

# Comparison inequalities for suprema of bounded empirical processes

Antoine Marchina\*

## Abstract

In this Note we provide comparison moment inequalities for suprema of bounded empirical processes. Our methods are only based on a decomposition in martingale and on comparison results concerning martingales proved by Bentkus and Pinelis.

**Keywords:** concentration inequalities; generalized moments; empirical processes; martingale method.

**AMS MSC 2010:** 60E15.

Submitted to ECP on November 5, 2017, final version accepted on May 9, 2018.

## 1 Introduction

Let  $X_1, \dots, X_n$  be a sequence of independent random variables valued in some measurable space  $(\mathcal{X}, \mathcal{F})$  with common distribution  $P$ . Let  $P_n$  denote for every integer  $n$  the empirical probability measure  $P_n := n^{-1}(\delta_{X_1} + \dots + \delta_{X_n})$ . Let  $\mathcal{F}$  be a countable class of measurable functions  $f : \mathcal{X} \rightarrow \mathbb{R}$  such that  $P(f) = 0$  for all  $f \in \mathcal{F}$ . In this Note, we are concerned with concentration properties around the mean of the random variable

$$Z := \sup\{nP_n(f) : f \in \mathcal{F}\}, \quad (1.1)$$

when  $\mathcal{F}$  satisfies a two-sided or a one-sided (from above) boundedness condition. Our approach is based on a decomposition of  $Z - \mathbb{E}[Z]$  into a sum of martingale increments together with comparison inequalities for martingales with (two-sided or one-sided) bounded increments proved by Bentkus [1] and Pinelis [7]. Before going further, let us introduce some notations.

**Definition 1.1.** (i) Let  $\alpha, \beta$  be two reals such that  $\alpha < \beta$ . We say that a random variable  $\theta$  follows a Bernoulli distribution if it assumes exactly two values and we write  $\theta \sim \mathcal{B}_m(\alpha, \beta)$  if

$$\mathbb{P}(\theta = \beta) = 1 - \mathbb{P}(\theta = \alpha) \in ]0, 1[, \text{ and } \mathbb{E}[\theta] = m. \quad (1.2)$$

Notice that

$$\text{Var}(\theta) = (m - \alpha)(\beta - m). \quad (1.3)$$

(ii) For any  $a \geq 0$ ,  $\Gamma_a$  stands for any centered Gaussian random variable with variance equals to  $a$ .

(iii) For any  $a > 0$ ,  $\Pi_a$  stands for any Poisson random variable with parameter  $a$ . We also denote by  $\tilde{\Pi}_a$  the centered Poisson random variable  $\tilde{\Pi}_a := \Pi_a - a$ .

---

\*Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay, 78035 Versailles, France. E-mail: antoine.marchina@uvsq.fr

Let us introduce the class of convex functions in which the comparison inequalities, stated in this Note, are valid.

**Definition 1.2.** Let  $k \in \mathbb{N}^*$ . As usual, we denote by  $\mathcal{C}^k$  the space of  $k$ -times continuously differentiable functions from  $\mathbb{R}$  to  $\mathbb{R}$ . We define the following class of functions:

$$\mathcal{G}^k := \{\varphi \in \mathcal{C}^{k-1} : \varphi, \varphi', \dots, \varphi^{(k-1)} \text{ are convex}\}. \tag{1.4}$$

Furthermore, we use the classical notations  $x_+ := \max(0, x)$  and  $x_+^\alpha = (x_+)^{\alpha}$ . We now recall the two following comparison results, proved by Bentkus [1] and Pinelis [7], which are the main tools in our proofs.

**Proposition 1.3** (Lemmas 4.4 and 4.5 in [1] and Theorem 2.1 and Remark 2.3 in [7]). Let  $s_1^2, \dots, s_n^2, \beta_1, \dots, \beta_n$  be positive reals. Let  $M_n := \sum_{k=1}^n X_k$  be a martingale with respect to a nondecreasing filtration  $(\mathcal{F}_k)$  such that  $M_0 = 0$ ,

$$X_k \leq 1, \text{ and } \mathbb{E}[X_k^2 | \mathcal{F}_{k-1}] \leq s_k^2 \text{ a.s.} \tag{1.5}$$

(i) Let  $s^2 := n^{-1}(s_1^2 + \dots + s_n^2)$  and  $S_n := \theta_1 + \dots + \theta_n$  be a sum of  $n$  independent copies of a random variable  $\theta$  with distribution  $\mathcal{B}_0(-s^2, 1)$  (defined by (1.2)). Then for any  $\varphi \in \mathcal{G}^2$ ,

$$\mathbb{E}[\varphi(M_n)] \leq \mathbb{E}[\varphi(S_n)]. \tag{a}$$

(ii) Additionally to (1.5), assuming that

$$\mathbb{E}_{k-1}[X_{k+}^3] \leq \beta_k \text{ a.s., and } \beta := \frac{1}{n} \sum_{k=1}^n \beta_k \leq s^2, \tag{1.6}$$

we have for any  $\varphi \in \mathcal{G}^3$ ,

$$\mathbb{E}[\varphi(M_n)] \leq \mathbb{E}[\varphi(\Gamma_{n(s^2-\beta)} + \tilde{\Pi}_{n\beta})], \tag{b}$$

where  $\Gamma_{n(s^2-\beta)}$  and  $\tilde{\Pi}_{n\beta}$  are independent and respectively defined by Definition 1.1 (ii) and (iii).

**Remark 1.4.** In fact, the results in the original papers are stated in the following slightly smaller class of functions:

$$\mathcal{H}_+^\alpha := \left\{ \varphi : \varphi(u) = \int_{-\infty}^{\infty} (u-t)_+^\alpha \mu(dt) \text{ for some Borel measure } \mu \geq 0 \text{ on } \mathbb{R} \right. \\ \left. \text{and all } u \in \mathbb{R} \right\},$$

for  $\alpha \in \{2, 3\}$ . The extensions to  $\mathcal{G}^\alpha$  follow from a result of Pinelis [8, Corollary 5.8] (see also [6, Section 2]).

**Remark 1.5.** From moment comparison inequalities in  $\mathcal{H}_+^\alpha$ , such as in the above Proposition, one can derive tail comparison inequalities. We refer the reader to [4, 5, 6] for the statements of these results and for some more details.

Finally, we use the notations:

$$E_k := \mathbb{E} \sup_{f \in \mathcal{F}} P_k(f) \text{ for any } k = 1, \dots, n \text{ and } \bar{E} := n^{-1}(E_1 + \dots + E_n). \tag{1.7}$$

## 2 Results

### 2.1 Two-sided boundedness condition

Here, by two-sided boundedness condition, we mean that  $\mathcal{F}$  is a countable class of measurable functions with values in  $[-a, 1]$  for some positive real  $a$ . Let  $\psi$  be the function defined on  $[0, 1]$  by

$$\psi(x) = x(1-x) \text{ if } x \in [0, 1/2] \text{ and } \psi(x) = 1/4 \text{ if } x \in ]1/2, 1]. \tag{2.1}$$

**Theorem 2.1.** Let  $\mathcal{F}$  be a countable class of measurable functions from  $\mathcal{X}$  into  $[-a, 1]$  such that  $P(f) = 0$  for all  $f \in \mathcal{F}$ . Let  $Z$  be defined by (1.1).

(i) Case  $a \geq 1$ .

Let  $\theta$  be a Bernoulli random variable with distribution  $\mathcal{B}_0(-a, 1)$  (defined by (1.2)). Let  $\theta_1, \dots, \theta_n$  be  $n$  independent copies of  $\theta$  and let  $S_n := \theta_1 + \dots + \theta_n$ . Then for any function  $\varphi \in \mathcal{G}^2$ ,

$$\mathbb{E}[\varphi(Z - \mathbb{E}Z)] \leq \mathbb{E}[\varphi(S_n)]. \tag{a}$$

(ii) Case  $a < 1$ .

Let  $\vartheta$  be a Bernoulli random variable with distribution

$$\mathcal{B}_0\left(- (a + 1)^2 \psi\left(\frac{a + \bar{E}}{a + 1}\right), 1\right)$$

(defined by (1.2)), where  $\psi$  is defined by (2.1), and  $\bar{E}$  is defined by (1.7). Let  $\vartheta_1, \dots, \vartheta_n$  be  $n$  independent copies of  $\vartheta$  and let  $T_n := \vartheta_1 + \dots + \vartheta_n$ . Then for any function  $\varphi \in \mathcal{G}^2$ ,

$$\mathbb{E}[\varphi(Z - \mathbb{E}Z)] \leq \mathbb{E}[\varphi(T_n)]. \tag{b}$$

**Remark 2.2.** If the class  $\mathcal{F}$  satisfies the uniform law of large numbers, that is the random variable  $\sup_{f \in \mathcal{F}} |P_n(f)|$  converges to 0 in probability, then  $E_n$  decreases to 0 (see, for instance, Section 2.4 of van der Vaart and Wellner [10]) and so  $\bar{E}$  also decreases to 0. This ensures that  $n \mapsto (a + 1)^2 \psi((a + \bar{E})/(a + 1))$  (which is also the variance of  $\vartheta$ ) is nonincreasing and tends to  $a$  as  $n$  tends to infinity.

**Example 2.3** (Set-indexed empirical processes). Let  $\mathcal{S}$  be a countable class of measurable sets of  $\mathcal{X}$ . We consider the class of functions

$$\mathcal{F} = \left\{ \frac{\mathbb{1}_S - P(S)}{1 - P(S)} : S \in \mathcal{S} \right\}. \tag{2.2}$$

Let  $p := \sup\{P(S) : S \in \mathcal{S}\}$  and we assume that  $p < 1$ . Since  $x \mapsto x/(1 - x)$  is increasing on  $[0, 1[$ , we can apply Theorem 2.1 with  $a = p/(1 - p)$ . Hence with the notations of Theorem 2.1:

(i) If  $p > 1/2$ , then for any  $\varphi \in \mathcal{G}^2$ ,  $\mathbb{E}[\varphi(Z - \mathbb{E}Z)] \leq \mathbb{E}[\varphi(S_n)]$ , and  $\theta$  has the distribution  $\mathcal{B}_0(-p/(1 - p), 1)$ .

(ii) If  $p < 1/2$ , then for any  $\varphi \in \mathcal{G}^2$ ,  $\mathbb{E}[\varphi(Z - \mathbb{E}Z)] \leq \mathbb{E}[\varphi(T_n)]$ , and  $\vartheta$  has the distribution  $\mathcal{B}_0(-(1 - p)^{-2} \psi(\bar{E} + p(1 - \bar{E})), 1)$ .

Theorem 3.1 of Rio [9], when applied to  $Z$  (see also his Theorem 4.2 (a)), provides a Bennett-type inequality for classes of sets with small measures under  $P$ . Precisely the condition is  $E_n + p(1 - E_n) \leq 1/2$ . Hence since  $\mathcal{G}^2$  contains all increasing exponential functions  $x \mapsto e^{tx}$ ,  $t > 0$ , the Case (ii) above completes Rio's result in this situation.

## 2.2 One-sided boundedness condition

Here, by one-sided boundedness condition, we mean that  $\mathcal{F}$  is a countable class of measurable functions with values in  $]-\infty, 1]$ . Let  $\rho$  be the function defined on  $[0, 1]$  by

$$\rho(x) = x(1 - x)^2 \text{ if } x \in [0, 1/3] \text{ and } \rho(x) = 4/27 \text{ if } x \in ]1/3, 1]. \tag{2.3}$$

**Theorem 2.4.** Let  $\mathcal{F}$  be a countable class of measurable functions from  $\mathcal{X}$  into  $]-\infty, 1]$  such that  $P(f) = 0$  for any  $f \in \mathcal{F}$ . Let  $Z$  be defined by (1.1). Define also

$$\sigma^2 := \sup\{P(f^2) : f \in \mathcal{F}\} \text{ and } m_+^3 := \sup\{P(f_+^3) : f \in \mathcal{F}\}, \tag{2.4}$$

and

$$v^2 := \sigma^2 + 2\psi(\bar{E}) \text{ and } \beta_+^3 := m_+^3 + 3\rho(\bar{E}). \tag{2.5}$$

(i) Let  $\theta$  be a Bernoulli random variable with distribution  $\mathcal{B}_0(-v^2, 1)$  (defined by (1.2)). Let  $\theta_1, \dots, \theta_n$  be  $n$  independent copies of  $\theta$  and let  $S_n := \theta_1 + \dots + \theta_n$ . Then for any function  $\varphi \in \mathcal{G}^2$ ,

$$\mathbb{E}[\varphi(Z - \mathbb{E}[Z])] \leq \mathbb{E}[\varphi(S_n)]. \tag{a}$$

(ii) Let  $\beta := \min\{\beta_+^3, v^2\}$ . Then for any function  $\varphi \in \mathcal{G}^3$ ,

$$\mathbb{E}[\varphi(Z - \mathbb{E}[Z])] \leq \mathbb{E}[\varphi(\Gamma_{n(v^2-\beta)} + \tilde{\Pi}_{n\beta})], \tag{b}$$

where  $\Gamma_{n(v^2-\beta)}$  and  $\tilde{\Pi}_{n\beta}$  are independent and respectively defined by Definition 1.1 (ii) and (iii).

**Remark 2.5.** Inequalities (a) and (b) of Theorem 2.4 are not easy to compare directly. To this end, we start by observing that Inequality (a) yields the following (see [1, Theorem 1.1]):

$$\mathbb{E}[\varphi(Z - \mathbb{E}[Z])] \leq \mathbb{E}[\varphi(\tilde{\Pi}_{n\beta})] \quad \text{for any } \varphi \in \mathcal{G}^2. \tag{a'}$$

One can find in Pinelis [6] a thorough study of the comparison between the right-hand sides of (b) and (a') as well as with other classical bounds. Let us mention here some facts.

(i) If  $\beta = v^2$ , then the right-hand side of (b) is equal to the right-hand side of (a'). Thus, since  $\mathcal{G}^3 \subset \mathcal{G}^2$ , Inequality (b) is relevant with respect to (a) if and only if  $\beta_+^3 < v^2$  (see also the point (ii) below).

(ii) For any  $v^2 > 0$  and any  $\varepsilon \in ]0, 1[$ ,

$$\mathbb{E}[\varphi(\Gamma_{(1-\varepsilon)v^2} + \tilde{\Pi}_{\varepsilon v^2})] \leq \mathbb{E}[\varphi(\tilde{\Pi}_{v^2})] \quad \text{for any } \varphi \in \mathcal{G}^2. \tag{2.6}$$

### 3 Proofs

The starting point of the proofs is a martingale decomposition of  $Z$  which we briefly recall. Firstly by virtue of the monotone convergence theorem, we can suppose that  $\mathcal{F}$  is a finite class of functions. Set  $\mathcal{F}_0 := \{\emptyset, \Omega\}$ ,  $\mathcal{F}_k := \sigma(X_1, \dots, X_k)$  and  $\mathcal{F}_n^k := \sigma(X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n)$  for all  $k = 1, \dots, n$ . Let  $\mathbb{E}_k$  (respectively  $\mathbb{E}_n^k$ ) denote the conditional expectation operator associated with  $\mathcal{F}_k$  (resp.  $\mathcal{F}_n^k$ ). Set also  $Z^{(k)} := \sup\{nP_n(f) - f(X_k) : f \in \mathcal{F}\}$  and  $Z_k := \mathbb{E}_k[Z]$ . Let us number the functions of the class  $\mathcal{F}$  and consider the random variables

$$\tau := \inf\{i > 0 : nP_n(f_i) = Z\} \text{ and } \tau_k := \inf\{i > 0 : nP_n(f_i) - f_i(X_k) = Z^{(k)}\}.$$

We notice that

$$f_{\tau_k}(X_k) \leq Z - Z^{(k)} \leq f_{\tau}(X_k). \tag{3.1}$$

Set now  $\xi_k := \mathbb{E}_k[f_{\tau_k}(X_k)]$  and let  $r_k$  be the nonnegative random variable such that  $\xi_k + r_k = Z_k - \mathbb{E}_k[Z^{(k)}]$ . Since  $\mathbb{E}_n^k[f_{\tau_k}(X_k)] = P(f_{\tau_k})$  (see, for instance, Section 3.2 in [3, Lemma 3.10]), the centering assumption on the elements of  $\mathcal{F}$  ensures that  $\mathbb{E}_{k-1}[\xi_k] = 0$  and thus

$$Z - \mathbb{E}[Z] = \sum_{k=1}^n \Delta_k, \quad \text{where } \Delta_k := \xi_k + r_k - \mathbb{E}_{k-1}[r_k]. \tag{3.2}$$

The important point is that  $\mathbb{E}_{k-1}[r_k]$  is a corrective term which is essentially small. This is the statement of the following lemma:

**Lemma 3.1.** We have  $0 \leq \mathbb{E}_{k-1}[r_k] \leq E_{n-k+1} \leq 1$  a.s.

The proof is based on a property of exchangeability of variables, proved in Marchina [3]. Since it is the fundamental tool of the paper, we give again the proof for sake of completeness.

*Proof of Lemma 3.1.* Since  $f \leq 1$  for any  $f \in \mathcal{F}$ ,  $E_k \leq 1$  for any  $k = 1, \dots, n$ . Moreover,  $r_k \geq 0$  which implies  $\mathbb{E}_{k-1}[r_k] \geq 0$ . Let us now prove the main part  $\mathbb{E}_{k-1}[r_k] \leq E_{n-k+1}$ . We start by the following property of exchangeability of variables.

**Lemma 3.2.** For any integer  $j \geq k$ ,  $\mathbb{E}_{k-1}[f_\tau(X_k)] = \mathbb{E}_{k-1}[f_\tau(X_j)]$  a.s.

*Proof of Lemma 3.2.* By the definition of the random variable  $\tau$ , for every permutation on  $n$  elements  $\sigma$ ,  $\tau(X_1, \dots, X_n) = \tau \circ \sigma(X_1, \dots, X_n)$  almost surely. Applying now this fact to  $\sigma = (k j)$  (the transposition which exchanges  $k$  and  $j$ ), it suffices to use Fubini's theorem (recalling that  $j \geq k$ ) to complete the proof.  $\square$

Hence,

$$\begin{aligned} \mathbb{E}_{k-1}[f_\tau(X_k)] &= \mathbb{E}_{k-1}[f_\tau(X_k) + \dots + f_\tau(X_n)] / (n - k + 1) \\ &\leq \mathbb{E}_{k-1} \sup_{f \in \mathcal{F}} \{f(X_k) + \dots + f(X_n)\} / (n - k + 1) = E_{n-k+1}. \end{aligned} \quad (3.3)$$

Recalling that  $\mathbb{E}_{k-1}[\xi_k] = 0$ , (3.1) yields that  $\mathbb{E}_{k-1}[r_k] \leq \mathbb{E}_{k-1}[f_\tau(X_k)]$ , which combined with (3.3) end the proof of Lemma 3.1.  $\square$

We are now in position to prove Theorem 2.1.

*Proof of Theorem 2.1.* Observe first that  $\Delta_k \leq 1$  by (3.1) and the uniform boundedness condition on  $\mathcal{F}$ . Then, in view of Proposition 1.3 (i), it remains us to bound up the conditional variance with respect to  $\mathcal{F}_{k-1}$  of  $\Delta_k$ . This is the subject of the following lemma:

**Lemma 3.3.** For any  $k = 1, \dots, n$ ,

$$\mathbb{E}_{k-1}[\Delta_k^2] \leq \sup_{m \in [0, E_{n-k+1}]} (m + a)(1 - m) \text{ a.s.}$$

*Proof of Lemma 3.3.* A classical result due to Hoeffding [2] (see his Inequalities (4.1) and (4.2)) states that any bounded random variable  $X$ , such that  $a \leq X \leq b$  for some reals  $a$  and  $b$ , satisfies

$$\mathbb{E}[\varphi(X)] \leq \mathbb{E}[\varphi(\theta)] \text{ for any convex function } \varphi, \quad (3.4)$$

where  $\theta$  is a Bernoulli random variable with distribution  $\mathcal{B}_{\mathbb{E}[X]}(a, b)$  (defined by (1.2)). In particular,  $\text{Var}(X)$  is lower than  $\text{Var}(\theta)$ . We apply this result conditionally to  $\mathcal{F}_{k-1}$  to the variable  $\xi_k + r_k$  which has its values in  $[-a, 1]$  by (3.1). Recalling now (1.3), one immediately obtains

$$\text{Var}(\xi_k + r_k \mid \mathcal{F}_{k-1}) \leq (\mathbb{E}_{k-1}[r_k] + a)(1 - \mathbb{E}_{k-1}[r_k]) \text{ a.s.}, \quad (3.5)$$

which combined with Lemma 3.1 conclude the proof of Lemma 3.3.  $\square$

Let us now complete the proof of Theorem 2.1. Define the function  $V$  by

$$V(m) := (m + a)(1 - m) \text{ for any } m \in [0, E_{n-k+1}]. \quad (3.6)$$

Case (i):  $a \geq 1$ . Since  $(1 - a)/2 \leq 0$ ,  $V$  is decreasing on  $[0, E_{n-k+1}]$  and then by Lemma 3.3,  $\mathbb{E}_{k-1}[\Delta_k^2] \leq a$ . Thus Inequality (a) follows by Proposition 1.3 (i).

Case (ii):  $a < 1$ . Here the maximum of  $V$  is reached at

$$m = \min(E_{n-k+1}, (1 - a)/2).$$

Thus Lemma 3.3 implies

$$\mathbb{E}_{k-1}[\Delta_k^2] \leq (a+1)^2 \psi\left(\frac{a + E_{n-k+1}}{a+1}\right). \quad (3.7)$$

Furthermore, since  $\psi$  is concave,

$$\frac{1}{n} \sum_{k=1}^n \psi\left(\frac{a + E_{n-k+1}}{a+1}\right) \leq \psi\left(\frac{a + \bar{E}}{a+1}\right). \quad (3.8)$$

Hence Inequality (b) follows again from Proposition 1.3 (i) together with (3.7)–(3.8). The proof of Theorem 2.1 is now complete.  $\square$

*Proof of Theorem 2.4.* Case (i). We start from (3.2). Since  $\mathbb{E}_{k-1}[\xi_k] = 0$ , we get

$$\mathbb{E}_{k-1}[\Delta_k^2] = \mathbb{E}_{k-1}[(\xi_k + r_k)^2] - (\mathbb{E}_{k-1}[r_k])^2. \quad (3.9)$$

Since  $\mathbb{E}_n^k[f_{\tau_k}^2(X_k)] = P(f_{\tau_k}^2)$ , the conditional Jensen inequality implies that  $\mathbb{E}_{k-1}[\xi_k^2] \leq \sigma^2$ . Therefore, from (3.9) and the fact that (3.1) implies  $\xi_k + r_k \leq 1$ , we get

$$\begin{aligned} \mathbb{E}_{k-1}[\Delta_k^2] &\leq \sigma^2 + \mathbb{E}_{k-1}[2\xi_k r_k + r_k^2] - (\mathbb{E}_{k-1}[r_k])^2 \\ &\leq \sigma^2 + 2\mathbb{E}_{k-1}[r_k] - (\mathbb{E}_{k-1}[r_k])^2 \\ &\leq \sigma^2 + 2\psi(\mathbb{E}_{k-1}[r_k]) \\ &\leq \sigma^2 + 2\psi(E_{n-k+1}), \end{aligned} \quad (3.10)$$

where the last inequality follows from Lemma 3.1 and  $\psi$  is the nondecreasing function already defined in (2.1). Then the same conclusion as in the proof of the Case (ii) of Theorem 2.1 allows us to conclude the proof of Inequality (a) of Theorem 2.4.

Case (ii). Since  $x \mapsto x_+^3$  is a convex function, we have

$$\Delta_{k+}^3 \leq (\xi_k - \mathbb{E}_{k-1}[r_k])_+^3 + 3r_k(\xi_k + r_k - \mathbb{E}_{k-1}[r_k])_+^2. \quad (3.11)$$

In the same way as previously, since  $\mathbb{E}_n^k[f_{\tau_{k+}}^3(X_k)] = P(f_{\tau_{k+}}^3)$ , one has  $\mathbb{E}_{k-1}[\xi_{k+}^3] \leq m_+^3$ . Therefrom, recalling that  $r_k \geq 0$  and  $\xi_k + r_k \leq 1$ , we get

$$\mathbb{E}_{k-1}[\Delta_{k+}^3] \leq m_+^3 + 3\rho(\mathbb{E}_{k-1}[r_k]) \leq m_+^3 + 3\rho(E_{n-k+1}), \quad (3.12)$$

where the last inequality follows from Lemma 3.1 and  $\rho$  is the nondecreasing function defined in (2.3). Since  $\rho$  is concave, we complete the proof in the same way as the Case (i) by using Proposition 1.3 (ii) in place of Proposition 1.3 (i).  $\square$

## References

- [1] V. Bentkus. (2004). On Hoeffding’s inequalities, *Ann. Probab.* **32**, no. 2, 1650–1673. MR-2060313
- [2] W. Hoeffding. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30.
- [3] A. Marchina. (2017). Concentration inequalities for suprema of unbounded empirical processes. Preprint on <hal-01545101>.
- [4] I. Pinelis. (1998). Optimal tail comparison based on comparison of moments. *High dimensional probability*, Springer, pp. 297–314. MR-1652335
- [5] I. Pinelis. (1999). Fractional sums and integrals of r-concave tails and applications to comparison probability inequalities. *Advances in stochastic inequalities* (Atlanta, GA, 1997) **234**, 149–168. MR-1694770

- [6] I. Pinelis. (2009). On the Bennett-Hoeffding inequality. Preprint. Available at arXiv:0902.4058 [math.PR].
- [7] I. Pinelis. (2014). On the Bennett-Hoeffding inequality. *Ann. Inst. H. Poincaré Probab. Statist.*, **50**, no. 1, 15–27. MR-3161520
- [8] I. Pinelis. (2016). Convex cones of generalized multiply monotone functions and the dual cones. *Banach J. Math. Anal.* **10**, no. 4, 864–897. MR-3555754
- [9] E. Rio. (2001). Inégalités de concentration pour les processus empiriques de classes de parties. *Probab. Theory Related Fields* **119**, no. 2, 163–175.
- [10] A.W. van der Vaart and J.A. Wellner. (1996). *Weak convergence and empirical processes: With applications to Statistics*. *Springer Series in Statistics*. New-York: Springer. MR-1385671

**Acknowledgments.** I am grateful to Emmanuel Rio for several helpful discussions.

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS<sup>1</sup>)
- Easy interface (EJMS<sup>2</sup>)

### Economical model of EJP-ECP

- Non profit, sponsored by IMS<sup>3</sup>, BS<sup>4</sup>, ProjectEuclid<sup>5</sup>
- Purely electronic

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>2</sup>EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

<sup>3</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>4</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>5</sup>Project Euclid: <https://projecteuclid.org/>

<sup>6</sup>IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>