

## SOME FINITE SAMPLE PROPERTIES OF NEGATIVELY DEPENDENT RANDOM VARIABLES

UDC 519.21

ALESSIO FARCOMENI

ABSTRACT. We discuss some finite sample properties of vectors of negatively dependent random variables. We extend some inequalities, widely used for independent random variables, and some basic tools such as the symmetrization lemma, to the case of negatively dependent random variables.

A sequence of random variable is said to be negatively (positively) dependent if

$$\mathbb{P}\left(\bigcap_{i=1}^n \{X_i \leq z_i\}\right) \leq (\geq) \prod \mathbb{P}(X_i \leq z_i)$$

and

$$\mathbb{P}\left(\bigcap_{i=1}^n \{X_i > z_i\}\right) \leq (\geq) \prod \mathbb{P}(X_i > z_i),$$

for  $z_i \in \mathbf{R}$ ,  $i = 1, \dots, n$ . Negative dependence is implied for instance by negative association. A sequence of random variables is said to be negatively (positively) associated if for all monotonically coordinatewise nondecreasing functions  $g_1$  and  $g_2$ ,

$$\text{Cov}[g_1(X_1, \dots, X_n), g_2(X_1, \dots, X_n)] \leq (\geq) 0,$$

when it exists. Positive association was introduced in [3], and negative association in [4]. Negative association implies negative dependence. Moreover, it is straightforward to prove that under either condition,  $\mathbb{E}[\prod X_i] \leq (\geq) \prod \mathbb{E}[X_i]$ . Any subset of a set of negatively associated or negatively dependent random variables is still negatively associated or negatively dependent; and any nondecreasing function of negatively (positively) associated random variables is still negatively (positively) associated. Some further basic properties are given for instance in [7], and [2] makes a review of concepts of negative dependence.

We now provide a brief list of the most important cases of negatively associated random variables: multivariate normal random variables with nonpositive (nonnegative) correlations are negatively (positively) associated. Independent random variables are both positively and negatively associated. Multinomial, multivariate hypergeometric, Dirichlet random variables are always negatively associated. For other examples, we refer for instance to [4].

In this paper we show some properties of negatively dependent random variables. The main goal is to extend tools and inequalities widely used for independent random variables. In Section 1 we provide the symmetrization lemma under arbitrary dependence.

---

2000 *Mathematics Subject Classification.* Primary 60E15, 47N30.

*Key words and phrases.* Negative dependence, association, Hoeffding inequality, exponential tail inequality, bounded difference inequality, empirical distribution, symmetrization lemma.

In Section 2 we provide exponential tail inequalities and a symmetrization argument for negatively dependent random variables. First, the well-known Hoeffding inequality is proved for negatively dependent random variables, then we provide extension of the bounded difference inequality. Finally, we present a symmetrization argument.

## 1. TOOLS

We begin by proving an extension of the symmetrization lemma under arbitrary dependence.

**Definition 1** (Separability). Let  $(Y(u), u \in \mathcal{U})$  be a family of random variables on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . The family is called *separable* if there exists a countable set  $U_0 \subseteq \mathcal{U}$  and a set  $E \in \mathcal{F}$  such that

- (1)  $\mathbb{P}(E) = 1$ ;
- (2) for any  $\omega \in E$  and for any  $u \in \mathcal{U}$  there exists a sequence  $(u_j, j \geq 1)$  in  $U_0$  such that  $Y(u_j, \omega) \rightarrow Y(u, \omega)$  as  $j \rightarrow \infty$ .

**Lemma 1** (Symmetrization Lemma). Let  $(Y(u), u \in \mathcal{U})$  be a separable family of random variables, and  $(Y'(u), u \in \mathcal{U})$  an independent copy of  $(Y(u), u \in \mathcal{U})$  with the same joint distribution for any  $u_1, \dots, u_n$  (that is, with the same dependency structure). Let  $\mathbb{P}(|Y'(u)| > \varepsilon/2) \leq 1/2$  for any  $u \in \mathcal{U}$ . Then

$$\mathbb{P}\left(\sup_u |Y(u)| > \varepsilon\right) \leq 2\mathbb{P}\left(\sup_u |Y(u) - Y'(u)| > \varepsilon/2\right)$$

for any  $\varepsilon > 0$ .

*Proof.* If  $(Y(u), u \in \mathcal{U})$  is separable, then so is  $(Y'(u), u \in \mathcal{U})$ . Moreover, there exists a countable set  $U_0 \subseteq \mathcal{U}$  such that  $\sup_{u \in \mathcal{U}} |Y(u)| = \sup_{u \in U_0} |Y(u)|$ . Let  $u_i$  be the  $i$ -th element of  $U_0$ . Let  $A_1 = \{|Y(u_1)| > \varepsilon\}$ , and

$$A_i = \{|Y(u_1)| \leq \varepsilon, \dots, |Y(u_{i-1})| \leq \varepsilon, |Y(u_i)| > \varepsilon\}$$

for  $i \geq 2$ . Note that if  $|Y(u_i)| > \varepsilon$  and  $|Y'(u_i)| \leq \varepsilon/2$ , then  $|Y(u_i) - Y'(u_i)| > \varepsilon/2$ . We have

$$\begin{aligned} \frac{1}{2}\mathbb{P}\left(\sup_{u \in \mathcal{U}} |Y(u)| > \varepsilon\right) &= \frac{1}{2}\sum_{i \in U_0} \mathbb{P}(A_i) \\ &\leq \sum \mathbb{P}(A_i) \mathbb{P}(|Y'(u_i)| \leq \varepsilon/2) \\ &= \sum \mathbb{P}(A_i, |Y'(u_i)| \leq \varepsilon/2) \\ &\leq \sum \mathbb{P}(A_i, |Y(u_i) - Y'(u_i)| > \varepsilon/2) \\ &\leq \sum \mathbb{P}\left(A_i, \sup_{u \in U_0} |Y(u) - Y'(u)| > \varepsilon/2\right) \\ &\leq \mathbb{P}\left(\sup_{u \in \mathcal{U}} |Y(u) - Y'(u)| > \varepsilon/2\right). \quad \square \end{aligned}$$

## 2. EXPONENTIAL TAIL INEQUALITIES

**2.1. Hoeffding inequality.** The key part in proving Hoeffding inequality for negatively dependent random variables is the thesis of Lemma 2. Note that it can hold also under different assumptions. It is straightforward to see, for instance, that Lemma 2 is true for a vector of binary random variables whenever the covariance between any two of them is nonpositive.

**Lemma 2.** *Suppose  $X_1, \dots, X_n$  is a vector of negatively associated random variables. Then*

$$\mathbb{E} \left( \exp \left\{ t \sum X_i \right\} \right) \leq \prod \mathbb{E}(\exp\{tX_i\})$$

for any  $t > 0$ .

*Proof.* This is a straightforward generalization of Lemma 1 in [6], stemming from the fact that if  $X_1, \dots, X_n$  is a vector of negatively dependent random variables, then  $e^{tX_1}, \dots, e^{tX_n}$  is also negatively dependent.  $\square$

**Theorem 1** (Hoeffding inequality). *Let  $X_1, \dots, X_n$  be a sequence of negatively dependent random variables. Let  $\mathbb{P}(a_i < X_i < b_i) = 1$ ,  $\mathbb{E}(X_i) = 0$ . Let  $S_n = \sum_i (b_i - a_i)^2/8$ . Let  $\varepsilon > 0$ . Then, for any  $t > 0$ ,*

$$(1) \quad \mathbb{P} \left( \sum X_i \geq \varepsilon \right) \leq e^{-t\varepsilon + t^2 S_n}.$$

*Proof.* By Markov's theorem and by Lemma 2,

$$\begin{aligned} \mathbb{P} \left( \sum X_i \geq \varepsilon \right) &= \mathbb{P} \left( t \sum X_i \geq t\varepsilon \right) \\ &= \mathbb{P} \left( e^{t \sum X_i} \geq e^{t\varepsilon} \right) \\ &\leq e^{-t\varepsilon} \mathbb{E} \left( e^{t \sum X_i} \right) \\ &\leq e^{-t\varepsilon} \prod \mathbb{E} \left( e^{tX_i} \right). \end{aligned}$$

The key difference between this proof and the one for independent random variables is at the third step, where we replace equality with an inequality sign. The rest of the proof is analogous to the proof of Hoeffding's inequality for independent random variables.  $\square$

It is interesting to note that defining  $S_n = \sum_i \mathbb{E}[X_i^2]$  in [1] shows an inequality (see (2.8) there) analogous to (1), with  $t$  close enough to zero, for *unbounded* negatively dependent random variables, only putting some additional assumptions on the higher order moments.

**2.2. Bounded difference inequality.** A generalization of Hoeffding's inequality is given by the bounded difference inequality, often used to convert bounds for the expected value to exponential tail inequalities. The bounded difference inequality for independent random variables was first derived in [5].

**Theorem 2** (Bounded difference inequality). *Suppose  $g(\cdot)$  satisfies the bounded difference assumption*

$$(2) \quad \sup_{x_1, \dots, x_n; x'_i \in A} |g(x_1, \dots, x_n) - g(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)| \leq c_i,$$

for  $1 \leq i \leq n$ , and any set  $A$ . Suppose  $X_1, \dots, X_n$  is a vector of negatively dependent random variables. Then, for all  $t > 0$ ,

$$\mathbb{P}\{|g(X_1, \dots, X_n) - \mathbb{E}[g(X_1, \dots, X_n)]| \geq t\} \leq 2 \exp \left\{ -2t^2 / \sum_{i=1}^n c_i^2 \right\}.$$

*Proof.* We will prove that

$$(3) \quad \mathbb{P}\{g(X_1, \dots, X_n) - \mathbb{E}[g(X_1, \dots, X_n)] \geq t\} \leq \exp \left\{ -2t^2 / \sum_{i=1}^n c_i^2 \right\}.$$

Similarly it can be proved that

$$\mathbb{P}\{\mathbb{E}[g(X_1, \dots, X_n)] - g(X_1, \dots, X_n) \geq t\} \leq \exp\left\{-2t^2 / \sum_{i=1}^n c_i^2\right\}.$$

Combination of these two results yields the thesis.

The hypothesis of negative dependence is needed for the application of Theorem 1 in this straightforward extension: let  $V$  and  $Z$  be such that  $\mathbb{E}[V|Z] = 0$ , and for some  $h(\cdot)$  and  $c > 0$ ,  $h(Z) \leq V \leq h(Z) + c$ . Then, for all  $s > 0$ ,

$$(4) \quad \mathbb{E}[e^{sV} | Z] \leq e^{s^2 c^2 / 8}.$$

Now denote  $V = g(X_1, \dots, X_n) - \mathbb{E}[g(X_1, \dots, X_n)]$  and

$$H_i(X_1, \dots, X_i) = \mathbb{E}[g(X_1, \dots, X_n) | X_1, \dots, X_i],$$

and for any  $i$  define

$$(5) \quad V_i = H_i(X_1, \dots, X_i) - H_{i-1}(X_1, \dots, X_{i-1}).$$

Let  $F_i(x) = \mathbb{P}(X_i < x | X_1, \dots, X_{i-1})$ . Clearly,  $V = \sum V_i$  and

$$H_{i-1}(X_1, \dots, X_{i-1}) = \int H_i(X_1, \dots, X_{i-1}, x) F_i(dx).$$

Moreover, define

$$W_i = \sup_u H(X_1, \dots, X_{i-1}, u) - H_{i-1}(X_1, \dots, X_{i-1})$$

and

$$Z_i = \inf_u H(X_1, \dots, X_{i-1}, u) - H_{i-1}(X_1, \dots, X_{i-1}).$$

Clearly,  $\mathbb{P}(Z_i \leq V_i \leq W_i) = 1$  and

$$(6) \quad W_i - Z_i = \sup_u \sup_v (H(X_1, \dots, X_{i-1}, u) - H(X_1, \dots, X_{i-1}, v)) \leq c_i$$

by the bounded difference assumption. Therefore, by (4), for any  $i$ ,

$$(7) \quad \mathbb{E}[e^{sV_i} | X_1, \dots, X_{i-1}] \leq e^{s^2 c_i^2 / 8}.$$

Finally, by Chernoff's bound, for any  $s > 0$ ,

$$\begin{aligned} \mathbb{P}\{g(X_1, \dots, X_n) - \mathbb{E}[g(X_1, \dots, X_n)] \geq t\} &\leq \frac{\mathbb{E}[\exp\{s \sum_{i=1}^n V_i\}]}{e^{st}} \\ &= \frac{\mathbb{E}\left[\exp\left\{s \sum_{i=1}^{n-1} V_i\right\} \mathbb{E}[e^{sV_n} | X_1, \dots, X_{n-1}]\right]}{e^{st}} \\ &\leq e^{s^2 c_n^2 / 8} \frac{\mathbb{E}\left[\exp\left\{s \sum_{i=1}^{n-1} V_i\right\}\right]}{e^{st}} \\ &\leq e^{-st} \exp\left\{s^2 \sum_{i=1}^n c_i^2 / 8\right\}, \end{aligned}$$

by repeating the same argument  $n$  times. Choosing  $s = 4t / \sum c_i^2$  yields inequality (3).  $\square$

**2.3. Inequalities for the empirical measure.** We now provide some inequalities for the empirical measure of negatively dependent random variables. In what follows we define the empirical measure of a vector of random variables  $X_1, \dots, X_n$  as

$$\mu_n(A) = \frac{1}{n} \sum_i 1_{x_i \in A},$$

while the empirical distribution considers a restriction to classes of sets  $A = (-\infty, z]$  and will be denoted by  $\hat{F}(z) = n^{-1} \sum_{i=1}^n 1_{X_i \leq z}$  for  $z \in \mathbf{R}$ .

We will now restrict ourselves to a class of sets such that

$$\mathbf{P}(X_i \in A, X_j \in A) \leq \mathbf{P}(X_i \in A) \mathbf{P}(X_j \in A)$$

for any  $A$  in the class, and do not explicitly request for negative dependence. A class of this kind will be, for instance, the class of sets of the form  $(-\infty, z]$ , for  $z$  real, under (pairwise) negative dependence for the random variables.

**Theorem 3.** *Let  $X_1, \dots, X_n$  be a set of identically distributed random variables, and  $X'_1, \dots, X'_n$  an independent copy of  $X_1, \dots, X_n$ . Let  $\mu(A) = \mathbf{P}(X_i \in A)$ ,*

$$\mu_n(A) = \frac{1}{n} \sum 1_{X_i \in A},$$

*and  $\mu'_n(A) = n^{-1} \sum 1_{X'_i \in A}$ . Suppose that there exists  $\mathcal{A}$ , a class of sets such that*

$$\mathbf{P}(X_i \in A_i, X_j \in A_j) \leq \mathbf{P}(X_i \in A_i) \mathbf{P}(X_j \in A_j)$$

*for  $A_i, A_j \in \mathcal{A}$ , and that  $(\mu_n(A), A \in \mathcal{A})$  is separable. Then for any  $\varepsilon > 0$  and  $n \geq 2/\varepsilon^2$ ,*

$$(8) \quad \mathbf{P} \left( \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu(A)| > \varepsilon \right) \leq 2 \mathbf{P} \left( \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu'_n(A)| > \varepsilon/2 \right).$$

*Furthermore,*

$$(9) \quad \mathbf{E} \left[ \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu(A)| \right] \leq \mathbf{E} \left[ \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu'_n(A)| \right].$$

*Proof.* Let  $X'_1, \dots, X'_n$  be an independent copy of  $X_1, \dots, X_n$ : a vector of random variables independent of the first one but with the same dependence structure. Let  $\mu'_n(A) = n^{-1} \sum 1_{X'_i \in A}$  and recall that under (pairwise) negative dependence,

$$(10) \quad \mathbf{P}(X_i \leq x_i \cap X_j \leq x_j) \leq \mathbf{P}(X_i \leq x_i) \mathbf{P}(X_j \leq x_j).$$

Our hypothesis is slightly more general, since we do not assume negative dependence but only that

$$(11) \quad \mathbf{P}(X_i \in A_i \cap X_j \in A_j) \leq \mathbf{P}(X_i \in A_i) \mathbf{P}(X_j \in A_j).$$

Negative dependence and sets of the form  $(-\infty, z]$  will suffice for the assumption (11).

Main steps of the proof are as follows: it is easy to see that under our assumptions,

$$(12) \quad V[\mu_n(A) - \mu(A)] \leq \frac{\mu(A)(1 - \mu(A))}{n}.$$

In fact, we have  $\mathbf{E}[\mu_n(A) - \mu(A)] = 0$ , which implies

$$V[\mu_n(A) - \mu(A)] = \mathbf{E} [(\mu_n(A) - \mu(A))^2].$$

Furthermore,

$$\begin{aligned}
\mathbb{E} [(\mu_n(A) - \mu(A))^2] &= \mathbb{E} \left[ \frac{1}{n^2} \sum_{i,j} 1_{X_i \in A} 1_{X_j \in A} \right] - \mu^2(A) \\
&= \frac{1}{n^2} \sum_{i,j} \mathbb{E}[1_{X_i \in A} 1_{X_j \in A}] - \mu^2(A) \\
&= \frac{1}{n^2} \sum_{i,j} \mathbb{P}(X_i \in A, X_j \in A) - \mu^2(A) \\
&= \frac{1}{n^2} \left( \sum_{i \neq j} \mathbb{P}(X_i \in A, X_j \in A) + \sum_{i=1}^n \mathbb{P}(X_i \in A) - n^2 \mu^2(A) \right) \\
&\leq \frac{1}{n^2} \left( \sum_{i \neq j} \mathbb{P}(X_i \in A) \mathbb{P}(X_j \in A) + n\mu(A) - n^2 \mu^2(A) \right),
\end{aligned}$$

where we used inequality (11) in the last step. The last expression is easily seen to be equal to  $\mu(A)(1 - \mu(A))/n$ , as desired; hence (12) is true.

We can now apply Chebyshev's inequality to the random variable  $(\mu_n(A) - \mu(A))$ , together with inequality (12), to obtain

$$\begin{aligned}
\mathbb{P} \left( |\mu_n(A) - \mu(A)| > \frac{\varepsilon}{2} \right) &\leq \frac{4}{\varepsilon^2} V[(\mu_n(A) - \mu(A))] \\
&\leq \frac{4}{\varepsilon^2} \frac{\mu(A)(1 - \mu(A))}{n} \\
&\leq \frac{1}{n\varepsilon^2} \leq \frac{1}{2} \quad \text{for all } n \geq \frac{2}{\varepsilon^2}.
\end{aligned}$$

We can then apply Lemma 1, since we have separability. Hence, for  $n \geq 2/\varepsilon^2$ ,

$$(13) \quad \mathbb{P} \left( \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu(A)| > \varepsilon \right) \leq 2 \mathbb{P} \left( \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu'_n(A)| > \varepsilon/2 \right).$$

To see the second inequality, by Jensen's inequality and the law of iterated expectation we get

$$\begin{aligned}
\mathbb{E} \left[ \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu(A)| \right] &\leq \mathbb{E} \left[ \sup_{A \in \mathcal{A}} \mathbb{E}[|\mu_n(A) - \mu'_n(A)| \mid X_1, \dots, X_n] \right] \\
&\leq \mathbb{E} \left[ \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu'_n(A)| \right].
\end{aligned} \quad \square$$

If we restrict ourselves to  $n = 2$ , we can provide a further randomization argument:

**Theorem 4.** *Let  $X_1, X_2$  be negatively dependent random variables, and  $X'_1, X'_2$  an independent copy. Let  $\sigma_1$  and  $\sigma_2$  be independent sign variables such that*

$$\mathbb{P}(\sigma_i = 1) = \mathbb{P}(\sigma_i = -1) = \frac{1}{2}.$$

*Let  $\mu(A) = \mathbb{P}(X_i \in A)$  and  $\mu_2(A) = \frac{1}{2} \sum 1_{x_i \in A}$ . Suppose that there exists  $\mathcal{A}$ , a class of sets such that  $\mathbb{P}(X_1 \in A, X_2 \in A) \leq \mathbb{P}(X_1 \in A) \mathbb{P}(X_2 \in A)$  for  $A \in \mathcal{A}$ , and that  $(\mu_2(A), A \in \mathcal{A})$  is separable. Then for any  $A \in \mathcal{A}$ ,*

$$\mathbb{P} \left( |\mu_2(A) - \mu'_n(A)| > \frac{\varepsilon}{2} \right) \leq \mathbb{P} \left( \frac{1}{2} \left| \sum \sigma_i (1_{X_i \in A} - 1_{X'_i \in A}) \right| > \frac{\varepsilon}{2} \right).$$

To prove Theorem 4, we need a preparatory lemma, whose results may be of interest per se when working with negatively dependent random variables.

**Lemma 3.** *Suppose that, for any  $A \in \mathcal{A}$ ,  $\mathbb{P}(X_2 \in A, X_1 \in A) \leq \mathbb{P}(X_2 \in A)\mathbb{P}(X_1 \in A)$ . Then*

$$\mathbb{P}(X_2 \in A \mid X_1 \in A) \leq \mathbb{P}(X_2 \in A),$$

*while, if at least one of the two marginal probabilities is smaller than 1,*

$$\mathbb{P}(X_2 \in A \mid X_1 \notin A) \geq \mathbb{P}(X_2 \in A).$$

*Moreover,*

$$\mathbb{P}(X_2 \in A, X_1 \notin A) \geq \mathbb{P}(X_2 \in A)\mathbb{P}(X_1 \notin A),$$

*and finally*

$$\mathbb{P}(X_2 \notin A, X_1 \notin A) \leq \mathbb{P}(X_2 \notin A)\mathbb{P}(X_1 \notin A).$$

*Proof.* The first inequality follows from the definition of conditional probability. To see the second, we can apply the Bayes theorem and the first inequality to get:

$$\begin{aligned} \mathbb{P}(X_2 \in A \mid X_1 \notin A) &= \mathbb{P}(X_1 \notin A \mid X_2 \in A) \mathbb{P}(X_2 \in A) / \mathbb{P}(X_1 \notin A) \\ &= (1 - \mathbb{P}(X_1 \in A \mid X_2 \in A)) \mathbb{P}(X_2 \in A) / \mathbb{P}(X_1 \notin A) \\ &\geq (1 - \mathbb{P}(X_1 \in A)) \mathbb{P}(X_2 \in A) / \mathbb{P}(X_1 \notin A) \\ &= \mathbb{P}(X_2 \in A), \end{aligned}$$

assuming without loss of generality that  $\mathbb{P}(X_1 \in A) < 1$ . The third inequality follows from

$$\begin{aligned} \mathbb{P}(X_2 \in A, X_1 \notin A) &= \mathbb{P}(X_1 \notin A \mid X_2 \in A) \mathbb{P}(X_2 \in A) \\ &= (1 - \mathbb{P}(X_1 \in A \mid X_2 \in A)) \mathbb{P}(X_2 \in A) \\ &\geq (1 - \mathbb{P}(X_1 \in A)) \mathbb{P}(X_2 \in A). \end{aligned}$$

The last inequality follows from

$$\begin{aligned} \mathbb{P}(X_2 \notin A, X_1 \notin A) &= (1 - \mathbb{P}(X_2 \in A \mid X_1 \notin A)) \mathbb{P}(X_1 \notin A) \\ &\leq (1 - \mathbb{P}(X_2 \in A)) \mathbb{P}(X_1 \notin A). \end{aligned} \quad \square$$

We can now prove Theorem 4.

*Proof.* Let  $X'_1, X'_2$  be an independent copy of  $X_1, X_2$ . In view of Theorem 3, we claim that

$$\mathbb{P}(|\mu_2(A) - \mu(A)| > \varepsilon) \leq 2\mathbb{P}(|\mu_2(A) - \mu'_n(A)| > \varepsilon/2),$$

since that is true if one takes the supremum over  $A \in \mathcal{A}$  inside.

Now let  $\sigma_1$  and  $\sigma_2$  be sign variables. We will see that a randomization argument can be used here. A possibility for further work is to try and extend the randomization argument to arbitrary  $n$ . We have

$$(14) \quad \mathbb{P}\left(|\mu_2(A) - \mu'_n(A)| > \frac{\varepsilon}{2}\right) \leq \mathbb{P}\left(\frac{1}{2} \left| \sum \sigma_i (1_{X_i \in A} - 1_{X'_i \in A}) \right| > \frac{\varepsilon}{2}\right).$$

If the random variables are independent, equality follows from identical distribution after randomization. Under dependence, identical distribution is not guaranteed any more.

To prove (14), let

$$Y_i = 1_{X_i \in A} - 1_{X'_i \in A}.$$

We now show that  $P(Y_2 = 1, Y_1 = 1) = P(Y_2 = -1, Y_1 = -1)$ :

$$\begin{aligned}
P(Y_2 = 1, Y_1 = 1) &= P(Y_2 = -1, Y_1 = -1) \\
&\iff P(Y_2 = 1 \mid Y_1 = 1) = P(Y_2 = -1 \mid Y_1 = -1) \\
&\iff P(X_2 \in A, X_2' \notin A \mid X_1 \in A, X_1' \notin A) \\
&\quad = P(X_2 \notin A, X_2' \in A \mid X_1 \notin A, X_1' \in A) \\
&\iff P(X_2 \in A \mid X_1 \in A) P(X_2' \notin A \mid X_1' \notin A) \\
&\quad = P(X_2 \notin A \mid X_1 \notin A) P(X_2' \in A \mid X_1' \in A).
\end{aligned}$$

The last equality is true since we took  $X_1', X_2'$  to be a copy of  $X_1, X_2$  with the same dependency structure. With the same strategy it can be seen that, for  $k = \{-1, 0, 1\}$ ,

$$\begin{aligned}
P(Y_2 = k, Y_1 = k) &= P(Y_2 = -k, Y_1 = -k), \\
P(Y_2 = k, Y_1 = -k) &= P(Y_2 = -k, Y_1 = k),
\end{aligned}$$

and finally that

$$P(Y_2 = k, Y_1 = 0) = P(Y_2 = -k, Y_1 = 0).$$

On the other hand,

$$(15) \quad P(Y_2 = 1, Y_1 = 1) \leq P(Y_2 = -1, Y_1 = 1).$$

We can, in fact, apply the results of Lemma 3 to show that:

$$\begin{aligned}
P(Y_2 = 1, Y_1 = 1) &= P(X_2 \in A, X_2' \in A, X_1 \notin A, X_1' \notin A) \\
&= P(X_2 \in A, X_1 \in A) P(X_2' \notin A, X_1' \notin A) \\
&\leq P(X_2 \in A) P(X_2' \notin A) P(X_1 \in A) P(X_1' \notin A)
\end{aligned}$$

and

$$\begin{aligned}
P(Y_2 = 1, Y_1 = -1) &= P(X_2 \in A, X_2' \notin A, X_1 \notin A, X_1' \in A) \\
&= P(X_2 \in A, X_1 \notin A) P(X_2' \notin A, X_1' \in A) \\
&\geq P(X_2 \in A) P(X_2' \notin A) P(X_1 \notin A) P(X_1' \in A).
\end{aligned}$$

By identical distribution of  $X_i$  and  $X_i'$ , (15) follows.

With a few explicit calculations the results on the joint distribution of  $(Y_1, Y_2)$  can be used to compute the distribution of the empirical measure

$$\begin{aligned}
P(|\mu_2(A) - \mu_2'(A)| = 1) &= P(Y_1 + Y_2 = 2) + P(Y_1 + Y_2 = -2) \\
&= 2P(Y_1 = 1, Y_2 = 1).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
P\left(\left|\sum \sigma_i (1_{X_i \in A} - 1_{X_i' \in A})\right| = 2\right) &= 2P(\sigma_1 Y_1 = 1, \sigma_2 Y_2 = 1) \\
&= 2[P(Y_1 = 1, Y_2 = 1) P(\sigma_1 = \sigma_2) + P(\sigma_1 \neq \sigma_2) P(Y_1 = 1, Y_2 = -1)] \\
&\geq 2P(Y_1 = 1, Y_2 = 1).
\end{aligned}$$

Moreover,

$$P(|\mu_2(A) - \mu_2'(A)| = 0) = P(Y_1 = 0, Y_2 = 0) + 2P(Y_1 = 1, Y_2 = -1),$$

while

$$\begin{aligned}
P(|\sigma_1 Y_1 + \sigma_2 Y_2| = 0) &= P(Y_1 = 0, Y_2 = 0) + 2P(\sigma_1 = \sigma_2) P(Y_1 = 1, Y_2 = -1) \\
&\quad + 2P(\sigma_1 \neq \sigma_2) P(Y_1 = 1, Y_2 = 1) \\
&\leq P(Y_1 = 0, Y_2 = 0) + 2P(Y_1 = 1, Y_2 = -1).
\end{aligned}$$



The results are combined to see that

$$\mathbb{P}\left(|\mu_2(A) - \mu'_n(A)| > \frac{\varepsilon}{2}\right) \leq \mathbb{P}\left(\frac{1}{2} \left| \sum \sigma_i (1_{X_i \in A} - 1_{X'_i \in A}) \right| > \frac{\varepsilon}{2}\right),$$

that is, inequality (14).  $\square$

### 3. DISCUSSION

We proved that negatively dependent random variables enjoy certain special properties of independent random variables, in particular in terms of Hoeffding and bounded difference inequalities and of the possibility to apply symmetrization. These tools pave the road to inequalities for the empirical distribution of negatively dependent random variables.

**Acknowledgements.** The author is grateful to Professor Enzo Orsingher for advice and encouragement, and to a referee for clarifying review and pointing out the reference [1].

### BIBLIOGRAPHY

1. M. D. Amini and A. Bozorgnia, *Complete convergence for negatively dependent random sequences*, Journal of Applied Mathematics and Stochastic Analysis **16** (2003), 121–126. MR1989578 (2004j:60037)
2. H. W. Block, T. H. Savits, and M. Shaked, *Some concepts of negative dependence*, The Annals of Probability **10** (1982), 765–772. MR659545 (83i:60015)
3. J. D. Esary, F. Proschan, and D. W. Walkup, *Association of random variables, with applications*, The Annals of Mathematical Statistics **38** (1967), 1466–1474. MR0217826 (36:915)
4. J. D. Kumar and F. Proschan, *Negative association of random variables with applications*, The Annals of Statistics **11** (1983), 286–295. MR684886 (85d:62058)
5. C. McDiarmid, *On the method of bounded differences*, Surveys in Combinatorics, Cambridge University Press, 1989, pp. 148–188. MR1036755 (91e:05077)
6. A. Volodin, *On the Kolmogorov exponential inequality for negatively dependent random variables*, Pakistan Journal of Statistics **18** (2002), 249–253. MR1944611 (2003k:60053)
7. Y. L. Tong, *Probability Inequalities in Multivariate Distributions*, Academic Press, 1980. MR572617 (82k:60038)

UNIVERSITY OF ROME “LA SAPIENZA”, PIAZZALE ALDO MORO 5, 00185 ROME, ITALY  
*E-mail address:* `alessio.farcomeni@uniroma1.it`

Received 10/AUG/2006  
 Originally published in English