ON TOURNAMENTS AND NEGATIVE DEPENDENCE

YAAKOV MALINOVSKY [iD],* *University of Maryland, Baltimore County*
YOSEF RINOTT,** *The Hebrew University of Jerusalem*

Abstract

Negative dependence of sequences of random variables is often an interesting characteristic of their distribution, as well as a useful tool for studying various asymptotic results, including central limit theorems, Poisson approximations, the rate of increase of the maximum, and more. In the study of probability models of tournaments, negative dependence of participants' outcomes arises naturally, with application to various asymptotic results. In particular, the property of *negative orthant dependence* was proved in several articles for different tournament models, with a special proof for each model. In this note we unify these results by proving a stronger property, *negative association*, a generalization leading to a very simple proof. We also present a natural example of a knockout tournament where the scores are negatively orthant dependent but not negatively associated. The proof requires a new result on a preservation property of negative orthant dependence that is of independent interest.

Keywords: Negative association; negative orthant dependence; multivariate inequalities

2020 Mathematics Subject Classification: Primary 62H05
Secondary 05C20; 60E15

1. Introduction

1.1. Tournaments

A tournament consists of competitions between several players where the final score or payoff of each player is determined by the sum of the scores of the player's matches. For a tournament with n players, let $\mathbf{S} = (S_1, \ldots, S_n)$ denote the vector of their final scores. Under natural probability models and in many kinds of tournaments, the components of $\mathbf{S}$ exhibit some type of negative dependence. We briefly define two concepts of dependence to be considered in this paper and then we discuss various tournaments where these concepts are relevant. We present a theorem on negative association that unifies and strengthens known results on negative dependence of tournament scores, and leads to new ones. Specifically, we prove negative association in various models. We also analyze a tournament in which, interestingly, negative association holds when the draw of matches is random, and otherwise only a weaker notion of negative dependence, negative orthant dependence, holds.

* Postal address: Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, USA. Email: yaakovm@umbc.edu
** Postal address: Department of Statistics and Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Israel. Email: yosef.rinott@mail.huji.ac.il

1.2. Two notions of negative dependence

We define the following negative dependence notions. See [11] and references therein for details. Throughout this paper, increasing (decreasing) stands for nondecreasing (nonincreasing).

Definition 1. ([11], *Definition* 2.3) The random variables $S_1, \ldots, S_n$ or the vector $\mathbf{S} = (S_1, \ldots, S_n)$ are said to be *negatively lower orthant dependent* (NLOD) if, for all $s_1, \ldots, s_n \in \mathbb{R}$,

$$\mathbb{P}(S_1 \leq s_1, \ldots, S_n \leq s_n) \leq \mathbb{P}(S_1 \leq s_1) \cdots \mathbb{P}(S_n \leq s_n), \tag{1}$$

and *negatively upper orthant dependent* (NUOD) if

$$\mathbb{P}(S_1 > s_1, \ldots, S_n > s_n) \leq \mathbb{P}(S_1 > s_1) \cdots \mathbb{P}(S_n > s_n). \tag{2}$$

Negative orthant dependence (NOD) is said to hold if both (1) and (2) hold.

Definition 2. ([11], *Definition* 2.1) The random variables $S_1, \ldots, S_n$ or the vector $\mathbf{S} = (S_1, \ldots, S_n)$ are said to be *negatively associated* (NA) if, for every pair of disjoint subsets A_1, A_2 of $\{1, 2, \ldots, n\}$,

$$\mathrm{Cov}(f_1(S_i, i \in A_1), f_2(S_j, j \in A_2)) \leq 0 \tag{3}$$

whenever f_1 and f_2 are real-valued functions, increasing in all coordinates.

Clearly, NA implies NOD (see [11]). In Section 3.2 we provide a natural example of a tournament where $S_1, \ldots, S_n$ are NOD but not NA.

1.3. Motivation

1.3.1. *General dependence structure* The study of dependence structure between random variables and related stochastic orders is of interest in itself; see, e.g., the books [12, 13, 21], and articles such as [7, 23] which concentrate on negative dependence and its applications. Dependence models are relevant to a large number of applications, such as system reliability and risk theory [21, 30], statistical mechanics [29], asymptotic approximations [6] and nonasymptotic bounds on the difference between certain functions of dependent random variables, and to simple models with independence [2, 9], in multiple testing hypotheses [3, 25, 31, 32], various optimization problems (see, e.g., [24]), and geometric probability [22].

1.3.2. *Negative Dependence and Tournaments* Negative association and other concepts of negative dependence are relevant to tournaments, as explained below. In the present paper we unify results which appear in the literature on tournaments, and extend them to the strong notion of NA, and to general classes of tournaments.

Pemantle [23] stated that 'the property of NA is reasonably useful but hard to verify'. We provide simple tools and examples where NA is verified in the context of tournaments.

A certain tournament model (details provided in the next section) was considered in [10] where player 1 is stronger than all the other players, who are all equally strong. It was proved that $\lim_{n \to \infty} \mathbb{P}(S_1 > \max\{S_2, \ldots, S_n\}) \to 1$; i.e. player 1 achieves the highest score with probability approaching 1. The proof is based on the fact (proved by a special coupling argument) that the components of $\mathbf{S}$ are NLOD; we give a simpler proof showing the stronger property of NA. A binomial tournament model (details provided in the next section) was studied in [28], establishing bounds for $\mathbb{P}(S_i > \max_{j \neq i} S_j)$ using a stochastic ordering property which required knowledge of a certain negative dependence structure of the scores (see also [27]). The convergence in probability of the normalized maximal score to a constant for a general tournament

model (details are given in the next section) was studied in [18] by using inequalities for the joint distribution function of the scores $S_1, \ldots, S_n$; the proof required the NLOD property of the scores. The asymptotic distribution of the maximal score, second maximal, etc. in a chess round-robin tournament model (details provided in the next section) were established in [16, 17] using a nonasymptotic bound on the total variation distance between the sum of indicators that the score of player j is larger than a given constant and a suitable Poisson approximation which would hold if the indicators were independent. This bound is based on the fact that the indicators have a certain negative dependence structure. It follows that we can use classical limiting results under independence and show that the maximal score and related functional have Gumbel-type distribution in the limit. In all these examples we provide a simple proof of NA which implies the required negative dependence, and in the last example our proof holds for a complete range of the parameters, unlike the proof in [17]. Thus, we unify and simplify many existing results in the literature, extending the range of tournament models and strengthening the dependence proved.

1.4. Constant-sum round-robin tournaments

We start with a formulation of a *general constant-sum round-robin tournaments*; see, e.g., [5, 19]. Assume that each of n players competes against each of the other $n-1$ players. When player i plays against player j, where $i < j$, player i's reward is a random variable X_{ij} having a distribution function F_{ij} with support on $[0, r_{ij}]$, and $X_{ji} = r_{ij} - X_{ij}$; for $i < j$ this determines $F_{ji}(t) = 1 - F_{ij}(r_{ij} - t)$ for $t \in [0, r_{ij}]$. Thus, each pair of players competes for a share of a given reward. We assume that the X_{ij} are independent for $i < j$, and also that $r_{ij} \geq 0$. The case where $r_{ij} = 0$ has the interpretation that players i and j do not compete against each other. The total reward for player i is defined for all the tournaments we consider by $S_i = \sum_{j=1, j \neq i}^{n} X_{ij}$, $i = 1, \ldots, n$. The sum of the rewards is constant: $\sum_{i=1}^{n} S_i = \sum_{i<j} r_{ij}$.

We prove that $S_1, \ldots, S_n$ are NA (Definition 2), extending and simplifying various results in the literature (to be specified below), and, more generally, if u_i are increasing functions, it follows that that $u_1(S_1), \ldots, u_n(S_n)$ are also NA. These functions can represent the utilities of the players. See Proposition 1 for a further generalization.

1.4.1. *A round-robin tournament with integer reward* The case of the above round-robin tournament model with integer support $\{0, 1, \ldots, r_{ij}\}$ of F_{ij} was considered recently in [18]. Our results on negative dependence for the general round-robin tournament generalize the negative dependence results in [18]. Specifically, the NLOD property is proved in [18], and our general result yields the NA property with a simpler proof.

We next discuss further special cases of our general formulation that have appeared in the literature.

1.4.2. *A round-robin tournament with pairwise repeated games* Recently, [28] considered a special case of the above two models where $X_{ij} \sim \text{Binomial}(r_{ij}, p_{ij})$ independently for all $i < j$, $r_{ij} = r_{ji}$, and $X_{ji} = r_{ij} - X_{ij} \sim \text{Binomial}(r_{ij}, 1 - p_{ij})$. As always, $S_i = \sum_{j=1, j \neq i}^{n} X_{ij}$. This model arises if each pair of players (i, j) plays r_{ij} independent games, and i wins with probability p_{ij}. [28] obtained NOD-type results for general p_{ij} using log-concavity, conditioning, and Efron's well-known theorem [8]. Again, we strengthen and simplify these results and prove the NA property. The results in [28] were used to study expressions such as $\mathbb{P}(S_i > \max_{j \neq i} S_j)$ and related inequalities, under a special model for p_{ij}, given, e.g., in [4, 34].

1.4.3. *A simple round-robin tournament* The above general model was considered in [10], where, for any $i \neq j$, $X_{ij} + X_{ji} = 1$, $X_{ij} \in \{0, 1\}$, and $\mathbb{P}(X_{ij} = 1) = p_{ij}$, and proved that $S_1, \ldots, S_n$ are NLOD by invoking coupling arguments. The latter fact was then used to prove that if $\mathbb{P}(X_{1j} = 1) = p > \frac{1}{2}$, and $\mathbb{P}(X_{ij} = 1) = \frac{1}{2}$ for all $1 < i \neq j \leq n$, then $\lim_{n \to \infty} \mathbb{P}(S_1 > \max\{S_2, \ldots, S_n\}) \to 1$; that is, player 1 achieves the highest score with probability approaching 1.

1.4.4. *A chess round-robin tournament with ties* The following round-robin tournament model appeared in [16, 17]: for $i \neq j$, $X_{ij} + X_{ji} = 1$, $X_{ij} \in \{0, \frac{1}{2}, 1\}$; this can be seen as a special case of the general model where the F_{ij} have the support $\{0, \frac{1}{2}, 1\}$. The case where all players are equally strong was considered, i.e. $\mathbb{P}(X_{ij} = 1) = \mathbb{P}(X_{ji} = 1)$, and where the probability of a tie, $p = \mathbb{P}(X_{ij} = \frac{1}{2})$, is common to all games. A type of negative dependence called *negative relation* was proved which is weaker than NA [2, Chapter 2] for $S_1, \ldots, S_n$ using the log-concavity of the probability function of $2X_{ij}$, which requires restricting the range of p to $p = 0$ or $p \in [\frac{1}{3}, 1)$. Results from [2] were then used to prove a Poisson approximation to the number of times S_i exceeds a certain threshold. We strengthen this result to NA, which in fact holds for all p and, more generally, for all values of $\mathbb{P}(X_{ij} = 1)$ and $\mathbb{P}(X_{ij} = \frac{1}{2})$; i.e. the above assumptions of equality of strength and a common probability of ties are dropped.

1.5. Random-sum *n*-player games

The following somewhat abstract description of a tournament is a generalization of all the above tournament models. Consider a sequence of K n-player games (rounds), where the random payoff for player $i \in \{1, \ldots, n\}$ in round $k \in \{1, \ldots, K\}$ is $X_i^{(k)}$ and the components of each of the payoff vectors $\mathbf{X}^{(k)} = (X_1^{(k)}, \ldots, X_n^{(k)})$ are NA, with the $\mathbf{X}^{(k)}$ being independent. In general, the sum of the components of each $\mathbf{X}^{(k)}$ is assumed to be a random variable. Constant-sum (or, equivalently, zero-sum) examples are formed when the payoff vectors $\mathbf{X}^{(k)}$ have the multinomial or Dirichlet distribution (see [11, Section 3.1] for these and further examples). An example where the sum of the players' payoffs in each game is random is the case where the vector $\mathbf{X}^{(k)}$ is jointly normal with correlations ≤ 0 [11, Section 3.4].

The total payoff for player i in the K rounds is $S_i = \sum_{k=1}^{K} X_i^{(k)}$, $i = 1, \ldots, n$. We prove in Section 2, Theorem 2, that $S_1, \ldots, S_n$ are NA. More generally, we can take $S_i = u_i(X_i^{(1)}, \ldots, X_i^{(K)})$ where u_i is any increasing function of player i's payoffs. Note that here, unlike in pairwise duels, several (and even all) players may compete in each round. The limiting distribution of the number of pure Nash equilibria in such random games was studied in [26].

1.5.1. *Two sport examples* A football (soccer in the US) league provides an example of a random-sum round-robin tournament. The winning team is awarded three points, and if the game ends in a tie each team receives one point. For a single match the score possibilities for the two teams are (3, 0), (1, 1), and (0, 3) with some probabilities, forming an NA distribution for any probabilities. Let the n-dimensional vectors $\mathbf{X}^{(\mathbf{k})}$, for $\mathbf{k} = (ij)$ with $i \neq j$, consist of zeros except for two coordinates i and j corresponding to the playing teams i and j, where one of the above three vectors appears. Then, $\mathbf{S} = \sum_{\mathbf{k}=(i,j) : 1 \leq i < j \leq n} \mathbf{X}^{(\mathbf{k})}$ represents the vector of total scores of the n teams after they all play each other. It is easy to see that each vector $\mathbf{X}^{(\mathbf{k})}$ is NA. Equivalently, we can assume that $\mathbf{X}^{(\mathbf{k})}$ contains the scores of all players in all matches in week k.

Under some assumptions (which are an approximation to reality), the Association of Tennis Professionals (ATP) ranking is another example. It can be seen as a tournament in which the

number of points awarded to the winner of each game depends on the tournament and the stage reached. Players' ranks are increasing functions of their total scores. Here we do not assume that each player plays against all others in the ATP ranking, which is expressed by setting some of the rewards to zero.

1.6. Knockout tournaments

Consider a knockout tournament with $n = 2^\ell$ players of equal strength; that is, player i defeats player j independently of all other duels with probability $\frac{1}{2}$ for all $1 \leq i \neq j \leq n$. The winner continues to a duel with another winner, and the defeated player is eliminated from the tournament. Let S_i denote the number of games won by player i. We could also replace S_i by the prize money of player i, which in professional tournaments is usually an increasing function of S_i. For a completely random schedule of matches (aka the draw; see [1]), we show in Section 3 that the vector $\mathbf{S} = (S_1, \ldots, S_n)$ is NA. Note that in tennis tournaments such as Wimbledon the draw is not completely random as top-seeded players' matches are drawn in a way that prevents them from playing against other top-seeded players in early rounds. For nonrandom draws we prove the NOD property via a new preservation result, and we provide a counterexample to the NA property; thus, it need not hold for fixed, nonrandom draws. We also provide an example where NOD and NA do not hold if players are not of equal strength.

2. Negative association and round-robin tournaments

The following theorem generalizes [11, Application 3.2(c)]; it implies that the scores $S_1, \ldots, S_n$ in the general round-robin model of Section 1.4, and therefore in all round-robin models of Section 1.4, are NA, and therefore also NLOD, NUOD, and NOD.

Theorem 1. *Let $X_1, \ldots, X_n$ be independent random variables, and let g_i, $i = 1, 2, \ldots, n$ be decreasing functions. Set $Y_1 = g_1(X_1)$, ..., $Y_n = g_n(X_n)$, and for $j = 1, \ldots, m$ set $S_j = f_j(\{X_i : i \in A_j\}, \{Y_i : i \in B_j\})$, where f_j are coordinate-wise increasing functions of $|A_j| + |B_j|$ variables, and the sets $A_1, \ldots, A_m$ and $B_1, \ldots, B_m$ are disjoint subsets of $\{1, 2, \ldots, n\}$. Then the random variables $S_1, \ldots, S_m$ are NA.*

Proof. The pair of variables X, $g(X)$ with g decreasing is NA. This is well known; for completeness, here is a simple proof. Let X^* be an independent copy of X. For increasing functions f_1 and f_2, we have

$$2\mathrm{Cov}(f_1(X), f_2(g(X))) = \mathbb{E}\{[f_1(X) - f_1(X^*)][f_2(g(X)) - f_2(g(X^*))]\} \leq 0,$$

since the expression in the expectation is ≤ 0. The pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$ are independent, and each pair is NA. Property P_7 of [11] states that the union of independent sets of NA random variables is NA. Therefore, the random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ are NA. Property P_6 in [11] states that increasing functions defined on disjoint subsets of a set of NA random variables are NA. Therefore, $S_1, \ldots, S_m$ are NA. $\square$

We now apply Theorem 1 to show the NA property in the general round-robin model of Section 1.4.

Proposition 1. *Let $X_{ij} \sim F_{ij}$ with support on $[0, r_{ij}]$ be independent for $1 \leq i < j \leq n$, where $r_{ij} \geq 0$, and let $X_{ji} = r_{ij} - X_{ij}$. Set $S_i = \sum_{j=1, j \neq i}^n X_{ij}$, $i = 1, \ldots, n$. Then $S_1, \ldots, S_n$ are NA. More generally, if we set $S_i = u_i(X_{i1}, \ldots, X_{i,i-1}, X_{i,i+1}, \ldots, X_{in})$, $i = 1, \ldots, n$, where the u_i are any increasing functions, we again have that the variables $S_1, \ldots, S_n$ are NA.*

Proof. Instead of a single index we apply Theorem 1 to the independent doubly indexed random variables X_{ij} for $i < j$. Let $g_{ij}(x) = r_{ij} - x$, so that $X_{ji} = g_{ij}(X_{ij}) = r_{ij} - X_{ij}$, with X_{ji} playing the role of the Y in Theorem 1. Since the S_i are sums of disjoint subsets of the variables defined above, the result follows by Theorem 1, and the same argument holds with the functions u_i replacing the sums. $\qquad\square$

Since all the round-robin models of Section 1.4 are special cases of the general round-robin model, we have the following result.

Corollary 1. *The NA property for $S_1, \ldots, S_n$ holds in all the round-robin models in Section 1.4. The NLOD results proved in the literature for these models follow; moreover, NUOD and hence NOD also follow.*

The football example of Section 1.5 is not a special case of the constant-sum general round-robin model; here, the NA property follows by Theorem 1, replacing the functions g_{ij} by g defined by $g(3) = 0$, $g(1) = 1$, and $g(0) = 3$. It also follows by Theorem 2 below.

We now consider the random-sum n-player games tournament of Section 1.5.

Theorem 2. *Consider the random-sum tournament model of Section 1.5, i.e. a sequence of K n-player games (rounds), where the random payoff for player $i \in \{1, \ldots, n\}$ in round $k \in \{1, \ldots, K\}$ is $X_i^{(k)}$, and the components of each payoff vector $\mathbf{X}^{(k)} = \left(X_1^{(k)}, \ldots, X_n^{(k)}\right)$ are NA. The vectors $\mathbf{X}^{(k)}$ are distributed independently. Let $S_i = \sum_{k=1}^{K} X_i^{(k)}$. Then, $S_1, \ldots, S_n$ are NA. More generally, the variables $S_i = u_i\left(X_i^{(k)}, \ldots, X_i^{(K)}\right)$, $i = 1, \ldots, n$, where the u_i are any increasing functions, are NA.*

Theorem 2 can be restated in the following lemma, which follows readily from properties of negative association given in [11]. The same result for positive association, with the same proof, is given in [14, Remark 4.2].

Lemma 1. *The convolution of NA vectors is NA.*

Proof. Let $\mathbf{X}^{(k)} \in \mathbb{R}^n$ be independent NA vectors, and let $\mathbf{S} = (S_1, \ldots, S_n) = \sum_{k=1}^{K} \mathbf{X}^{(k)}$. By Properties P_7 and then P_6 of [11], the union of all variables in these vectors is NA; hence, $S_1, \ldots, S_n$ are NA since they are increasing functions of disjoint subsets of the above union. $\qquad\square$

This argument also holds when $S_i = u_i\left(X_i^{(1)}, \ldots, X_i^{(K)}\right)$, thus proving the last part of Theorem 2.

The next corollary shows that the NA property of the general round-robin model of Section 1.4, and hence in all the models of 1.4, also follows from Theorem 2.

Corollary 2. *The scores $S_1, \ldots, S_n$ of the general round-robin models in Section 1.4 are NA.*

Proof. For clarity, we start with the simple case of $n = 3$. Define the vectors $\mathbf{Y}^{12} = (X_{12}, r_{12} - X_{12}, 0)$, $\mathbf{Y}^{13} = (X_{13}, 0, r_{13} - X_{13})$, and $\mathbf{Y}^{23} = (0, X_{23}, r_{23} - X_{23})$ with X_{ij} of the general round-robin model. It is easy to see that $S_i = \sum_{1 \leq k < \ell \leq 3} Y_i^{k\ell}$.

In general, starting with the rewards X_{ij} of the general round-robin model, form the $K =: n(n-1)/2$ vectors $\mathbf{Y}^{ij} \in \mathbb{R}^n$, $1 \leq i < j \leq n$, with ith component, $Y_i^{ij} = X_{ij}$, jth component $Y_j^{ij} = r_{ij} - X_{ij}$, and the remaining components equal to zero. The components $\left(Y_1^{ij}, \ldots, Y_n^{ij}\right)$ of each of the K vectors $\mathbf{Y}^{ij}$ are obviously NA. Setting $S_i = \sum_{1 \leq k < \ell \leq n} Y_i^{k\ell}$, $i = 1, \ldots, n$, it is easy

to see that these S_i coincide with those of the general round-robin model. Theorem 2 applied to the K vectors $\mathbf{Y}^{kl}$ implies that the variables S_i are NA. □

3. Knockout tournaments

We now discuss negative dependence in the knockout tournament of Section 1.6.

3.1. Knockout tournaments with a random draw

Proposition 2. *Consider a knockout tournament starting with $n = 2^\ell$ players, where player i defeats player j independently of all other duels with probability $\frac{1}{2}$ for all $1 \leq i \neq j \leq n$; the winner continues to a duel with another winner, and the defeated player is eliminated from the tournament. Let S_i denote the number of games won by player i. Assume a completely random schedule (draw) of the matches. Then $S_1, \ldots, S_n$ are NA.*

Proof. First note that for a given ℓ, the vector $\mathbf{S} = (S_1, \ldots, S_n)$ contains the components $i = 0, \ldots, \ell$ with $i < \ell$ appearing $2^{\ell-1-i}$ times, and ℓ appearing once. For example, if $n = 4$ ($\ell = 2$) then there are two players with zero wins, one player (the losing finalist) with one win, and one player (the champion) with two wins. Thus, the vector $\mathbf{S}$ is a permutation of the vector $(0, 0, 1, 2)$. If $n = 8$ ($\ell = 3$) then $\mathbf{S}$ is a permutation of the vector $(0, 0, 0, 0, 1, 1, 2, 3)$. Under the assumption of a random draw, all permutations are equally likely as all players play a symmetric role. Theorem 2.11 of [11] states that if $\mathbf{X} = (X_1, \ldots, X_n)$ is a random permutation of a given list of real numbers, then $\mathbf{X}$ is NA, and the result follows. □

Without the assumption that players have equal probabilities in each duel, negative association as in Proposition 2 need not hold. To see this, consider the case of four players and assume first that the relations between the players are deterministic; specifically, player 1 beats player 2 with probability 1, and loses to players 3 and 4 with probability 1. Player 2 beats players 3 and 4 with probability 1, and player 3 beats player 4 with probability 1. These relations are not transitive (for example, player 1 beats player 2 who beats player 3, but player 3 beats player 1), which is not uncommon in various sports. With a random draw, the vector $S = (S_1, S_2, S_3, S_4)$ can only take the outcomes $(1, 0, 2, 0)$ (when player 1 meets player 2 in the first round), $(0, 2, 1, 0)$ (player 1 meets player 3 in the first round), and $(0, 2, 0, 1)$ (player 1 meets player 4 in the first round), each with probability $\frac{1}{3}$. Let $f_1(S_1) = S_1$ and $f_2(S_3) = S_3$. Then $\mathbb{E}f_1(S_1)f_2(S_3) = \frac{2}{3} > \mathbb{E}f_1(S_1)\mathbb{E}f_2(S_3) = \frac{1}{3}$, whereas $\mathbb{E}f_1(S_1) = \frac{1}{3}$ and $\mathbb{E}f_2(S_3) = 1$, which contradicts negative association. If we replace the probabilities of 1 by $1 - \varepsilon$ for small ε then the same result holds by an obvious continuity argument, so deterministic relations are not necessary for this example. In the above example the vector $\mathbf{S}$ is not even NLOD. In fact, $\mathbb{P}(S_1 \leq 0, S_3 \leq 0) = \frac{1}{3} > \frac{2}{3} \cdot \frac{1}{3} = \mathbb{P}(S_1 \leq 0)\mathbb{P}(S_3 \leq 0)$.

3.2. Knockout tournaments with a nonrandom draw

This section provides a counterexample showing that for knockout tournaments with a given nonrandom draw, the scores $S_1, \ldots, S_n$ need not be NA; however, we prove that they are NOD. To obtain the latter result we prove a result on NOD (and NLOD and NUOD) of independent interest.

Consider a knockout tournament with $n = 4$ players of equal strength and a draw where in the first round player 1 plays against player 2, and player 3 against player 4. In this case only eight permutations of $(0, 0, 1, 2)$ are possible and one of the first two coordinates must be positive, so $(0, 0, 1, 2)$ itself is not a possible outcome. Consider the functions $f_1(S_1, S_3)$

taking the value 0 everywhere apart from $f_1(0, 1) = f_1(0, 2) = 1$, and $f_2(S_2, S_4)$, which is 0 everywhere apart from $f_2(2, 0) = 1$. We have $\mathbb{E}[f_1(S_1, S_3)f_2(S_2, S_4)] = \frac{1}{8}$, $\mathbb{E}f_1(S_1, S_3) = \frac{2}{8}$, and $\mathbb{E}f_2(S_2, S_4) = \frac{1}{8}$, and (3) does not hold. In a tennis tournament, this arrangement of matches occurs if players 1 and 3 are top-seeded and the draw prevents them from being matched against each other in the first round.

Finally, we prove that in a knockout tournament with a nonrandom schedule, $\mathbf{S} = (S_1, \ldots, S_n)$ is NOD. We need the following theorem, which may be of independent interest.

Theorem 3. *Let $\mathbf{X}^{(k)} = \left(X_1^{(k)}, \ldots, X_n^{(k)}\right) \in \mathbb{R}^n$, $k = 1, \ldots, K$ satisfy the following two assumptions:*

(i) *for all $k = 1, \ldots, K$, $\mathbf{X}^{(k)} \mid \mathbf{X}^{(k-1)} + \cdots + \mathbf{X}^{(1)}$ is NLOD;*

(ii) *for all k and i, $X_i^{(k)} \mid \mathbf{X}^{(k-1)} + \cdots + \mathbf{X}^{(1)} \stackrel{d}{=} X_i^{(k)} \mid X_i^{(k-1)} + \cdots + X_i^{(1)}$;*

i.e. the conditional distribution of $X_i^{(k)}$ depends only on the ith coordinate of the sum of its predecessors. Then $\mathbf{X}^{(1)} + \cdots + \mathbf{X}^{(K)}$ is NLOD. Moreover, the result holds if we replace NLOD by NUOD, and hence also by NOD.

Proof. It is well known that a random vector $\mathbf{Z} = (Z_1, \ldots, Z_n)$ is NLOD if and only if $\mathbb{E} \prod_{i=1}^n \phi_i(Z_i) \leq \prod_{i=1}^n \mathbb{E}\phi_i(Z_i)$ for any nonnegative decreasing functions ϕ_i ([33, Theorem 6.G.1(b)] or [21, Theorem 3.3.16]). The proof proceeds by induction, and it is easy to see that it suffices to prove it for $K = 2$. Set $\mathbf{X} := \mathbf{X}^{(1)}$ and $\mathbf{Y} := \mathbf{X}^{(2)}$. We have

$$\mathbb{E} \prod_{i=1}^n \phi_i(X_i + Y_i) = \mathbb{E}\left\{ \mathbb{E}\left[\prod_{i=1}^n \phi_i(X_i + Y_i) \mid \mathbf{X} \right] \right\} \leq \mathbb{E} \prod_{i=1}^n \mathbb{E}[\phi_i(X_i + Y_i) \mid \mathbf{X}] = \mathbb{E}\left[\prod_{i=1}^n g_i(X_i) \right],$$

where $g_i(X_i) = E[\phi_i(X_i + Y_i) \mid \mathbf{X}]$, and the inequality holds by assumption (i). By assumption (ii) we have that $g_i(X_i)$ indeed depends only on X_i, and it is obviously nonnegative and decreasing. By the NLOD property of $\mathbf{X}$ we have

$$\mathbb{E} \prod_{i=1}^n g_i(X_i) \leq \prod_{i=1}^n \mathbb{E}g_i(X_i) = \prod_{i=1}^n \mathbb{E}\phi_i(X_i + Y_i),$$

and the result follows. The same proof holds for NUOD with the functions ϕ_i taken to be increasing. □

A special case of Theorem 3 is the following corollary, which, for nonnegative vectors, follows from [33, Theorem 6.G.19] and can be obtained from [15, Theorem 1] (for vectors in $\mathbb{R}^2$) and from [20, Theorem 4.2(e)].

Corollary 3. *The sum of independent NOD (NLOD, NUOD) vectors is NOD (NLOD, NUOD).*

Proposition 3. *For the knockout tournament with a nonrandom draw, the vector $\mathbf{S} = (S_1, \ldots, S_n)$ is NOD.*

Proof. Without loss of generality, assume that in the first round player $2i - 1$ plays against player $2i$ for $i = 1, \ldots, n/2$. Let $X_j^{(1)} = 0$ (1) if player j loses (wins) the first round, $j = 1, \ldots, n$. The pairs of variables $X_{2i-1}^{(1)}, X_{2i}^{(1)}$ are independent and NOD (in fact they are NA), taking the values (0, 1) or (1, 0). It follows readily that the 0–1 vector $\mathbf{X}^{(1)} = \left(X_1^{(1)}, \ldots, X_n^{(1)}\right)$,

whose jth coordinate indicates a win or a loss for player j in the first round, is NOD. Now the second round is similar with only half the players (those who won the first round), where the value 0 is set for players who lost in the first round. Continuing this way, we see that the vector $(S_1, \ldots, S_n)$ is the sum of the 0–1 vectors of all the rounds. These vectors are not independent because the value of 0 in a coordinate of a vector pertaining to a given round must by followed by a zero there in the next round. However, assumptions (i) and (ii) of Theorem 3 are easily seen to hold, and the NOD property follows. □

Acknowledgement

Funding information

Competing interests

There were no competing interests to declare that arose during the preparation or publication process of this article.

References

[1] ADLER, I., CAO, Y., KARP, R., PEKÖZ, E. A. AND ROSS, S. M. (2017). Random knockout tournaments. *Operat. Res.* **65**, 1589–1596.

[2] BARBOUR, A. D., HOLST, L. AND JANSON, S. (1992). *Poisson Approximation* (Oxford Studies in Probability). The Clarendon Press, Oxford.

[3] BENJAMINI, Y. AND YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29**, 1165–1188.

[4] BRADLEY, R. A. AND TERRY, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* **39**, 324–345.

[5] BRUSS, F. T. AND FERGUSON, T. S. (2018). Testing equality of players in a round-robin tournament. *Math. Sci.* **43**, 125–136.

[6] BULINSKI, A. AND SHASHKIN, A. (2007). *Limit Theorems for Associated Random Fields and Related Systems* (Adv. Ser. Statist. Sci. Appl. Prob. **10**). World Scientific, Hackensack, NJ.

[7] DALY, F. (2016). Negative dependence and stochastic orderings. *ESAIM Prob. Statist.* **20**, 45–65.

[8] EFRON, B. (1965). Increasing properties of Pólya frequency functions. *Ann. Math. Statist.* **36**, 272–279.

[9] GOLDSTEIN, L. AND WIROONSRI, N. (2018). Stein's method for positively associated random variables with applications to the Ising and voter models, bond percolation, and contact process. *Ann. Inst. H. Poincaré Prob. Statist.* **54**, 385–421.

[10] HUBER, P. J. (1963). A remark on a paper of Trawinski and David entitled: Selection of the best treatment in a paired comparison experiment. *Ann. Math. Statist.* **34**, 92–94.

[11] JOAG-DEV, K. AND PROSCHAN, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11**, 286–295.

[12] JOE, H. (1997). *Multivariate Models and Dependence Concepts* (Monographs Statist. Appl. Prob. **73**). Chapman & Hall, London.

[13] JOE, H. (2015). *Dependence Modeling with Copulas* (Monographs Statist. Appl. Prob. **134**). CRC Press, Boca Raton, FL.

[14] KARLIN, S. AND RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivar. Anal.* **10**, 467–498.

[15] LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.*, **37**, 1137–1153.
[16] MALINOVSKY, Y. (2022). On the distribution of winners' scores in a round-robin tournament. *Prob. Eng. Inf. Sci.* **36**, 1098–1102.
[17] MALINOVSKY, Y. (2022). Correction to "On the distribution of winners' scores in a round-robin tournament." To appear in *Prob. Eng. Inf. Sci.* DOI: https://doi.org/10.1017/S0269964822000158.
[18] MALINOVSKY, Y. AND MOON, J. W. (2022). On the negative dependence inequalities and maximal score in round-robin tournaments. *Statist. Prob. Lett.* **185**, 109432.
[19] MOON, J. W. (2013). *Topics on Tournaments*. Available at https://www.gutenberg.org/ebooks/42833.
[20] Müller, A. (1997). Stochastic orders generated by integrals: A unified study. *Adv. Appl. Prob.* **29**, 414–428.
[21] MÜLLER, A. AND STOYAN, D. (2002). *Comparison Methods for Stochastic Models and Risks*. John Wiley, Chichester.
[22] NEWMAN, C. M. (1984). Asymptotic independence and limit theorems for positively and negatively dependent random variables. In *Proc. Symp. Inequalities in Statist. Prob.* (IMS Statist. Lect. Notes Monograph Ser. **5**), ed. Y. L. TONG. Institute of Mathematical Statistics, pp. 127–140.
[23] PEMANTLE, R. (2000). Towards a theory of negative dependence. *J. Math. Phys.* **41**, 1371–1390.
[24] PUCEETI, G. AND WANG, R. (2015). Extremal dependence concepts. *Statist. Sci.* **30**, 485–517.
[25] RINOTT, Y. AND POLLAK, M. (1980). A stochastic ordering induced by a concept of positive dependence and monotonicity of asymptotic test sizes. *Ann. Statist.* **190**, 8–198.
[26] RINOTT, Y. AND SCARSINI, M. (2000). On the number of pure strategy Nash equilibria in random games. *Games Econom. Behavior* **33**, 274–293.
[27] ROSS, S. M. (2016). Improved Chen–Stein bounds on the probability of a union. *J. Appl. Prob.* **53**, 1265–1270.
[28] ROSS, S. M. (2022). Team's seasonal win probabilities. *Prob. Eng. Inf. Sci.* **36**, 988–998
[29] ROUSSAS, G. G. (1999). Positive and negative dependence with some statistical applications. In *Asymptotics, Nonparametrics, and Time Series* (Statist. Textbooks Monogr. **158**), ed. S. Ghosh. Dekker, New York, pp. 757–788.
[30] RÜSCHENDORF, L. (2013). *Mathematical Risk Analysis: Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, Heidelberg.
[31] SAMUEL-CAHN, E. (1996). Is the Simes improved Bonferroni procedure conservative? *Biometrika* **83**, 928–933.
[32] SARKAR, S. K. AND CHANG, C.-K. (1997). The Simes method for multiple hypothesis testing with positively dependent test statistics. *J. Amer. Statist. Assoc.* **92**, 1601–1608.
[33] SHAKED, M. AND SHANTHIKUMAR, J. G. (2007). *Stochastic Orders*. Springer, New York.
[34] ZERMELO, E. (1929). Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. *Math. Z.* **29**, 436–460.