

## A NOTE ON SOME INEQUALITIES FOR CERTAIN CLASSES OF POSITIVELY DEPENDENT RANDOM VARIABLES

BY

PRZEMYSŁAW MATUŁA (LUBLIN)

*Abstract.* We prove an inequality for the difference between a joint distribution and the product of its marginals for certain classes of positively dependent random variables. Some multivariate extensions are also discussed.

**1991 Mathematics Subject Classification:** Primary 60E15, 62H20, 62G30.

**Key words and phrases:** Positive dependence, copula, covariance.

### 1. INTRODUCTION

Let  $X$  and  $Y$  be absolutely continuous random variables (r.v.'s) with densities  $f_X(x)$  and  $f_Y(y)$ , let us denote by  $F_{X,Y}(x, y)$ ,  $F_X(x)$ ,  $F_Y(y)$  their joint distribution function (d.f.) and marginal d.f.'s. Furthermore we introduce the following notation:

$$\begin{aligned} H_{X,Y}(x, y) &= F_{X,Y}(x, y) - F_X(x)F_Y(y) \\ &= P(X \leq x, Y \leq y) - P(X \leq x)P(Y \leq y). \end{aligned}$$

The random variables  $X$  and  $Y$  are said to be *positively quadrant dependent* (PQD) if (cf. [4])

$$H_{X,Y}(x, y) \geq 0 \quad \text{for all } x, y \in \mathbb{R}.$$

In [5] the following inequality for PQD r.v.'s has been established:

$$(1.1) \quad \sup_{x,y \in \mathbb{R}} H_{X,Y}(x, y) \leq C \cdot \text{Cov}_H^{1/3}(X, Y),$$

where  $C = \frac{3}{2} \sqrt[3]{2} (\|f_X\|_\infty + \|f_Y\|_\infty)^{2/3}$  and

$$\text{Cov}_H(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_{X,Y}(x, y) dx dy$$

is the so-called Hoeffding covariance (if the usual product moment covariance exists, then it is equal to the Hoeffding covariance). Here and in the sequel

$$\|f\|_{\infty} = \inf \{M: \mu \{x \in R: |f(x)| > M\} = 0\}$$

is the essential supremum of  $f$  with respect to the Lebesgue measure  $\mu$  on  $R$ .

The inequalities of the type (1.1) found the applications in kernel estimation of the density and distribution function and in studying convergence of empirical processes for positively dependent r.v.'s (cf. [9] and [5] where further references are given). Our goal is to show that the inequality (1.1) may be improved for certain classes of PQD r.v.'s. Let us recall some concepts of positive dependence for pairs of r.v.'s which will be considered in the sequel (cf. [2] and [7]).

DEFINITION 1.

•  $Y$  is *stochastically increasing in  $X$*  (in symbols,  $SI(Y|X)$ ) if  $P(Y > y | X = x)$  is a nondecreasing function of  $x$  for all  $y$ .

•  $X$  is *stochastically increasing in  $Y$*  ( $SI(X|Y)$ ) if  $P(X > x | Y = y)$  is a nondecreasing function of  $y$  for all  $x$ .

• The r.v.'s  $X$  and  $Y$  with joint density  $f_{X,Y}(x, y)$  are *positively likelihood ratio dependent* ( $PLR(X, Y)$ ) if

$$f_{X,Y}(x, y) f_{X,Y}(x', y') \geq f_{X,Y}(x, y') f_{X,Y}(x', y)$$

for all  $x, y, x', y' \in R$  such that  $x \leq x'$  and  $y \leq y'$ ; in this case the density is said to be *totally positive of order 2* ( $TP_2$ ) and it is said that the r.v.'s are  $TP_2$ .

• The r.v.'s  $X$  and  $Y$  are *associated* ( $A$ ) if

$$\text{Cov}(f(X, Y), g(X, Y)) \geq 0$$

for any coordinatewise nondecreasing functions  $f, g: R^2 \rightarrow R$  for which this covariance exists.

Let us recall that the following implications between these concepts of dependence hold (cf. [7]):

PROPOSITION 1.

•  $TP_2 \Rightarrow SI(X|Y) \Rightarrow A$ .

•  $TP_2 \Rightarrow SI(Y|X) \Rightarrow A$ .

•  $A \Rightarrow PQD$ .

Let us note that for continuous r.v.'s  $X$  and  $Y$ , by Sklar's theorem (cf. [7]), there exists a unique function  $C(u, v)$ ,  $u, v \in I^2 := [0, 1]^2$  such that

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)).$$

This function is called the *copula* of  $X$  and  $Y$  and it is a distribution function on  $I^2$  with uniform marginals. The inequalities of the form (1.1) may be obtained

by studying the related inequalities for the copula of  $X$  and  $Y$ . At first let us observe that

$$(1.2) \quad \sup_{x,y \in R} H_{X,Y}(x, y) = \max_{u,v \in I} (C(u, v) - uv).$$

Furthermore, by the equality (2.12) of Yu (cf. [10]) for PQD r.v.'s we have

$$(1.3) \quad \int_0^1 \int_0^1 (C(u, v) - uv) dudv = \text{Cov}(F_X(X), F_Y(Y)) \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F'_X(x) F'_Y(y) H_{X,Y}(x, y) dx dy \leq \|f_X\|_{\infty} \|f_Y\|_{\infty} \text{Cov}_H(X, Y).$$

In view of (1.2) and (1.3) we shall obtain bounds for  $\max_{u,v \in I} (C(u, v) - uv)$  in terms of  $\int_0^1 \int_0^1 (C(u, v) - uv) dudv$ ; therefore we need some additional properties of the copula under dependence conditions mentioned in Definition 1 (cf. [7]).

**PROPOSITION 2.**

- $X, Y$  are PQD iff  $C(u, v) \geq uv$  for all  $u, v \in I^2$ .
- $X, Y$  are SI( $Y|X$ ) iff for any  $v \in I$ ,  $C(u, v)$  is a concave function of  $u$ .
- $X, Y$  are SI( $X|Y$ ) iff for any  $u \in I$ ,  $C(u, v)$  is a concave function of  $v$ .

From the next proposition it follows that the  $TP_2$  property in some cases may be easily verified (cf. [6], p. 128).

**PROPOSITION 3.** Assume that  $X$  and  $Y$  have a twice differentiable positive density  $f_{X,Y}(x, y)$ . Then  $X, Y$  are  $TP_2$  iff

$$\frac{\partial^2}{\partial x \partial y} \ln f_{X,Y}(x, y) \geq 0.$$

## 2. THE MAIN RESULTS

**THEOREM 1.** Let  $X$  and  $Y$  be absolutely continuous r.v.'s which are both SI( $X|Y$ ) and SI( $Y|X$ ). Then

$$(2.1) \quad \sup_{x,y \in R} H_{X,Y}(x, y) \leq 4 \|f_X\|_{\infty} \|f_Y\|_{\infty} \text{Cov}_H(X, Y).$$

**Proof.** Let  $C(u, v)$  be the copula of  $X$  and  $Y$ . Obviously,  $C(u, v) - uv$  is nonnegative, continuous and  $C(u, v) - uv = 0$  on the boundary of  $I^2$ . The case  $C(u, v) - uv \equiv 0$  is trivial, so that we may and do assume that for some  $(u_0, v_0) \in (0, 1)^2$

$$(2.2) \quad a := C(u_0, v_0) - u_0 v_0 = \max_{u,v \in I} (C(u, v) - uv).$$

It is easy to see that, by Proposition 2, also  $C(u, v) - uv$  is concave with respect to  $u$  for fixed  $v$ . Thus we have

$$(2.3) \quad C(u, v_0) - uv_0 \geq \begin{cases} \frac{au}{u_0} & \text{for } u \in \langle 0, u_0 \rangle, \\ \frac{a(1-u)}{1-u_0} & \text{for } u \in \langle u_0, 1 \rangle. \end{cases}$$

Finally, by concavity of  $C(u, v) - uv$  with respect to  $v$  for fixed  $u$  and by (2.3), we get

$$(2.4) \quad C(u, v) - uv \geq \frac{auv}{u_0 v_0} \quad \text{for } u \in \langle 0, u_0 \rangle, v \in \langle 0, v_0 \rangle,$$

$$C(u, v) - uv \geq \frac{au(1-v)}{u_0(1-v_0)} \quad \text{for } u \in \langle 0, u_0 \rangle, v \in \langle v_0, 1 \rangle,$$

$$C(u, v) - uv \geq \frac{a(1-u)v}{(1-u_0)v_0} \quad \text{for } u \in \langle u_0, 1 \rangle, v \in \langle 0, v_0 \rangle,$$

$$C(u, v) - uv \geq \frac{a(1-u)(1-v)}{(1-u_0)(1-v_0)} \quad \text{for } u \in \langle u_0, 1 \rangle, v \in \langle v_0, 1 \rangle.$$

Therefore, the inequalities (2.4) and elementary calculations yield

$$\int_0^1 \int_0^1 (C(u, v) - uv) dudv \geq \frac{a}{4}.$$

Consequently,

$$\max_{u, v \in I} (C(u, v) - uv) \leq 4 \int_0^1 \int_0^1 (C(u, v) - uv) dudv,$$

and the conclusion follows from (1.2) and (1.3). ■

From the relation between  $TP_2$  and  $SI$  described in Proposition 1 we get the following corollary:

**COROLLARY 1.** *If  $X$  and  $Y$  are  $TP_2$ , then (2.1) holds true.*

It is well known (cf. [2]) that positively correlated Gaussian r.v.'s are  $TP_2$  (this may be directly checked by Proposition 3); furthermore  $H_{X,Y}(x, y) = -H_{X,-Y}(x, -y)$ , so that from Corollary 1 we easily derive the following inequality for Gaussian vectors.

**THEOREM 2.** *Assume that  $X$  and  $Y$  have the joint normal distribution with the correlation coefficient  $\rho \in (-1, 1)$ . Then*

$$(2.5) \quad \sup_{x, y \in \mathbb{R}} |H_{X,Y}(x, y)| \leq \frac{2|\rho|}{\pi}.$$

The inequality (2.1) cannot be improved; let us consider the following example.

EXAMPLE 1. Consider the r.v.'s  $X$  and  $Y$  with joint density with uniform marginals of the form  $f(x, y) = 1 + h(x)h(y)$  for  $(x, y) \in I^2$ , where

$$h(x) = \begin{cases} a & \text{if } x \in \langle 0, \frac{1}{2} \rangle, \\ -a & \text{if } x \in \langle \frac{1}{2}, 1 \rangle, \\ 0 & \text{otherwise} \end{cases}$$

for some  $a \in \langle 0, 1 \rangle$ . The distribution (equal to the copula) of  $X$  and  $Y$  takes the form  $F(x, y) = xy + H(x)H(y)$  for  $(x, y) \in I^2$ , where  $H(x) = \int_0^x h(t) dt$ . Then

$$\sup_{x, y \in R} |F(x, y) - xy| = (\max_{x \in I} H(x))^2 = a^2/4$$

and

$$\text{Cov}(X, Y) = \int_0^1 \int_0^1 H(t)H(s) dt ds = a^2/16,$$

so that we have equality in (2.1). It is easy to see that, by Proposition 2, the r.v.'s are  $SI(Y|X)$  and  $SI(X|Y)$ .

Now let us introduce another family of PQD copulas. Let us denote by  $\mathcal{B}$  a family of measurable functions  $h: I \rightarrow R$  such that:

$$(2.6) \quad \|h\|_\infty \leq 1,$$

$$(2.7) \quad \int_0^1 h(t) dt = 0,$$

$$(2.8) \quad \int_0^x h(t) dt \geq 0 \quad \text{for } x \in I.$$

It is easy to see that for any  $h_1, h_2 \in \mathcal{B}$  the function  $f(x, y) = 1 + h_1(x)h_2(y)$  for  $(x, y) \in I^2$  and  $f(x, y) = 0$  otherwise is a density function with the PQD property and uniform marginals with copula

$$(2.9) \quad F(x, y) = xy + H_1(x)H_2(y), \quad \text{where } H_i(x) = \int_0^x h_i(t) dt, \quad i = 1, 2.$$

Let us observe that  $H_i(0) = H_i(1) = 0$  and  $H_i$  is nonnegative and continuous on  $I$  so that it attains its maximum value at some point  $x_i \in (0, 1)$ ,  $\sup_{x \in I} H_i(x) = H_i(x_i) = a_i$ , say. By properties (2.6)–(2.8) we get for  $i = 1, 2$

$$0 \leq x_i - a_i \leq x_i \leq x_i + a_i \leq 1$$

and

$$H_i(s) \geq \begin{cases} a_i - x_i + s & \text{for } s \in [x_i - a_i, x_i], \\ a_i + x_i - s & \text{for } s \in [x_i, x_i + a_i]. \end{cases}$$

Thus

$$\int_0^1 H_i(t) dt \geq \int_{x_i - a_i}^{x_i + a_i} H_i(t) dt \geq a_i^2.$$

Therefore

$$\sup_{x, y \in I} (F(x, y) - xy) = \sup_{x \in I} H_1(x) \sup_{y \in I} H_2(y) = a_1 a_2$$

and

$$\int_0^1 \int_0^1 (F(x, y) - xy) dx dy = \int_0^1 H_1(x) dx \int_0^1 H_2(x) dx \geq (a_1 a_2)^2.$$

Thus we have the following inequality:

$$\sup_{x, y \in I} (F(x, y) - xy) \leq \left( \int_0^1 \int_0^1 (F(x, y) - xy) dx dy \right)^{1/2}.$$

From the consideration presented above, (1.2) and (1.3) we get the following theorem:

**THEOREM 3.** *Let  $X$  and  $Y$  be absolutely continuous r.v.'s with copula of the form (2.9) for some  $h_1, h_2 \in \mathcal{B}$ . Then*

$$(2.10) \quad \sup_{x, y \in \mathbb{R}} H_{X, Y}(x, y) \leq \|f_X\|_\infty^{1/2} \|f_Y\|_\infty^{1/2} \text{Cov}_H^{1/2}(X, Y).$$

In order to see that the above inequality cannot be improved let us consider the following example.

**EXAMPLE 2.** Let

$$h_1(x) = h_2(x) = \begin{cases} 1 & \text{if } x \in [0, a), \\ -1 & \text{if } x \in [a, 2a), \\ 0 & \text{if } x \in [2a, 1] \end{cases}$$

for some  $a \in [0, \frac{1}{4}]$ . It is easily checked that in this case, for the random variables  $X$  and  $Y$  with the joint distribution of the form (2.9), we have

$$\sup_{x, y \in \mathbb{R}} H_{X, Y}(x, y) = \text{Cov}_H^{1/2}(X, Y).$$

## 3. MULTIVARIATE EXTENSIONS

In this section we extend the results obtained in Section 2 to the multivariate case. We need some concepts of positive dependence for families of random variables. The random variables  $X_1, X_2, \dots, X_n$  are *associated* if

$$\text{Cov}(f(X_1, \dots, X_n), g(X_1, \dots, X_n)) \geq 0$$

for any coordinatewise nondecreasing functions  $f, g: R^n \rightarrow R$  for which this covariance exists. The random variables  $X_1, X_2, \dots, X_n$  are *multivariate totally positive of order 2* ( $MTP_2$ ) (cf. [3]) if their joint density  $f$  satisfies the following inequality:

$$f(x \vee y) f(x \wedge y) \geq f(x) f(y),$$

where

$$x = (x_1, \dots, x_n), \quad y = (y_1, \dots, y_n)$$

and

$$x \vee y = (\max(x_1, y_1), \dots, \max(x_n, y_n)), \quad x \wedge y = (\min(x_1, y_1), \dots, \min(x_n, y_n)).$$

Let us note that  $MTP_2$  r.v.'s are associated, and r.v.'s with joint strictly positive density which are  $TP_2$  in pairs are also  $MTP_2$ . In view of Proposition 3 it is often easier to check the  $MTP_2$  property than association.

Now let us prove a generalization of the Lebowitz inequality (cf. [8]).

**THEOREM 4.** *If the random variables  $X_1, \dots, X_n$  are associated and  $A_1, \dots, A_k$  are disjoint subsets of  $\{1, 2, \dots, n\}$ , then for any  $x_i \in R, i \in A_1 \cup \dots \cup A_k$ ,*

$$(3.1) \quad 0 \leq P(X_i \leq x_i, i \in A_1 \cup \dots \cup A_k) - \prod_{i=1}^k P(X_j \leq x_j, j \in A_i) \\ \leq \sum_{1 \leq i < j \leq k} \sum_{r \in A_i, s \in A_j} H_{X_r, X_s}(x_r, x_s).$$

*In particular,*

$$0 \leq P(X_1 \leq x_1, \dots, X_n \leq x_n) - \prod_{i=1}^n P(X_i \leq x_i) \leq \sum_{1 \leq i < j \leq n} H_{X_i, X_j}(x_i, x_j).$$

**Proof.** Let us put  $I_{A_i} = I(X_j \leq x_j, j \in A_i)$ . Then the left-hand side of (3.1) may be expressed and bounded as follows:

$$P(X_i \leq x_i, i \in A_1 \cup \dots \cup A_k) - \prod_{i=1}^k P(X_j \leq x_j, j \in A_i) = E \prod_{i=1}^k I_{A_i} - \prod_{i=1}^k E I_{A_i} \\ = \sum_{i=1}^{k-1} \left( \prod_{m=0}^{i-1} E I_{A_m} \right) \text{Cov}(I_{A_i}, \prod_{j=i+1}^k I_{A_j}) \leq \sum_{i=1}^{k-1} \text{Cov}(I_{A_i}, \prod_{j=i+1}^k I_{A_j}),$$

where  $I_{A_0} = 1$ . Let us introduce the following notation:

$$S_1 = \sum_{r \in A_i} I(X_r \leq x_r), \quad S_2 = \sum_{j=i+1}^k \sum_{s \in A_j} I(X_s \leq x_s).$$

Then  $S_1 - I_{A_i}$ ,  $S_1 + I_{A_i}$ ,  $S_2 - \prod_{j=i+1}^k I_{A_j}$ ,  $S_2 + \prod_{j=i+1}^k I_{A_j}$  are nonincreasing functions of  $X_1, \dots, X_n$ . Thus, by the properties of associated r.v.'s, we get

$$\begin{aligned} 0 &\leq \text{Cov}(S_1 + I_{A_i}, S_2 - \prod_{j=i+1}^k I_{A_j}) + \text{Cov}(S_1 - I_{A_i}, S_2 + \prod_{j=i+1}^k I_{A_j}) \\ &= 2\text{Cov}(S_1, S_2) - 2\text{Cov}(I_{A_i}, \prod_{j=i+1}^k I_{A_j}). \end{aligned}$$

Consequently,

$$\text{Cov}(I_{A_i}, \prod_{j=i+1}^k I_{A_j}) \leq \sum_{j=i+1}^k \sum_{r \in A_i, s \in A_j} \text{Cov}(I(X_r \leq x_r), I(X_s \leq x_s))$$

and the proof is completed. ■

It is easily seen that Theorem 4 and Corollary 1 yield the following result.

**COROLLARY 2.** *Let  $X_1, \dots, X_n$  be  $MTP_2$  with bounded densities. Then*

$$\begin{aligned} 0 &\leq P(X_1 \leq x_1, \dots, X_n \leq x_n) - \prod_{i=1}^n P(X_i \leq x_i) \\ &\leq 4 \sum_{1 \leq i < j \leq n} \|f_{X_i}\|_\infty \|f_{X_j}\|_\infty \text{Cov}_H(X_i, X_j). \end{aligned}$$

It is known that Gaussian random vectors positively correlated are  $MTP_2$ . Therefore, by Theorems 2 and 4, we have the next corollary.

**COROLLARY 3.** *Assume that  $[X_1, \dots, X_n]$  is a Gaussian vector and the r.v.'s  $X_i$  are positively correlated with correlation coefficients  $\rho_{ij} = \text{Corr}(X_i, X_j)$ . Then*

$$0 \leq P(X_1 \leq x_1, \dots, X_n \leq x_n) - \prod_{i=1}^n P(X_i \leq x_i) \leq \frac{2}{\pi} \sum_{1 \leq i < j \leq n} \rho_{ij}.$$

#### 4. CONCLUDING REMARKS

It was mentioned in the Introduction that inequalities of type (1.1) are the main tool in estimation of the density and distribution function for associated sequences. It is easy to see that in many cases one can replace the assumption of association by a stronger  $MTP_2$  property and apply the inequality (2.1) instead of (1.1); this results in relaxing the conditions on the covariance structure of the considered sequence.



As an example let us state the result corresponding to the one presented in [1].

**THEOREM 5.** Let  $(X_n)_{n \in \mathbb{N}}$  be a stationary sequence of  $MTP_2$  random variables with bounded densities and common distribution function  $F(x)$ . Denote by  $F_n(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x)$  the empirical distribution function. If

$$\sum_{j=n+1}^{\infty} \text{Cov}_H(X_1, X_j) = o(1),$$

then

$$n^{1/2} (F_n(x) - F(x)) / \sigma(x) \xrightarrow{d} N(0, 1)$$

for all  $x$  such that  $0 < F(x) < 1$ , where

$$\sigma^2(x) = F(x)(1 - F(x)) + 2 \sum_{j=2}^{\infty} (P(X_1 \leq x, X_j \leq x) - F^2(x)).$$

If, for some  $r > 0$ ,

$$\sum_{j=n+1}^{\infty} \text{Cov}_H(X_1, X_j) = O(n^{-r+1}),$$

then there exists some constant  $C > 0$  such that

$$\sup_{x \in \mathbb{R}} P(|F_n(x) - F(x)| \geq \varepsilon) \leq C \varepsilon^{-2r} n^{-r}$$

for every  $\varepsilon > 0$  and  $n \geq 1$ .

**Proof.** Let us observe that for fixed  $x$  the random variables  $I(X_1 \leq x)$ ,  $I(X_2 \leq x)$ , ... are associated. Then the inequality (2.1) may be applied and the proof runs as the proofs of Theorems 3.1 and 3.3 in [1], so we omit some obvious modifications. ■

**Remark 1.** Finally, let us observe that from our main results one can also deduce some inequalities for negatively dependent random variables. If the random variables  $X$  and  $Y$  are negatively quadrant dependent (NQD), then  $-X$  and  $Y$  are PQD, so that a result similar to that stated in Theorem 3 also holds for a family of NQD random variables. The inequality (3.1) with the reverse sign also holds for negatively associated random variables.

#### REFERENCES

- [1] I. Bagai and B. L. S. Prakasa Rao, *Estimation of the survival function for stationary associated processes*, Statist. Probab. Lett. 12 (1991), pp. 385–391.
- [2] R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing. Probability Models*, Holt, Rinehart and Winston, New York 1975.

- [3] S. Karlin and Y. Rinott, *Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions*, J. Multivariate Anal. 10 (1980), pp. 467–498.
- [4] E. L. Lehmann, *Some concepts of dependence*, Ann. Math. Statist. 37 (1966), pp. 1137–1153.
- [5] P. Matuła, *On some inequalities for positively and negatively dependent random variables with applications*, Publ. Math. Debrecen 63/4 (2003), pp. 511–521.
- [6] A. Müller and D. Stoyan, *Comparison Methods for Stochastic Models and Risks*, Wiley, 2002.
- [7] R. B. Nelsen, *An Introduction to Copulas*, Springer, New York 1999.
- [8] C. M. Newman, *Asymptotic independence and limit theorems for positively and negatively dependent random variables*, in: *Inequalities in Statistics and Probability*, Y. L. Tong (Ed.), Hayward CA, Inst. Math. Statist., 1984.
- [9] G. G. Roussas, *Kernel estimates under association – strong uniform consistency*, Statist. Probab. Lett. 12 (1991), pp. 393–403.
- [10] H. Yu, *A Glivenko–Cantelli lemma and weak convergence for empirical processes of associated sequences*, Probab. Theory Related Fields 95 (1993), pp. 357–370.

Institute of Mathematics  
Maria Curie-Skłodowska University  
pl. Marii Curie-Skłodowskiej 1  
20-031 Lublin, Poland  
E-mail: matula@golem.umcs.lublin.pl

Received on 6.12.2003

---