Tchebichef [17], pages 159–160 and Cantelli [5], Inequality (19), p. 53) is efficient for any $x > 0$ and equivalent to the upper bound

$$Q_X(u) \le \mathbb{E}(X) + \sigma\sqrt{(1/u) - 1} \quad \text{for any } u \in]0, 1[. \tag{1.11}$$

Wald [20] tried to generalize the Tchebichef-Cantelli inequality in the case $p \ne 2$ to a nonnegative random variable X in L^p with prescribed expectation and L^p-norm. However he obtained only an implicit upper bound on the tail of X. In Section 4, we will obtain a sharp and explicit extension of (1.11) to the CVaR of a random variable X in L^p for $p > 1$. We emphasize that, in contrast to the tail inequalities, quantile inequalities support explicit extensions to L^p for arbitrary p.

2 A covariance inequality involving entropy

Throughout this section Y is a nonnegative real-valued random variable. We assume that $\mathbb{E}(Y \log^+ Y) < \infty$ and $\mathbb{E}(Y) > 0$.

Theorem 2.1. Let Y 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

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

Proof. Define the two-parameter family of functions $\varphi_{a,b}$ by $\varphi_{a,b}(x) = (x/b)\log(x/a)$ for any $x \ge 0$ and any positive reals a and b, with the convention $0 \log 0 = 0$. Clearly

$$Y\eta \le \varphi_{a,b}(Y) + \varphi_{a,b}^*(\eta), \quad \text{where } \varphi_{a,b}^*(y) = \sup\{xy - \varphi_{a,b}(x) : x \in [0, \infty[\,\}. \tag{2.1}$$

Next the function $x \mapsto xy - \varphi_{a,b}(x)$ takes its maximum at point $x = ae^{by} - 1$, from which $\varphi_{a,b}^*(y) = (a/b)\exp(yb - 1)$. It follows that

$$Y\eta \le b^{-1}\bigl(Y \log Y - Y \log a + a \exp(b\eta - 1)\,\bigr). \tag{2.2}$$

Taking the expectation in the above inequality,

$$\mathbb{E}(Y\eta) \le b^{-1}\bigl(\mathbb{E}(Y \log Y) - \mathbb{E}(Y) \log a + (a/e)\mathbb{E}\bigl(e^{b\eta}\bigr)\,\bigr). \tag{2.3}$$

Let us now minimize the upper bound. Differentiating the upper bound with respect to a, we get that the optimal value of a is $a = e\mathbb{E}(Y/\mathbb{E}(e^{b\eta})$. Choosing this value in (2.3), we get that

$$\mathbb{E}(Y\eta) \le b^{-1}\big(\mathbb{E}(Y\log Y) - \mathbb{E}(Y)\log\mathbb{E}(Y) + \mathbb{E}(Y)\log\mathbb{E}\big(e^{b\eta}\big)\,\big) \text{ for any } b > 0, \tag{2.4}$$

which implies Theorem 2.1. □

Remark 2.1. Applying Theorem 2.1 to $Y = Q_{M_n}(U)$ and $\eta = \log(1/U)$, one can prove (1.8) for any nonnegative submartingale $(M_k)_{0\le k\le n}$ in $L\log L$ such that $\mathbb{E}(M_n) = 1$. (see Rio [16], Remark 2.2). Notice that the proof of (1.5) in Gilat [8] is based on the inequality $Q_X(U)\eta \le \varphi_{1,b}(Q_X(U)) + \varphi^*_{1,b}(\eta)$, where $b = 1/c$. The minimization with respect to a is omitted, which leads to a suboptimal inequality.

3 Bounds on the CVaR involving entropy

The main result of this section is the optimal upper bound below on conditional value at risk of a nonnegative random variable X. We refer to Rio [16] for applications of Theorem 3.1 to upper bounds on the tail of the maximum of a nonnegative submartingale in the class $L\log L$.

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 $z > 1$

$$\tilde Q_X(1/z) \le \psi_H(z) \text{ where } \psi_H(z) = \inf_{t>0} t^{-1}\big(H - \log z + \log\big(e^{zt} + z - 1\big)\,\big). \tag{a}$$

An other formulation of ψ_H is

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

Furthermore

$$\psi_H(z) = z \text{ for any } z \le e^H \text{ and } \psi_H(z) < z \text{ for any } z > e^H. \tag{c}$$

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

$$\mathbb{E}(Y) = 1,\ \mathcal{H}(Y) = H \text{ and } \tilde Q_Y(1/z) = \psi_H(z). \tag{d}$$

Proof. By Theorem 2.1 applied to $Y = Q_X(U)$ and $\eta = z\,\mathbf{1}_{zU\le 1}$,

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

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 $H \ge \log z$,

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

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

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

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

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

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

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

Set

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

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

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

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. □

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 $\mathcal{H}(X) < \infty$,

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

Remark 3.2. From (1.4) and the above Remark, Theorem 3.1 applied to any positive submartingale $(M_k)_{0 \le k \le n}$ in the class $L \log L$ yields

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

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 (1.1). If $\mathbb{E}(M_n)\log z \le \mathcal{H}(M_n)$, then $\psi_H(z) = z$ and (3.7) does not improve (1.1).

Remark 3.3. Let μ be any law on $[0, \infty[$ with finite entropy. From Lemma 2 in Dubins and Gilat [7], there exists a nonnegative continuous time martingale $(M_t)_{t \in [0,1]}$ such that M_1 has the law μ and $M_1^* = \sup\{M_t : t \in [0,1]\}$ satisfies $Q_{M_1^*} = \tilde{Q}_{M_1}$. Hence Theorem 3.1 provides an optimal upper bound for continuous time martingales, which shows that (1.1) cannot be improved if $z \le e^H$.

4 Tchebichef type inequalities

In this section we give an extension of the Tchebichef-Cantelli inequality to random variables 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). Set $\sigma_p = \|X - \mathbb{E}(X)\|_p$. Then

$$\tilde{Q}_X(1/z) \le \mathbb{E}(X) + \sigma_p z^{1/p}(1 + (z-1)^{1-p})^{-1/p} \text{ for any } z > 1. \tag{a}$$

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

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

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, which implies (1.11).

Applying (1.4), we immediately get the corollary below. We refer to Rio ([16], Sections 4 and 5) for applications of this corollary.

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

$$Q_{M_n^*}(1/z) \le \mathbb{E}(M_n) + z^{1/p}(1 + (z-1)^{1-p})^{-1/p}\|M_n - \mathbb{E}(M_n)\|_p \text{ for any } z > 1.$$

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 $[0,1]$,

$$u\,\tilde{Q}_X(u) = u\,\tilde{Q}_X(u) - b\,\tilde{Q}_X(1) = \int_0^1 Q_X(s)\big(\mathbf{1}_{s\le u} - b\big)ds. \tag{4.1}$$

Applying then the Hölder inequality on $[0,1]$ with exponents p and $q = p/(p-1)$ to Q_X and $\mathbf{1}_{[0,u]} - b$, we get that $u\,\tilde{Q}_X(u) \le \sigma_p \big(u(1-b)^q + (1-u)b^q\big)^{1/q}$ or, equivalently,

$$\tilde{Q}_X(u) \le \sigma_p z^{1/p}\big((1-b)^q + (z-1)b^q\big)^{1/q}. \tag{4.2}$$

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 $b_0 = 1/\big(1 + (z-1)^{p-1}\big)$ is the unique critical point. Setting $b = b_0$ in (4.2), we then get that

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

which gives Theorem 4.1(a).

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

$$\mathbb{P}\big(X = z^{\frac{1}{p}}(1+(z-1)^{1-p})^{-\frac{1}{p}}\big) = 1/z = 1 - \mathbb{P}\big(X = -(z-1)^{-1}z^{\frac{1}{p}}(1+(z-1)^{1-p})^{-\frac{1}{p}}\big). \tag{4.3}$$

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

$$\mathbb{E}\big(|X|^p\big) = \big((1/z)z + (1-1/z)(z-1)^{-p}\,z\big)(1 + (z-1)^{1-p})^{-1} = 1. \qquad \square$$

Numerical comparisons. Below we compare Corollary 4.1 with the usual upper bound

$$Q_{M_n^*}(1/z) \le \mathbb{E}(M_n) + z^{1/p}\|M_n - \mathbb{E}(M_n)\|_p, \tag{4.4}$$

which can be immediately derived from (1.10), in the case $p = 3/2$, $\mathbb{E}(M_n) = 0$ and $\|M_n - \mathbb{E}(M_n)\|_p = 1$. We consider here values z of statistical interest. One can see that Corollary 4.1 is significantly better than (4.4) even for large values of z.

Ineq.	z=2	z=4	z=10	z=20	z=40	z=100	z=200
(4.4)	1.59	2.52	4.64	7.37	11.70	21.54	34.20
Cor. 4.1	1.00	1.86	3.83	6.42	10.59	20.21	32.67

5 SubGaussian random variables

In this section, for any real-valued random variable X we denote by ℓ_X the logarithm of the Laplace transform of X, defined by $\ell_X(t) = \log \mathbb{E}\big(e^{tX}\big)$ for any real t.

Definition 5.1. Let b be any positive real. The random variable X is said to be subGaussian on the right with parameter b if $X \in L^2$, X has a finite Laplace transform on $\mathbf{R}_+$ and

$$\ell_X(t) \leq t\,\mathbb{E}(X) + b^2(t^2/2) \ \text{ for any } \ t > 0. \tag{5.1}$$

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

The subGaussian constant is the smallest real b such that (5.1) holds. We refer to Bobkov et al [3] for estimates of this constant on some examples. We now introduce a slightly stronger condition on the Laplace transform.

Definition 5.2. Let b be any positive real. A real-valued random variable X is said to be entropic subGaussian on the right with parameter b if $X \in L^2$, X has a finite Laplace transform on $\mathbf{R}_+$ and

$$t\ell_X'(t) - \ell_X(t) \leq b^2(t^2/2) \ \text{ for any } \ t > 0. \tag{5.2}$$

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

If X belongs to $\mathcal{G}_{\mathcal{E}}(b)$, then X satisfies (5.1) with the same parameter b (*see* Ledoux [13], pages 69–70), which implies that $\mathcal{G}_{\mathcal{E}}(b) \subset \mathcal{G}(b)$. However $\mathcal{G}_{\mathcal{E}}(b)$ and $\mathcal{G}(b)$ are not equal. Thus, there is some hope to get better bounds for the quantiles of entropic subGaussian random variables. Nevertheless, under the median the bounds are exactly the same, as shown by Proposition 5.1 below.

Proposition 5.1. For any p in $[1/2, 1[$,

$$\sup_{X \in \mathcal{G}_{\mathcal{E}}(b)} \big(Q_X(p) - \mathbb{E}(X)\big) = \sup_{X \in \mathcal{G}(b)} \big(\tilde{Q}_X(p) - \mathbb{E}(X)\big) = b\sqrt{(1/p) - 1}.$$

Proposition 5.1 also proves that the Tchebichef-Cantelli inequality cannot be improved under the median, both in $\mathcal{G}_{\mathcal{E}}(b)$ and $\mathcal{G}(b)$. Before proving this proposition, we state the main result, which concerns the deviations over the median.

Theorem 5.1. Let p be any real in $]0, 1/2[$. Set $v = p/(1-p)$. Then

$$\sup_{X \in \mathcal{G}(b)} \big(Q_X(p) - \mathbb{E}(X)\big) = \sup_{X \in \mathcal{G}(b)} \big(\tilde{Q}_X(p) - \mathbb{E}(X)\big) = b\left(\frac{2|\log v|}{1 - v^2}\right)^{1/2}. \tag{a}$$

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

Define the function L_v^* by $L_v^*(y) = +\infty$ if $y > 1$ and

$$L_v^*(y) = \left(\frac{v + y}{v + 1}\right)\log\left(1 + \frac{y}{v}\right) + \left(\frac{1 - y}{v + 1}\right)\log(1 - y) \ \text{ if } y \in [0, 1].$$

Then, for any p in $]0, 1/2[$,

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

Notice that, from (1.4), one can immediately deduce upper bounds on the quantiles of the maximum of a subGaussian martingale. The statement is left to the reader.

Proof of Proposition 5.1. Notice that $\operatorname{Var} X \leq b^2$ for any X in $\mathcal{G}(b)$. Hence, from Theorem 4.1 and Remark 4.1,

$$\sup_{X\in\mathcal{G}(b)} \big(\tilde{Q}_X(p) - \mathbb{E}(X)\big) \leq b\sqrt{(1/p) - 1}. \tag{5.3}$$

Consequently, it suffices to prove that, for any p in $[1/2, 1[$,

$$\sup_{X\in\mathcal{G}_{\mathcal{E}}(b)} \big(Q_X(p) - \mathbb{E}(X)\big) \geq b\sqrt{(1/p) - 1}. \tag{5.4}$$

In order to prove (5.4), we will mainly use the lemma below.

Lemma 5.1. For any $p \geq 1/2$, the Bernoulli law $b(p)$ is entropic subGaussian with parameter $\sqrt{p(1-p)}$.

Proof of Lemma 5.1. 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 $\ell_X''(t) \leq b^2$ 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 $(1-p)e^{-t/2} + pe^{t/2} = \cosh(t/2) + (2p-1)\sinh(t/2) \geq 1$ for any $p \geq 1/2$ and any positive t. Hence $\ell_X''(t) \leq p(1-p)$ for any positive t, which implies Lemma 5.1. □

We now prove (5.4). Let U be a random variable with uniform law over $[0,1]$. Let q be any real in $]p, 1[$. Set $B_q = b(\,q(1-q)\,)^{-1/2}\,\mathbf{1}_{U\leq q}$. From Lemma 5.1, the random variable B_q is entropic sub-Gaussian with parameter b. Now

$$Q_{B_q}(p) = b(\,q(1-q)\,)^{-1/2} \text{ and } \mathbb{E}(B_q) = b(\,q/(1-q)\,)^{1/2},$$

whence

$$Q_{B_q}(q) - \mathbb{E}(B_q) = b(\,q(1-q)\,)^{-1/2}(1-q) = b\sqrt{(1/q) - 1}\,.$$

Now the right hand term in the above equality converges to $b\sqrt{(1/p) - 1}$ as $q \searrow p$, which completes the proof of (5.4). Finally (5.3) and (5.4) imply Proposition 5.1. □

Proof of Theorem 5.1. We need to introduce the Legendre-Fenchel dual. For a convex nondecreasing function $L : \mathbf{R}_+ \mapsto \mathbf{R}_+$, the Legendre-Fenchel dual L^* of L is defined by

$$L^*(x) = \sup\{xt - L(t) : t > 0\} \text{ for any } x \geq 0. \tag{5.5}$$

Clearly it is enough to prove Theorem 5.1 in the case $b = 1$. Let us prove (a). Notice first that, if X belongs to $\mathcal{G}(1)$, then, by the Jensen inequality, $\mathbb{E}(X \,|\, \mathcal{A})$ belongs to $\mathcal{G}(1)$ for any σ-field $\mathcal{A}$. Recall that, if U has the uniform distribution over $[0,1]$, $Q_X(U)$ has the same law as X. Now let $\mathcal{A}$ be the σ-field generated by the event $(U \leq p)$. Then $\beta_p = \mathbb{E}(Q_X(U) \,|\, \mathcal{A}) - \mathbb{E}(X)$ is a binary centered random variable in the class $\mathcal{G}(1)$. Furthermore $\mathbb{P}(\beta_p = \tilde{Q}_X(p) - \mathbb{E}(X)) = p$, which implies that β_p has the same law as $(\tilde{Q}_X(p) - \mathbb{E}(X))\xi_v$, where $v = p/(1-p)$ and ξ_v is the random variable defined by

$$\mathbb{P}(\xi_v = 1) = v/(1+v) \text{ and } \mathbb{P}(\xi_v = -v) = 1/(1+v). \tag{5.6}$$

Since $\tilde{Q}_X(p) - \mathbb{E}(X) \geq 0$, it follows that

$$\sup\{\tilde{Q}_X(p) - \mathbb{E}(X) \;:\; X \in \mathcal{G}_{\mathcal{E}}(1)\} = a_1, \text{ where } a_1 = \sup\{c \geq 0 : c\xi_v \in \mathcal{G}(1)\}. \tag{5.7}$$

Now, for a centered random variable X, define the subGaussian constant $C_{\mathcal{G}}(X)$ by

$$C_{\mathcal{G}}(X) = \sup\left\{ \left(2\ell_X(t)/t^2\right)^{1/2} \;:\; t > 0 \right\}. \tag{5.8}$$

By Lemma 2.22, page 25 in Bercu et al [1],

$$1/C_{\mathcal{G}}(X) = \inf\left\{ \left(2\ell_X^*(x)/x^2\right)^{1/2} \;:\; x > 0 \right\}. \tag{5.9}$$

We now apply this result to ξ_v. From Formula (2.55), page 29 in Bercu et al [1], $\ell^*_{\xi_v} = L^*_v$, where L^*_v is the function already defined in Theorem 5.1. Hence

$$1/C_{\mathcal{G}}(\xi_v) = \inf\left\{ \left(2L_v^*(x)/x^2\right)^{1/2} \;:\; x > 0 \right\} = \left(2|\log v|/(1-v^2)\right)^{1/2}, \tag{5.10}$$

according to Lemma 2.26, page 29 in Bercu et al [1] (*see* also Hoeffding [12] for a proof of this result). Consequently $a_1 = 1/C_{\mathcal{G}}(\xi_v)$. Now (5.7) implies the equality on the right in Theorem 5.1(a). Finally

$$\sup_{X\in\mathcal{G}(1)} \left(Q_X(p) - \mathbb{E}(X)\right) \geq \sup_{w>v}\left(1/C_{\mathcal{G}}(\xi_w)\right) = \left(2|\log v|/(1-v^2)\right)^{1/2},$$

which completes the proof of (a).

We now prove that, for $p < 1/2$,

$$2|\log v|/(1-v^2) < \min((1/p) - 1, 2\log(1/p)). \tag{5.11}$$

Since $v = p/(1-p)$, we have $v < 1$ and $(1/p) = 1 + (1/v)$. Now

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

This inequality ensures that

$$2|\log v|/(1-v^2) < 1/v \;\text{ for any } v < 1. \tag{5.12}$$

Next, the strict convexity of the function $x \mapsto x\log x$ ensures that

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

which completes the proof of (5.11).

It remains to prove (b). Let X be any random variable in the class $\mathcal{G}_{\mathcal{E}}(1)$ and λ be any positive real. Set

$$Y_\lambda = \exp(\lambda X - \ell_X(\lambda)). \tag{5.13}$$

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

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

which is equivalent to

$$\tilde{Q}_X(p) \leq \lambda^{-1}\left(\ell_X(\lambda) + \log \tilde{Q}_{Y_\lambda}(p)\right). \tag{5.14}$$

By definition $\mathbb{E}(Y_\lambda) = 1$. Hence, we may apply Theorem 3.1(a) with $z = 1/p$ to Y_λ:

$$\tilde{Q}_X(p) \leq \lambda^{-1}\left(\ell_X(\lambda) + \log \psi_{H_\lambda}(1/p)\right), \;\text{ where } H_\lambda = \mathcal{H}(Y_\lambda) \tag{5.15}$$

and ψ_H is the function already defined in Theorem 3.1(a). Let us now give a more tractable formula for $\psi_H(1/p)$. One can easily prove that, if $v = p/(1-p)$ and ξ_v is defined by (5.6), then

$$\psi_H(1/p) = 1 + v^{-1}\inf_{t>0} t^{-1}\left(H + \ell_{\xi_v}(t)\right). \tag{5.16}$$

Now the inversion formula below holds true (*see* Bercu et al [1], page 57): if X is a nondegenerate centered random variable with a finite Laplace transform on $[0,\infty[$,

$$\ell_X^{*-1}(x) = \inf\{t^{-1}(\ell_X(t)+x) : t>0\} \text{ for any } x \geq 0. \tag{5.17}$$

Since $\ell_{\xi_v}^* = L_v^*$ (*see* Bercu et al [1], p. 29), it follows that

$$\psi_{H_\lambda}(1/p) \leq 1 + v^{-1} L_v^{*-1}(H_\lambda) \text{ where } H_\lambda = \mathcal{H}(Y_\lambda) \text{ and } v = p/(1-p). \tag{5.18}$$

Next X belongs to $\mathcal{G}_\mathcal{E}(1)$, which ensures that $H_\lambda \leq \lambda^2/2$. Hence, using the monotonicity of L_v^{*-1},

$$\psi_{H_\lambda}(1/p) \leq 1 + v^{-1} L_v^{*-1}(\lambda^2/2). \tag{5.19}$$

From (5.15), (5.19) and the fact that X is also subGaussian with parameter 1, we get that, for any $\lambda > 0$,

$$\tilde{Q}_X(p) \leq \lambda^{-1}\big(\lambda^2/2 + \log\big(1 + v^{-1}L_v^{*-1}(\lambda^2/2)\big)\big). \tag{5.20}$$

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

$$\tilde{Q}_X(p) \leq (2L_v^*(x))^{-1/2}\big(L_v^*(x) + \log(1+x/v)\big) := \varphi(x). \tag{5.21}$$

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 (5.21). For this choice of x,

$$\log(1+x/v) = |\log v| = -\log v = -\log(1-x). \tag{5.22}$$

Hence $L_v^*(1-v) = (1+v)^{-1}(1-v)|\log v|$ and $\varphi(1-v) = \sqrt{2|\log v|/(1-v^2)}$. Now

$$\varphi'(1-v) = \big(2L_v^*(1-v)\big)^{-1/2}\big(1 - 2v|\log v|/(1-v^2)\big) > 0 \tag{5.23}$$

by (5.12), which gives the second part of (b). Theorem 5.1 is proved. □

6 Martingales with bounded increments

In this section, $(M_k)_{0\leq k\leq n}$ is a martingale with bounded increments. Applying the results of Section 5, we derive the new upper bound below on the CVaR of M_n from Lemma 2.4 in van de Geer [18].

Theorem 6.1. Let $(M_k)_{0\leq k\leq n}$ be a martingale with bounded increments, adapted to some filtration $(\mathcal{F}_k)_{0\leq k\leq n}$, satisfying $M_0 = 0$. Assume that for each k in $[1,n]$, there exist two $\mathcal{F}_{k-1}$-measurable bounded random variables A_k and B_k such that $A_k \leq M_k \leq B_k$ almost surely. Then, for any real $z > 2$,

$$\tilde{Q}_{M_n}(1/z)) \leq \sqrt{\frac{D_n|\log v|}{2(1-v^2)}}, \text{ where } v = 1/(z-1) \text{ and } D_n = \Big\| \sum_{k=1}^n (B_k - A_k)^2 \Big\|_\infty. \tag{6.1}$$

Proof. Let $W_0 = 0$ and $W_k = \sum_{j=1}^k (B_j - A_j)^2$ for k in $[1,n]$. By Lemma 2.4 in van de Geer [18], $\zeta_k(t) := \exp(tM_k - W_k t^2/8)$ is a supermartingale for any positive t. Hence $\mathbb{E}(\zeta_n(t) \leq \mathbb{E}(\zeta_0(t)) = 1$. Next $\zeta_n(t) \geq \exp(tM_n - D_n t^2/8)$ almost surely, whence

$$\log \mathbb{E}(\exp(tM_n)) \leq t^2 D_n/8 \text{ for any } t > 0. \tag{6.2}$$

Theorem 6.1 follows now from (6.2) and Theorem 5.1(a) applied with $p = 1/z$. □

Numerical comparisons. By Theorem 2.5 in van de Geer [18] applied with $c^2 = D_n$,

$$Q_{M_n}(1/z) \le \sqrt{D_n/4}\,\sqrt{2\log z}. \tag{6.3}$$

By the comparison inequality (2.8) in Pinelis [15],

$$\tilde{Q}_{M_n}(1/z) \le \sqrt{\Delta_n/4}\,Q_Y((5/e)^5/(5!z)\,),\ \text{where } \Delta_n = \sum_{k=1}^{n} \|B_k - A_k\|_\infty^2. \tag{6.4}$$

Below we give the values of the above upper bounds in the case $\Delta_n = D_n = 4$ for integer values of z, including the quartile and the decile. One can see that (6.4) is better than (6.1) as soon as $z > 10$. However D_n is often strictly less than Δ_n in the applications.

Ineq.	z=4	z=6	z=10	z=20	z=40	z=100
(6.3)	1.67	1.89	2.15	2.45	2.72	3.035
(6.4)	1.71	1.89	2.11	2.37	2.62	2.92
(6.1)	1.57	1.83	2.11	2.43	2.71	3.032

7 Bounded functions of independent random variables

Throughout this section E is a Polish space and f is a bounded and measurable function from E^n into $\mathbf{R}$. For any k in $[1, n]$, let δ_k be a bounded and measurable function from E^{n-1} into $[0, \infty[$ such that, for any $x = (x_1, x_2, \ldots, x_n)$ in E^n,

$$\sup_{(y,z)\in E\times E} |f(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_{k-1}, z, x_{k+1}, \ldots, x_n)| \le \delta_k(x^{(k)}), \tag{7.1}$$

where $x^{(k)} = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$. Define then

$$V_n(x) = \frac{1}{4}\sum_{k=1}^{n} \delta_k^2(x^{(k)})\ \text{ and } T_n = \|V_n\|_\infty^{1/2}. \tag{7.2}$$

Let X be a random vector with values in E^n and independent components. Applying the results of Section 5, we now derive the new upper bound below on the CVaR of $f(X)$ from the inequalities of Section 6.2 in Boucheron et al [4].

Theorem 7.1. Let f be a bounded and measurable function from E^n into $\mathbf{R}$. Let X be any random vector with values in E^n and independent components. Set $Z = f(X) - \mathbb{E}(f(X))$. Then, for any $z > 2$,

$$\tilde{Q}_Z(1/z) \le T_n \inf_{x\in]0,1[} \frac{L_v^*(x) + \log(1 + x/v)}{\sqrt{2L_v^*(x)}}, \tag{7.3}$$

where $v = 1/(z-1)$ and L_v^* is defined as in Theorem 5.1.

Proof. Let $\ell_Z(t) = \log \mathbb{E}(\exp(tZ))$. Then, $t\ell'(t) - \ell(t) \le T_n^2 t^2/2$ for any positive t, by the inequality page 175, line 9 in Boucheron et al [4]. Consequently Z belongs to $\mathcal{G}_{\mathcal{E}}(T_n)$. Now Theorem 5.1(b) implies (7.3). □

Numerical comparisons. By Theorem 6.5 in Boucheron et al [4],

$$Q_Z(1/z) \le T_n\,\sqrt{2\log z} \text{ for any } z > 1. \tag{7.4}$$

By the comparison inequality (2.8) in Pinelis [15],

$$\tilde{Q}_Z(1/z) \le \Theta_n\,Q_Y((5/e)^5/(5!z)\,) \text{ for any } z > 1,\ \text{where } \Theta_n = \frac{1}{2}\sqrt{\sum_{k=1}^{n} \|\delta_k\|_\infty^2}. \tag{7.5}$$

Below we give the values of the above upper bounds in the case $\Theta_n = T_n = 1$ for integer values of z, including the quartile and the decile. One can see that (7.5) is better than (7.3) for $z > 16$ and almost equivalent for $z = 16$. However T_n is often strictly less than Θ_n in the applications. Note also that (7.3) is significantly better than (7.4) for $z \leq 40$.

Ineq.	z=4	z=6	z=10	z=16	z=20	z=40
(7.4)	1.67	1.89	2.15	2.35	2.45	2.72
(7.5)	1.71	1.89	2.11	2.29	2.37	2.62
(7.3)	1.55	1.80	2.07	2.29	2.39	2.67

References

[1] Bercu, B., Delyon, B. and Rio, E. Concentration inequalities for sums and martingales. SpringerBriefs in Mathematics. Springer, Cham, x+120 pp. (2015). MR-3363542

[2] Blackwell, D. and Dubins, L. A converse to the dominated convergence theorem. *Illinois J. Math.* **7**, 508–514 (1963). MR-0151572

[3] Bobkov, S., Houdré, C. and Tetali, P. The subgaussian constant and concentration inequalities. *Israel J. Math.* **156**, 255–283 (2006). MR-2282379

[4] Boucheron, S., Lugosi, G. and Massart, P. Concentration inequalities. A nonasymptotic theory of independence. Oxford University Press, Oxford (2013). MR-3185193

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

[6] Doob, J. L. Regularity properties of certain families of chance variables. *Trans. Amer. Math. Soc.* **47**, 455–486 (1940). MR-0002052

[7] Dubins, L. and Gilat, D. On the distribution of maxima of martingales. *Proc. Amer. Math. Soc.* **68**, no. 3, 337–338 (1978). MR-0494473

[8] 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). MR-0840624

[9] 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). MR-0960529

[10] Hardy, G. and Littlewood, J. A maximal theorem with function-theoretic applications. *Acta Math.* **54**, no. 1, 81–116 (1930). MR-1555303

[11] Harremoës, P. Some new maximal inequalities. *Statist. Probab. Lett.* **78**, no. 16, 2776–2780 (2008). MR-2465121

[12] Hoeffding, W. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30 (1963). MR-0144363

[13] Ledoux, M. On Talagrand's deviation inequalities for product measures. *ESAIM Probab. Statist.* **1**, 63–87 (1996). MR-1399224

[14] Osekowski, A. Sharp martingale and semimartingale inequalities. *Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series)* **72**. Birkhauser/Springer (2012). MR-2964297

[15] Pinelis, I. On normal domination of (super)martingales. *Electron. J. Probab.* **11**, 1049–1070 (2006). MR-2268536

[16] Rio, E. About Doob's inequality, entropy and Tchebichef. hal-01630272, version 1 (2017).

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

[18] van de Geer, S. A. On Hoeffding's inequality for dependent random variables. *Empirical process techniques for dependent data,* 161–169, Birkhauser, Boston, MA, 2002. MR-1958780

[19] 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).

[20] Wald, A. Generalization of the inequality of Markoff. *The Annals of Mathematical Statistics* **9**, no 4, 244–255 (1938).