Electron. Commun. Probab. **18** (2013), no. 54, 1–6.
DOI: 10.1214/ECP.v18-2690
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Extensions of the Hoeffding-Azuma inequalities

Emmanuel Rio*

Abstract

In this paper we give extensions of the Hoeffding-Azuma inequalities for weighted sums of uniformly bounded martingale differences. Our results improve previous results of Antonov (1979).

Keywords: Hoeffding inequality ; Azuma inequality ; Discrete time martingales ; McDiarmid inequality .
AMS MSC 2010: 60E15.
Submitted to ECP on March 21, 2013, final version accepted on June 27, 2013.

1 Introduction

In this paper, we are interested in deviation inequalities for martingales with bounded differences. Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ denote an increasing filtration and $(X_n)_{n>0}$ is a sequence of real-valued integrable random variables, adapted to the above filtration. We consider a sequence $(\Delta_n)_{n>0}$ of nonnegative deterministic reals. The martingale $(M_n)_{n\geq 0}$ is defined by $M_0 = 0$ and

$$M_n = \sum_{k=1}^{n} \Delta_k \big(X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) \big) \ \text{ for } n > 0. \tag{1.1}$$

Throughout the paper, we assume that the sequences $(X_n)_{n>0}$ and $(\Delta_n)_{n>0}$ satisfy the additional integrability condition below:

$$\sup_{k>0} \mathbb{E}(|X_k|^p) \leq 1 \ \text{ and } \ \sum_{k>0} \Delta_k^p < \infty, \ \text{ for some } p \in]1,2]. \qquad H(p)$$

Let $Y_n = M_n - M_{n-1}$. By the Whittle inequality - see Inequality (13) in Whittle (1969) -

$$\mathbb{E}(|M_n|^p) \leq 2^{2-p} \sum_{k=1}^{n} \mathbb{E}(|Y_k|^p) \leq 4 \sum_{k>0} \Delta_k^p$$

under assumption $H(p)$. Hence, by the martingale convergence theorem, $(M_n)_n$ converges almost surely and in L^p to the random variable

$$M_\infty = \sum_{k=1}^{\infty} \Delta_k \big(X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) \big). \tag{1.2}$$

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

In this paper, we are interested in exponential decay of the tail function of M_∞ under additional conditions on the random variables X_k. Let us first recall the Hoeffding-Azuma inequality (see Devroye and Lugosi (2001), Chapter 2 for a proof). Assume that the sequence $(\Delta_n)_{n>0}$ belongs to $\ell^2(\mathbb{N}^*)$ and the random variables X_k take their values in $[0,1]$. In that case $H(2)$ holds true. For $p \geq 1$, set

$$\|\Delta\|_p = \Big(\sum_{k>0} \Delta_k^p\Big)^{1/p}. \tag{1.3}$$

The Hoeffding-Azuma inequality states that

$$\mathbb{P}(M_\infty \geq \|\Delta\|_2\, x) \leq \exp(-2x^2) \text{ for any positive } x. \tag{1.4}$$

Assume now that the sequence $(\Delta_k)_{k>0}$ satisfies the stronger assumption $\|\Delta\|_p < \infty$ for some $p < 2$. Under this stronger condition, Antonov (1979) proved that, for independent and centered random variables X_k with values in $[-1/2, 1/2]$,

$$\mathbb{P}(M_\infty \geq \|\Delta\|_p\, x) \leq \exp(-C_q x^q) \text{ with } q = p^* = p/(p-1) \text{ and } C_q = 8q^{-q}(q-1)^{q-1},$$

which is a much better tail estimate. For $q = 2$, $C_q = 2$. However the constant C_q is decreasing with respect to q and converges to 0 as q tends to ∞. For example, $C_q = 32/27$ for $q = 3$ and $C_q = 27/32$ for $q = 4$. In this paper, we will give more efficient constants. In particular we will prove that, for random variables X_k with values in $[0,1]$ (this assumption is weaker than Antonov's assumption),

$$\mathbb{P}(M_\infty \geq \|\Delta\|_p\, x) \leq \exp(-2x^q) \text{ with } q = p^* = p/(p-1), \tag{1.5}$$

which extends the Hoeffding-Azuma inequality with the same constant as for $p = 2$. This inequality will be derived from a more general result, which is stated and proved in Section 2. Next, in Section 3, we apply this general result to the Azuma inequality under symmetric conditions of boundedness. In Section 4, we extend the classical Hoeffding-Azuma inequality (1.4).

2 The main inequality

Throughout this section, we assume that $H(p)$ holds true for some p in $]1,2]$. Our main result is the following extension of the Hoeffding-Azuma inequalities.

Theorem 2.1. *Assume that there exists a convex and increasing function* $\ell : \mathbb{R}^+ \to \mathbb{R}^+$, *such that* $\ell(0) = \ell'(0) = 0$ *and*

$$\log \mathbb{E}\big(\exp\big(tX_k - t\mathbb{E}(X_k \mid \mathcal{F}_{k-1})\big) \mid \mathcal{F}_{k-1}\big) \leq \ell(t) \text{ a.s., for any } t \geq 0 \text{ and any } k > 0. \tag{2.1}$$

Let ℓ^* *denote the Young transform of* ℓ, *which is defined by* $\ell^*(x) = \sup_{t>0}(tx - \ell(t))$. *Suppose that* $H(p)$ *holds true, for some* p *in* $]1,2]$. *Set* $q = p/(p-1)$. *For any function* $\psi : \mathbb{R}^+ \to \mathbb{R}^+$, *let*

$$C_\psi(q) = \inf_{x>0} (\psi(x)/x^q). \tag{2.2}$$

Then, for any positive x,

$$\mathbb{P}(M_\infty \geq \|\Delta\|_p x) \leq \exp(-C_{\ell^*}(q) x^q).$$

Proof of Theorem 2.1. Since M_n converges to M_∞ almost surely, it is enough to prove Theorem 2.1 with M_n instead of M_∞. Now, from (2.1), by induction on n,

$$\log \mathbb{E}\big(\exp(tM_n)\big) \leq \ell(\Delta_1 t) + \ell(\Delta_2 t) + \cdots + \ell(\Delta_n t) := L(t). \tag{2.3}$$

Let $a_\ell(p) = \sup_{t>0}(\ell(t)/t^p)$. By (2.3), for any positive t,

$$L(t) \le a_\ell(p)\big(\Delta_1^p + \Delta_2^p + \cdots + \Delta_n^p\big)\, t^p \le a_\ell(p)\, \|\Delta\|_p^p\, t^p. \tag{2.4}$$

Since ℓ is convex, $\ell = (\ell^*)^*$, wich ensures that $\ell(t) = \sup_{x>0}(xt - \ell^*(x))$ for any positive t. Consequently

$$a_\ell(p) = \sup_{x>0}\sup_{t>0} t^{-p}(xt - \ell^*(x)). \tag{2.5}$$

Let then $f(t) = t^{-p}(xt - \ell^*(x))$. We have to compute the maximum of f. Since $\ell(0) = \ell'(0) = 0$, $\ell^*(x) > 0$ for any positive x. Now $f'(t) = t^{-1-p}(p\ell^*(x) - (p-1)xt)$. Hence f has an unique maximum at the point $t_x = q\ell^*(x)/x$. Therefore

$$\sup_{t>0} t^{-p}(xt - \ell^*(x)) = (q-1)\ell^*(x)t_x^{-p} = q^{1-p}p^{-1}\big(x^q/\ell^*(x)\big)^{p-1}. \tag{2.6}$$

Both (2.5) and (2.6) imply that

$$a_\ell(p) = p^{-1}\big(qC_{\ell^*}(q)\big)^{1-p}. \tag{2.7}$$

Combining (2.7) and (2.4), we then get that

$$L(t) \le p^{-1}\Big(\big(qC_{\ell^*}(q)\big)^{-1/q}\|\Delta\|_p\, t\Big)^p. \tag{2.8}$$

Now, recall that the Young transform of $t \to (t^p/p)$ is $x \to (x^q/q)$. Hence, from (2.8) and elementary computations, $L^*(x) \ge C_{\ell^*}(q)\big(x/\|\Delta\|_p\big)^q$, which, together with (2.3), completes the proof of Theorem 2.1. □

3 Application to Azuma type inequalities

In this section, we assume that the sequence $(X_n)_{n>0}$ satisfies Azuma's condition

$$\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) = 0 \;\text{ and }\; X_n \in [-1,1] \;\text{ almost surely, for any } n > 0. \tag{3.1}$$

Under condition (3.1), it is well known (see Bentkus (2004), for example) that

$$\log \mathbb{E}\big(\exp(tX_k) \mid \mathcal{F}_{k-1}\big) \le \log\cosh(t) \text{ a.s., for any } t \ge 0 \text{ and any } k > 0. \tag{3.2}$$

Hence condition (2.1) holds true with $\ell(t) = \log\cosh(t)$. Now (see Bentkus (2004) for example), $\ell^*(x) = \psi(x)$, with

$$\psi(x) = \tfrac{1}{2}\big((1+x)\log(1+x) + (1-x)\log(1-x)\big) \text{ for } x \le 1 \text{ and } \psi(x) = \infty \text{ for } x > 1. \tag{3.3}$$

Starting from (3.3) and Theorem 2.1, we now prove the following extension of the Azuma inequality (1967).

Theorem 3.1. *Let p be any real in $]1,2]$. Set $q = p/(p-1)$. Assume that condition (3.1) holds true and that the sequence $(\Delta_k)_{k>0}$ belongs to $\ell^p(\mathbb{N}^*)$. Let the constant $c(q)$ be defined by*

$$c(q) = 2^{-2/3} \;\text{ for } q \ge 8/3 \;\text{ and } c(q) = \frac{\sqrt[3]{3}}{2}\,\frac{\big((q/2)-(2/3)\big)^{(q/2)-(2/3)}}{(q-2)^{(q-2)/2}} \;\text{ for } q \in [2, 8/3],$$

with the convention that $0^0 = 1$. Then, for any positive x,

$$\mathbb{P}(M_\infty \ge \|\Delta\|_p x) \le \exp(-c_s(q)x^q) \text{ with } c_s(q) \ge c(q) \ge 1/2.$$

Remark 3.1. *From Equation (3.7) below,* $c(q)$ *is increasing with respect to* q *on* $[2, 8/3]$*. Note also that* $c(2) = 1/2$*. The threshold* $q = 8/3$ *in Theorem 3.1 comes from the function* φ *defined in (3.6) below. This threshold is purely technical.*

Remark 3.2. *From Equation (3.4) below, Theorem 3.1 holds with* $c_s(q) = C_\psi(q)$*. Since* $\psi(1) = \log 2$ *and* $\lim_{x\uparrow 1} \psi'(x) = \infty$*,* $C_\psi(q) < \log 2$ *for any* $q > 2$*. Furthermore, from the continuity of* ψ*,* $\lim_{q\uparrow\infty} C_\psi(q) = \log 2$*.*

Proof of Theorem 3.1. Clearly condition $H(p)$ holds true. Consequently we may apply Theorem 2.1. In particular Theorem 3.1 holds true with

$$c_s(q) = C_\psi(q) = \inf_{x\in]0,1]} (\psi(x)/x^q). \tag{3.4}$$

Now $\psi(0) = \psi'(0) = 0$ and $\psi''(x) = 1/(1-x^2) = 1 + x^2 + \cdots + x^{2k} + \cdots$. Hence, for any x in $[0, 1]$,

$$\psi(x) = \sum_{k>0} \frac{x^{2k}}{2k(2k-1)}. \tag{3.5}$$

We now prove that

$$\psi(x) \geq (1 - x^2/2)^{-1/3}(x^2/2) := \varphi(x)\,. \tag{3.6}$$

Let $u = x^2/2$. Then

$$\varphi(x) = u(1-u)^{-1/3} = \sum_{k\geq 0} a_k u^{k+1} \ \text{ with } a_0 = 1 \text{ and } a_k = \frac{1.4\ldots(3k-2)}{3.6\ldots(3k)} \text{ for } k > 0.$$

Now

$$\psi(x) = \sum_{k\geq 0} b_k u^{k+1} \ \text{ with } b_k = 2^k/((2k+1)(k+1)) \text{ for any } k \geq 0.$$

In order to compare φ and ψ, we will compare a_k/a_{k-1} and b_k/b_{k-1}. Clearly

$$a_k/a_{k-1} = (3k-2)/(3k) \ \text{ and } b_k/b_{k-1} = 2k(2k-1)/((2k+1)(k+1)),$$

from which $(a_1/a_0) = (1/3) = (b_1/b_0)$ and, for $k \geq 2$,

$$\frac{a_k b_{k-1}}{a_{k-1} b_k} = \frac{(3k-2)(2k+1)(k+1)}{6k^2(2k-1)} \leq \frac{6k^3 + 5k^2 - 3k - 2}{6k^3 + 6k^2} \leq 1.$$

Since $a_0 = b_0 = 1$, by induction on k, $a_k \leq b_k$ for any natural integer k, which implies (3.6). Now, by (3.4) and (3.6),

$$c_s(q) \geq c(q) := \inf_{x\in]0,1]} x^{-q}\varphi(x) = 2^{-q/2}\Big(\sup_{u\in]0,1/2]} u^{(q-2)/2}(1-u)^{1/3}\Big)^{-1}, \tag{3.7}$$

using again the change of variable $u = x^2/2$. Let $f(u) = u^{(q-2)/2}(1-u)^{1/3}$. For $q \geq 8/3$, the function f is increasing on $[0, 1/2]$. Then the maximum is reached at the point $u = 1/2$, from which $c(q) = 2^{-2/3}$. For $q \leq 8/3$, the maximum of f is reached at the point $u_q = (3q-6)/(3q-4)$ (which belongs to $[0, 1/2]$). Then

$$c(q) = 2^{-q/2}\,\frac{(3q-4)^{(q/2)-(2/3)}}{\sqrt[3]{2}\,(3q-6)^{(q-2)/2}} = \frac{\sqrt[3]{3}}{2}\ \frac{\big((q/2)-(2/3)\big)^{(q/2)-(2/3)}}{(q-2)^{(q-2)/2}}\,,$$

which completes the proof of Theorem 3.1. □

We now give another formulation of Theorem 3.1, which provides better estimates than the usual Azuma inequality. The proof, being immediate, is omitted.

Corollary 3.1. *Assume that (3.1) holds true and that the sequence* $(\Delta_k)_{k>0}$ *belongs to* $\ell^p(\mathbb{N}^*)$ *for some* $p < 2$. *Let* $c(q)$ *be defined as in Theorem 3.1. Then, for any positive* z *and any* r *in* $]1, 2[$ *such that* $(\Delta_k)_{k>0}$ *belongs to* $\ell^r(\mathbb{N}^*)$,

$$\mathbb{P}(M_\infty \geq z) \leq \exp\Big(-\sup_{q\in[2,r^*]} c(q)\Big(\frac{z}{\|\Delta\|_{q^*}}\Big)^q\Big), \text{ with } q^* = \tfrac{q}{q-1} \text{ and } r^* = \tfrac{r}{r-1}.$$

Remark 3.3. *For any fixed positive* z, *let* $g(q) = \log\big(c(q)\big(z/\|\Delta\|_{q^*})^q\big)$. *Then* g *is differentiable on* $]2, r^*[\setminus \{\frac{8}{3}\}$, *continuous on* $[2, r^*]$, *and, for* $q \neq \frac{8}{3}$,

$$g'(q) = \frac{c'(q)}{c(q)} + \log z - q^* \log(\|\Delta\|_{q^*}) + (q^* - 1)\Big(\sum_{k>0} \Delta_k^{q^*} \log(\Delta_k) / \sum_{k>0} \Delta_k^{q^*}\Big). \tag{3.8}$$

Since $\lim_{q\downarrow 2} c'(q) = +\infty$, *(3.8) ensures that* $\lim_{q\downarrow 2} g'(q) = +\infty$. *Hence the optimal value* q_{opt} *of* q *in Corollary 3.1 satisfies* $q_{opt} > 2$. *Therefore Corollary 3.1 strictly improves the Azuma inequality.*

4 Application to Hoeffding type inequalities

In this section, we assume that $(X_n)_{n>0}$ satisfies the Hoeffding type condition

$$X_n \in [0, 1] \text{ almost surely, for any } n > 0. \tag{4.1}$$

Under (4.1), by Inequality (4.6) in Rio (2013), for any positive integer k and any $t > 0$,

$$\log \mathbb{E}\big(\exp\big(tX_k - t\mathbb{E}(X_k \mid \mathcal{F}_{k-1})\big) \mid \mathcal{F}_{k-1}\big) \leq \ell(t) \text{ almost surely,} \tag{4.2}$$

with $\ell(t) = (t - \log t - 1) + t(e^t - 1)^{-1} + \log(1 - e^{-t})$. Furthermore, by Inequality (2.2) in Rio (2013),

$$\ell^*(x) \geq \max(\psi_1(x), \psi_2(x)), \tag{4.3a}$$

where ψ_1 and ψ_2 are defined by

$$\psi_1(x) = 2x^2 + (4x^4/9) \text{ and } \psi_2(x) = (x^2 - 2x)\log(1 - x), \tag{4.3b}$$

with the convention that $\psi_2(x) = +\infty$ for $x \geq 1$. The inequality $\ell^*(x) \geq \psi_1(x)$ is in fact due to Krafft (1969) and the second inequality $\ell^*(x) \geq \psi_2(x)$ is proved in Rio (2013). From (4.2), (4.3) and Theorem 2.1, we now derive the following extension of the Hoeffding inequality (1963).

Theorem 4.1. *Assume that (4.1) holds true and that the sequence* $(\Delta_k)_{k>0}$ *belongs to* $\ell^p(\mathbb{N}^*)$ *for some* $p < 2$. *Let* $c_1(q)$ *be defined by*

$$c_1(q) = \tfrac{22}{9} \text{ for } q \geq \tfrac{26}{11} \text{ and } c_1(q) = 4\left(\tfrac{\sqrt{2}}{3}\right)^{q-2} \frac{(4-q)^{(q-4)/2}}{(q-2)^{(q-2)/2}} \text{ for } q \in [2, \tfrac{26}{11}],$$

with the convention that $0^0 = 1$. *Then, for any positive* z *and any* r *in* $]1, 2[$ *such that* $(\Delta_k)_{k>0}$ *belongs to* $\ell^r(\mathbb{N}^*)$,

$$\mathbb{P}(M_\infty \geq z) \leq \exp\Big(-\sup_{q\in[2,r^*]} c_a(q)\Big(\frac{z}{\|\Delta\|_{q^*}}\Big)^q\Big), \text{ with } c_a(q) \geq c_1(q).$$

Remark 4.1. *From Equation (4.4) below, $c_1(q)$ is nondecreasing with respect to q. Furthermore $c_1(2) = 2$. Hence $c_1(q) \geq 2$, which yields Inequality (1.5).*

Remark 4.2. *For any fixed positive z, let $g_1(q) = \log\bigl(c_1(q)\bigl(z/\|\Delta\|_{q^*}\bigr)^q\bigr)$. Then g_1 is differentiable on $]2, r^*[\setminus \{\frac{26}{11}\}$ and continuous on $[2, r^*]$. Since $\lim_{q\downarrow 2} c_1'(q) = +\infty$, using the same arguments as in Remark 3.3, we get that $\lim_{q\downarrow 2} g_1'(q) = +\infty$. Hence Theorem 4.1 strictly improves the usual Hoeffding inequality.*

Remark 4.3. *From (4.3) and (4.4) below, $c_a(q) \geq C_{\psi_2}(q)$. Since $\lim_{x\uparrow 1} \psi_2(x) = +\infty$, it implies that $\lim_{q\uparrow\infty} c_a(q) = \infty$. Hence the lower bound $c_a(q) \geq c_1(q)$ is suboptimal for large values of q.*

Proof of Theorem 4.1. As in Section 3, we may apply Theorem 2.1. Both Theorem 2.1 and (4.3) ensure that Theorem 4.1 holds true with

$$c_a(q) = \inf_{x\in]0,1]} (\ell^*(x)/x^q) \geq c_1(q) := \inf_{x\in]0,1]} (\psi_1(x)/x^q). \tag{4.4}$$

It remains to compute $c_1(q)$. Let $f_q(x) = x^{-q}\psi(x) = 2x^{2-q} + (4/9)x^{4-q}$. Then

$$f_q'(x) = 2x^{1-q}((2-q) + (2/9)(4-q)x^2).$$

Hence, for $q \geq 4$, $f_q'(x) \leq 0$ for any positive x. In that case $c_1(q) = f_q(1) = 22/9$. For $q < 4$,

$$f_q'(x) = (4/9)(4-q)x^{1-q}(x^2 - 9(q-2)/(8-2q)).$$

If $q \geq 26/11$, then $9(q-2)/(8-2q) \geq 1$. In that case, $f_q'(x) \leq 0$ for any $x \leq 1$ and therefrom $c_1(q) = f_q(1) = 26/11$. Now, for q in $]2, 26/11]$, the positive real x_q defined by $x_q^2 = 9(q-2)/(8-2q)$ belongs to $]0,1]$. Then

$$c_1(q) = f_q(x_q) = 2x_q^{2-q}(1 + (4/9)x_q^2) = 4\left(\frac{\sqrt{2}}{3}\right)^{q-2} \frac{(4-q)^{(q-4)/2}}{(q-2)^{(q-2)/2}},$$

which completes the proof of Theorem 4.1. □

References

[1] Antonov, S. N. Probability inequalities for a series of independent random variables. *Teor. Verojatnost. i Primenen.* **24**, no. 3, (1979), 632–636. MR-0541380

[2] Azuma, Kazuoki. Weighted sums of certain dependent random variables. *Tôhoku Math. J.* **19**, (1967) 357–367. MR-0221571

[3] Bentkus, Vidmantas. On Hoeffding's inequalities. *Ann. Probab.* **32**, no. 2 (2004), 1650–1673. MR-2060313

[4] Devroye, L., Lugosi, G. Combinatorial methods in density estimation. Springer Series in Statistics. Springer-Verlag, New York, 2001. MR-1843146

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

[6] Krafft, O. A note on exponential bounds for binomial probabilities. *Annals of the Institute of Statistical Mathematics* **21**, (1969), 219–220. Zbl 0176-49106.

[7] Rio, E. On McDiarmid's concentration inequality. *Electron. Commun. Probab.* **18**, (2013), no. 44, 1–11.

[8] Whittle, P. Refinements of Kolmogorov's inequality. *Teor. Verojatnost. i Primenen.* **14**, (1969), 315–317. MR-0246353