On Round-Robin Tournaments with a Unique Maximum Score and Some Related Results

Yaakov Malinovsky *
Department of Mathematics and Statistics
University of Maryland, Baltimore County, Baltimore, MD 21250, USA

John W. Moon †
Department of Mathematical and Statistical Sciences
University of Alberta, Edmonton, AB T6G 2G1, Canada

September 1, 2022

Abstract

Richard Arnold Epstein (1927-2016) published the first edition of "The Theory of Gambling and Statistical Logic" in 1967. He introduced some material on round-robin tournaments (complete oriented graphs) with n labeled vertices in Chapter 9; in particular, he stated, without proof, that the probability that there is a unique vertex with the maximum score tends to one as n tends to infinity. Our object here is to give a proof of this result along with some historical remarks and comments. We also give related results on pairs of equal scores and degrees in tournaments and graphs.

Keywords: Complete graph, large deviations, maximal score, probabilistic inequalities, random graph, round-robin tournaments

MSC2020: 05C20,05C07,60F10, 60E15

*email: yaakovm@umbc.edu
†email: jwmoon@ualberta.ca

1 Introduction

In a classical round-robin tournament, each of n players wins or loses a game against each of the other $n-1$ players (Moon, 1968). Let X_{ij} equal 1 or 0 according as player i wins or loses the game played against player j, for $1 \leq i, j \leq n$, $i \neq j$, where $X_{ij} + X_{ji} = 1$. We assume that all $\binom{n}{2}$ pairs (X_{ij}, X_{ji}) are independently distributed with $P(X_{ij} = 1) = P(X_{ji} = 0) = 1/2$. Let

$$s_i = \sum_{j=1, j\neq i}^{n} X_{ij}$$

denote the score of player $i, 1 \leq i \leq n$, after playing against all the other $n-1$ players. We refer to $(s_1, s_2, \ldots, s_n)$ as the score sequence of the tournament.

Round-robin tournaments can be considered as a model of paired comparison experiments used in an attempt to rank a number of objects with respect to some criterion–or at least to determine if there is any significant difference between the objects–when it is impracticable to compare all the objects simultaneously. Measuring players' strengths in tournaments by paired comparisons of strength has a long history and there is a considerable literature on such experiments; see, e.g., Zermelo (1929), David (1988), David and Edwards (2001), and Aldous and Kolesnik (2022). In particular, David (1959) generated the score sequences of tournaments with n players for $3 \leq n \leq 8$ and their frequencies by expanding products of the form

$$G(n) = \prod_{1\leq i<j\leq n} (a_i + a_j).$$

For example,

$$G(3) = a_1^2 a_2 + a_1 a_2^2 + a_1^2 a_3 + a_1 a_3^2 + a_2^2 a_3 + a_2 a_3^2 + 2a_1 a_2 a_3,$$

so the (ordered) score sequence $(1, 1, 1)$ occurs twice while $(0, 1, 2)$ occurs six time in the $2^3 = 8$ possible outcomes when $n = 3$. He used this information to develop, among other things, tests for deciding whether the maximum in a given outcome was significantly larger than the expected value $(n-1)/2$ of a given score.

Let r_n denote the probability that an ordinary tournament with n labeled vertices has a unique vertex with maximum score, assuming all the $2^{\binom{n}{2}}$ such tournaments are equally likely. Epstein (1967), p. 353 gave the values $r_4 = .5, r_5 = .586, r_6 = .627, r_7 = .581$, and

$r_8 = .634$ with no explanation of how these numbers were calculated. However, the paper of David (1959) is included among the references Epstein gave at the end of the chapter containing these values. So it seems plausible to assume that he deduced these values from Table 1 in David (1959) since the values agree except that it follows from Table 1 that the value for r_8 should be $160,241,152/2^{28} = .596$.

Stockmeyer (2022) has recently pointed out that MacMahon (1923) generated the score sequences and their frequencies for tournaments with up to 9 vertices and his results agree with David's for $n = 8$. It follows from MacMahon's data that $r_9 = 42,129,744,768/2^{36} = .613$.

As a partial check we obtained the same result by another approach that made use of information about 8-vertex score frequencies but did not require prior information about 9-vertex degree distributions. Then one of us began trying to apply this approach to determine r_{10}, whose value was not known to us at the time. In the course of this process an error was discovered in the frequency that MacMahon (1923) (p.26) gave for the sequence $(2,2,3,3,4,4,6,6,6)$ and its complement; the frequency he gave, $361,297,520$, is not divisible by 9, as it should be since there are 9 choices for the label of the vertex that represents the winner of the match between, say, the two vertices of degree 2. It is a fairly straightforward exercise to show that the correct frequency for each of these two sequences is $10,000$ more than the frequency stated. With these corrections, the sum of all the frequencies is what it should be and this approach led to the conclusion that $r_{10} = 21,293,228,876,800/2^{45} = .605$.

During this process the other author discovered that Doron Zeilberger (Zeilberger, 2016) had extended MacMahon's work and had generated the score vectors and their frequencies for tournaments with up to 15 vertices using the Maple program. (We remark that Zeilberger's frequency for the two sequences mentioned earlier agree with the corrected values we gave.) Using a Matlab program, the values of r_n were deduced from Zeilberger's data for $n = 10$, 11, and 12. The value obtained when $n = 10$ agreed with the value stated above. And, as a partial check, we confirmed that the value for r_{12} obtained by using 12-vertex frequency data agreed with the value obtained by applying the other procedure described earlier to 11-vertex data. The values for $n = 4, 5, \ldots, 12$ are given in Table 1.

It was around this time that we discovered that the sequence `https://oeis.org/A013976` contains the values of the number of tournaments with a unique vertex of maximal score for $n = 1, 2, \ldots, 16$; these values are attributed to Michael Stob and Andrew Howroyd.

We remark that Hamming (1968) reviewed the 1st edition of Epstein's book. Three later editions of Epstein (1967) are reviewed in Mathematical Reviews; the review by Edward Thorp (1977, MR0446535) is the most thorough. The later editions contain more material and references, but the material about r_n remains unchanged (Epstein, 2013).

Epstein also stated that as n increases indefinitely, r_n approaches unity without a proof or reference for such a result.

It is not feasible to test Epstein's claim for large values of n by generating score sequences and their frequencies directly because of the length of time this would take. For example, executing the Zeilberger (2016) Maple code for $n = 17$ on a powerful computer (a dual CPU Intel 2620 v4, 1TB of memory, Unix operating systems) took 157:04 hours. We note that in this case, the number of different score sequences is 6,157,058 (see `https://oeis.org/A000571` and references therein.) So we have used Monte-Carlo simulations (Metropolis and Ulam, 1949) to test Epstein's statement for larger values of n. For a given value of n we sample $n(n-1)/2$ values of random Bernoulli variables $X_{ij}, 1 \leq i < j \leq n$, where $P(X_{ij} = 0) = P(X_{ij} = 1) = 1/2$; this determines a random n-vertex tournament and its score sequence. We repeated this process M times for a predetermined integer M. We let I_t denote a random indicator function that equals one if the tournament obtained at the t-th repetition has a unique score of maximum value, and equals zero otherwise, for $1 \leq t \leq M$. Then $\hat{r}_n(M) = 1/M \sum_{t=1}^{M} I_t$ is an unbiased estimator of r_n (i.e., $E\left(\hat{r}_n(M)\right) = r_n$ for any M); also $\hat{r}_n$ is a consistent estimator of r_n for large M, (i.e., $\lim_{M\to\infty} P\left(|\hat{r}_n(M) - r_n| > \varepsilon\right) = 0$ for any $\varepsilon > 0$). We used smaller values of M for some of the larger values of n because of time constraints. The results of these simulations are given in the Table 1 below.

n	M	r_n	$\hat{r}_n(M)$
4	10^6	0.5	$0.5003\cdots$
5	10^6	$600/2^{10} = 0.5859\cdots$	$0.5862\cdots$
6	10^6	$20,544/2^{15} = 0.6269\cdots$	$0.6262\cdots$
7	10^6	$1,218,224/2^{21} = 0.5808\cdots$	$0.5806\cdots$
8	10^6	$160,241,152/2^{28} = 0.5969\cdots$	$0.5966\cdots$
9	10^6	$42,129,744,768/2^{36} = 0.6130\cdots$	$0.6129\cdots$
10	10^6	$21,293,228,876,800/2^{45} = 0.6051\cdots$	$0.6054\cdots$
11	10^6	$22,220,602,090,444,032/2^{55} = 0.6167\cdots$	$0.6169\cdots$
12	10^6	$45,959,959,305,969,143,808/2^{66} = 0.6228\cdots$	$0.6231\cdots$
13	10^6		$0.6240\cdots$
14	10^6		$0.6323\cdots$
15	10^6		$0.6355\cdots$
30	10^6		$0.6881\cdots$
50	10^6		$0.7290\cdots$
100	10^6		$0.7808\cdots$
500	10^4		$0.8673\cdots$
1,000	10^4		$0.8996\cdots$
10,000	300		$0.9533\cdots$

Table 1: $r_4, \ldots, r_8$ were calculated from the scores distribution given in Table 1 of David (1959); r_9 from MacMahon (1923) data; r_{10}, r_{11}, r_{12} from Zeilberger (2016) data; see also `https://oeis.org/A013976`.

One of us inherited a copy of the 1st edition of Epstein's book a number of years ago and in the early 1980's wrote a letter to him, at the address on the title page of his book, to enquire about his statement concerning the uniqueness of the maximum scores in tournaments; the letter was returned stamped "return to sender, unable to locate". So a second letter was sent, this time to Academic Press, the publisher of the book; the reply to this letter stated that a certain address in California was the last known address. But there was no response to a letter sent to that address; so the matter was dropped for around 40 years until the present authors began considering an inequality by Huber (1963) for the distribution function of the scores in tournaments that had implications for the maximum scores in tournaments. Professor Noga Alon (Alon, 2022) referred us to a paper by Erdős and Wilson (1977) that contained a Lemma stating, in effect, that almost all labeled graphs in which pairs of vertices are joined by an edge with probability 1/2 have a unique vertex of maximum degree. In pursuing his suggestion we learned that Bollobás (1981) had derived numerous results on the distributions of the degree sequences $d_1 \geq d_2 \geq \cdots \geq d_n$ of ordinary n-vertex graphs in which edges are present with probability p; see also, Bollobás (2001) and Frieze and Karoński (2016). And then, after we had started writing this note, we came upon Guy (1984) by accident and learned that Epstein's statement had, in effect, been proposed as a Monthly problem and evidently was still unresolved in 1984.

2 Main Result

For expository convenience we introduce some notation and relations that we shall need later. Let

$$b(n-1,j) = P(s_i = j) = C(n-1,j)1/2^{n-1}, \;\; C(n-1,j) = \binom{n-1}{j}$$

and

$$B(n-1,k) = P(s_i > k) = \sum_{j>k} b(n-1,j)$$

for $0 \leq k, j \leq n-1$ and $1 \leq i \leq n$.

Next, let

$$t_{n-1} = (n-1)/2 + x_{n-1}((n-1)/4)^{1/2} \tag{1}$$

where

$$x_{n-1} = (2\log(n-1) - (1+\epsilon)\log(\log(n-1)))^{1/2} \tag{2}$$

for an arbitrary constant ϵ between 0 and 1, say. Then $x_{n-1} \to \infty$ and $x_{n-1} = o(n^{1/6})$ as $n \to \infty$ it follows that

$$b(n-1,t_{n-1}) \sim \left(\frac{2}{\pi(n-1)}\right)^{1/2} e^{-\frac{x_{n-1}^2}{2}} \sim \frac{\sqrt{2}(\log(n-1))^{(1+\epsilon)/2}}{\sqrt{\pi(n-1)^3}} \tag{3}$$

and

$$B(n-1,t_{n-1}) \sim \frac{1}{\sqrt{2\pi}}\frac{1}{x_{n-1}}e^{-\frac{x_{n-1}^2}{2}} \sim \frac{(\log(n-1))^{\epsilon/2}}{\sqrt{4\pi}(n-1)}; \tag{4}$$

see, e.g., Feller (1968), Ch. VII.2, Ch. VII.6 and Huber (1963).

Theorem 1. *The probability that a random n-vertex tournament T_n has a unique vertex of maximum score tends to one as n tends to infinity. In particular, if t_{n-1} is defined as in* (1) *and* (2) *and $s^\star$ denotes the maximum value of the scores $s_1, \ldots, s_n$ in T_n, then the following statements hold:*

(i) $P(s^\star > t_{n-1}) \to 1$ *as* $n \to \infty$;

(ii) If $W_n = W_n(T_n)$ denotes the number of ordered pairs of distinct vertices u and v in T_n such that $s_u = s_v = h$ for some integer h such that $t_{n-1} \leq h \leq n-1$, then $P(W_n > 0) \to 0$ *as* $n \to \infty$.

Proof. Conclusion (i) was proved by Huber (1963) and makes use of the following inequality for the joint distribution function for the scores $s_1, \ldots, s_n$ in a round-robin tournament.

$$P\left(s_1 < k_1, \ldots, s_m < k_m\right) \leq P\left(s_1 < k_1\right) \cdots P\left(s_m < k_m\right), \tag{5}$$

where $m \leq n$. Huber (1963) proved the inequality for any probability matrix (p_{ij}), where $p_{ij} + p_{ji} = 1$, and any numbers $(k_1, \ldots, k_m), m \leq n$. In our present case we assume $p_{ij} = 1/2$ for all $i \neq j$. Huber's derivation of (5) involves coupling arguments, a straightforward iterative procedure in which the dependent variables X_{ij} and X_{ji}, defined earlier, are replaced, one pair at a time, by independent variables Z_{ij} and Z_{ji} without changing the marginal distributions. The procedure leads to the conclusion that the left hand side of (5) is bounded above by a product of the corresponding probabilities for independent variables with the same distribution as the variables in the left hand side. Huber's inequality holds if we replace $<$ by $\leq$ in (5), and in a more general round-robin tournament model (Malinovsky and Moon, 2022); also in different tournaments and games models Malinovsky and Rinott (2022).

Once we have this inequality, the required conclusion follows upon observing that

$$P\left(s^\star < t_{n-1}\right) \leq \left(1 - B(n-1, t_{n-1})\right)^n \leq e^{-nB(n-1,t_{n-1})} \leq (1+o(1))e^{-\frac{(\log(n-1))^{\epsilon/2}}{\sqrt{4\pi}}} \to 0, \tag{6}$$

as $n \to \infty$, appealing to the definition of $B\left(n-1, t_{n-1}\right)$, inequality (5), the inequality $1 - c \leq e^{-c}$, and relation (4).

We now give a second proof of (i). This approach is based on an application of the Chebyshev's or Cauchy-Schwarz inequalities and is often called the 2nd Moment Method when it is applied in probabilistic problems in Graph Theory. See also Bollobás (1981), Sect.3 and Frieze and Karoński (2016), Sect. 3.2 where this method is applied to more general problems on the degrees in ordinary graphs.

Let $Y_t = Y_t(T_n)$ denote the number of vertices in T_n with score larger than $t = t_{n-1}$, i.e. $Y_t = \sum_{j=1}^{n} I\left(s_j > t\right)$. Since $P\left(s_u > t\right) = B\left(n-1, t\right)$ for any u and t, it follows that

$$E\left(Y_t\right) = nB(n-1, t) \sim n\frac{(\log(n-1))^{\epsilon/2}}{\sqrt{4\pi}(n-1)}, \tag{7}$$

upon appealing to (4) with $t = t_{n-1}$.

To determine the variance $Var(Y_t)$ we first observe that for any ordered pair of vertices u and v we can write their scores as $s_u = s'_u + X_{uv}$ and $s'_v + X_{vu}$, where s'_u and s'_v are the number of games u and v win against the remaining $n-2$ players; since s'_u and s'_v are independent variables, it follows that

$$\begin{aligned}
&P\left(s_u > t, s_v > t\right) \\
&= P\left(X_{uv} = 0\right) P\left(s'_u > t\right) P\left(s'_v > t-1\right) + P\left(X_{uv} = 1\right) P\left(s'_u > t-1\right) P\left(s'_v > t\right) \\
&= B(n-2, t)B(n-2, t-1).
\end{aligned}$$

Consequently,

$$\begin{aligned}
&Var(Y_t) = E(Y_t) + E(Y_t(Y_t-1)) - (E(Y_t))^2 \\
&= E(Y_t) + n(n-1)B(n-2, t)B(n-2, t-1) - (nB(n-1, t))^2.
\end{aligned} \tag{8}$$

Now

$$B(n-1, t) = 1/2(B(n-2, t) + B(n-2, t-1)),$$

so relation (8) simplifies to

$$\begin{aligned}
&Var(Y_t) = E(Y_t) - nB(n-2, t)B(n-2, t-1) - (1/4)n^2\left\{B(n-2, t-1) - B(n-2, t)\right\}^2 \\
&= E(Y_t) - nB(n-2, t)B(n-2, t-1) - (1/4)n^2 b(n-2, t)^2 \leq E(Y_t).
\end{aligned} \tag{9}$$

Therefore,

$$P(Y_t = 0) \leq P(|Y_t - E(Y_t)| \geq E(Y_t)) \leq \frac{Var(Y_t)}{(E(Y_t))^2} \leq \frac{1}{E(Y_t)} \to 0,$$

as $n \to \infty$, by Chebyshev's Inequality, relation (9), and (7). This implies conclusion (i).

We now turn to conclusion (ii). In view of conclusion (i), we may restrict our attention to tournaments T_n in which the maximum value $s^\star$ of the scores realized by the vertices is at least as large as $t = t_{n-1}$. Recall that $W_n = W_n(T_n)$ denotes the number of ordered pairs of distinct vertices u and v of T_n such that $t < s_u = s_v$ where $t \leq n-1$, i.e.

$$W_n = \sum_{1 \leq v < u \leq n} I(t < s_u = s_v).$$

Let s'_u and s'_v denote the scores of two such vertices u and v in their matches with the remaining $n-2$ players and note that s'_u and s'_v are independent variables. Then it follows that

$$\begin{aligned} &P(s_u = h, s_v = h) = 1/2P(s'_u = h-1)P(s'_v = h) + 1/2P(s'_u = h)P(s'_v = h-1) \\ &= C(n-2, h-1)(1/2)^{n-2}C(n-2, h)(1/2)^{n-2} \\ &\leq 4\frac{h}{n-1}\left(1 - \frac{h}{n-1}\right) C(n-1, h)(1/2)^{n-1}C(n-1, h)(1/2)^{n-1} \leq (b(n-1, h))^2 . \qquad (10) \end{aligned}$$

Hence,

$$\begin{aligned} &E(W_n) = E\left(\sum_{1 \leq v < u \leq n} I(t < s_u = s_v)\right) = n(n-1)E(I(t < s_1 = s_2)) \\ &= n(n-1)P(t < s_1 = s_2) = n(n-1)\sum_{h=t+1}^{n-1} P(s_u = h, s_v = h) \leq n(n-1)\sum_{h=t+1}^{n-1} b(n-1, h)^2 \\ &\leq n(n-1)b(n-1, t+1)B(n-1, t) \leq n(n-1)b(n-1, t)B(n-1, t) \\ &\sim \frac{(\log(n-1))^{1/2+\epsilon}}{\pi\sqrt{2(n-1)}} \to 0, \end{aligned}$$

as $n \to \infty$. Consequently, appealing to (3), (4), and to the fact that $W_n = W_n I(W_n > 0) \geq I(W_n > 0)$, we find that

$$1 - P(W_n = 0) = P(W_n > 0) \leq E(W_n) \to 0,$$

as required.

$\square$

For the sake of completeness, we mention an upper bound Huber (1963) gave for the maximum score $s^\star$ in almost all tournaments T_n. Let $t' = t'_{n-1}$ be defined as $t = t_{n-1}$ was defined earlier except that the ϵ in relation (2) is replaced by $-\epsilon$, it turns out that a relation corresponding to (4) is

$$B(n-1, t'_{n-1}) \sim \frac{(\log(n-1))^{-\epsilon/2}}{\sqrt{4\pi}(n-1)}.$$

Hence, it follows from Boole's inequality that

$$P\left(s^\star > t'\right) \le nB(n-1, t') = O\left((\log(n-1))^{-\epsilon/2}\right), \tag{11}$$

as $n \to \infty$. From (6) and (11) Huber (1963) concluded that

$$s^\star - \frac{n-1}{2} - \sqrt{\frac{n-1}{4}}\sqrt{2\log(n-1)} \to 0$$

in probability as $n \to \infty$.

3 Pairs of Equal Scores and Degrees in Tournaments and Graphs

We saw in the last section that pairs of vertices with equal scores at least as large as t_{n-1} are rare for large values of n. The suggests the problem of determining the expected value of the number $F_n = F_n(T_n)$ of unordered pairs of vertices u and v in a tournament T_n such that $s_u = s_v$. It follows from the first part of relation (10) that if $n \ge 3$, then

$$\begin{aligned} E(F_n) &= C(n,2)(1/2)^{2(n-2)} \sum_h C(n-2, h-1)C(n-2, h) \\ &= C(n,2)(1/2)^{2(n-2)} \sum_h C(n-2, h-1)C(n-2, n-2-h) \\ &= C(n,2)C(2(n-2), n-3)(1/2)^{2(n-2)} \sim \frac{1}{\sqrt{\pi(n-2)}}C(n,2), \end{aligned}$$

appealing to Vandermonde's identity at the next to the last step, and to equation (2.6) in Feller (1968), p.80.

Similarly, one could consider the expected value of the number $L_n = L_n(G_n)$ of unordered pairs of vertices u and v in an ordinary graph G_n with equal degrees d_u and d_v. Let d'_u and d'_v denote the number of edges joining u and v to the remaining $n-2$ vertices in G_n, and note that d'_u and d'_v are independent variables. When we consider the cases when there is or is not an edge joining u and v, we find that

$$P\left(d_u = h, d_v = h\right) = 1/2(P(d'_u = h-1)P(d'_v = h-1) + P(d'_u = h)P(d'_v = h))$$
$$= (1/2)^{2n-3}((C(n-2, h-1))^2 + (C(n-2, h))^2);$$

see, e.g. Bollobás (2001), p. 69. Hence,

$$E\left(L_n\right) = C(n,2)(1/2)^{2n-3}\sum_h \left[C(n-2,h-1)C(n-2,n-h-1) + C(n-2,h)C(n-2,n-2-h)\right]$$
$$= C(n,2)C(2(n-2), n-2)(1/2)^{2(n-2)}. \tag{12}$$

Note that

$$E(F_n) = \frac{n-2}{n-1}E(L_n).$$

Suppose no two vertices of a tournament T_n or an ordinary graph G_n have the same score or degree, assuming that $n \geq 2$; then the corresponding (non-decreasing) score or degree sequences must be $(0, 1, \ldots, n-1)$. The following observations are well known. The transitive tournament T_n, in which the vertices can be linearly ordered in such a way that each vertex wins against each of its predecessors in the sequence, is the only tournament with this score sequence (Moon (1968), chapter 7). But there is no graph G_n with this degree sequence since it is not possible for there to be vertices of degree 0 and $n-1$ in the same graph.

Remark 1. *Every non-trivial graph G_n contains at least one pair of vertices with equal degrees.*

As it turns out, the situation is reversed, in a sense, when we consider tournaments and graphs with exactly one pair of vertices with the same score or degree. From now on we shall frequently refer to such pairs of vertices as pairs of *special* vertices.

Theorem 2. *There is no round-robin tournament T_n with exactly one pair of special vertices.*

Proof. This certainly holds for small values of n, so we may assume that $n \geq 4$, say. If there is such a tournament T_n then its scores must be members of the set $N_n = (0, 1, \ldots, n-1)$. Let h denote the score of the two (and only two) vertices with the same score realized by the two special vertices. The remaining $n-2$ scores must be distinct and not equal to h. Hence, these scores can collectively realize only $n-2$ of the $n-1$ possible values remaining when h is removed from the set N_n; that is, there is an integer j, where $j \in N_n$ and $j \neq h$, that is not realized as the score of any vertex in T_n. This implies that the sum of the scores of all the vertices of T_n–bearing in mind that there are two vertices of score h and no vertices of score j–must equal

$$0 + 1 + 2 + \cdots + (n-1) + h - j = C(n, 2) + h - j = C(n, 2),$$

the total number of matches in the tournament. This would imply that $h = j$, which contradicts our assumptions. The result now follows. □

Let A and B denote two (disjoint) ordinary graphs (without loops or multiple edges). In what follows we let $A \cup B$ denote the graph consisting of a copy of A and a copy of B regarded as a single graph. And we let $A + B$ denote the graph obtained from $A \cup B$ by introducing additional edges that collectively join each vertex in A to each vertex in B. For example, if K_1 denotes the trivial graph consisting of a single isolated vertex, then $K_1 \cup K_1$ denotes the graph consisting of two vertices not joined by an edge and $K_1 + K_1$ denotes the graph consisting of two vertices that are joined by an edge.

If $n \geq 2$, let J_n denote the class of $n-$vertex graphs G_n with a unique pair of special vertices u and v with equal degrees.

Theorem 3. *If $n \geq 2$, then $G_n \in J_n$ if and only if G_n can be expressed in one of the following ways:*

$$G_n = H'_n \quad or \quad G_n = H''_n$$

where

$$H'_2 = K_1 \cup K_1 \quad and \quad H''_2 = K_1 + K_1$$

and

$$H'_n = H''_{n-1} \cup K_1 \quad and \quad H''_n = H'_{n-1} + K_1 \tag{13}$$

for $n \geq 3$.

Proof. The conclusion certainly holds when $n = 2$. The general argument will be by induction on n and will involve two different graphs, that are the complements or duals of each other, for each value of n. We will present the argument in three Assertions.

Assertion 1. *Suppose $n \geq 3$ and $G_n \in J_n$. If $d_u = d_v = h \in N_n$, for the unique pair of special vertices u and v of G_n, then $h \neq 0, n-1$.*

Proof. This is certainly true when $n = 3$, so we may assume that $n \geq 4$. The graph G_{n-2} determined by $n-2$ vertices of G_n other than u and v must contain at least one pair of special vertices w and z, since, by Remark 1, every non-trivial graph contains at least one such pair. Suppose that h does equal 0 or $n-1$. Then each vertex of G_{n-2} was originally joined to neither or to both of the vertices u and v in G_n, according as $h = 0$ or $h = n-1$. Consequently, the vertices w and z are also special vertices of G_n. But this would imply that G_n had at least two pairs of special vertices, namely, u and v and w and z, contradicting our assumption. □

When we refer to members of the degree sequence $(d_1, d_2, \ldots, d_n)$ of the graph G_n we assume the vertices are labeled so that the degrees appear in non-decreasing order.

Assertion 2. *If $G_n \in J_n$ and $n \geq 3$, then either*

(a) $d_1 = 0$ *or*

(b) $d_n = n - 1$

but not both.

Proof. The original degrees of the $n-2$ vertices of G_{n-2} determined by the $n-2$ vertices of G_n other than u and v are distinct members of the $n-1$ set $N' = N_n - \{h\}$. We note that 0 and $n-1$ are still in the set N', by Assertion 1. But there cannot be a vertex of degree 0 and one of degree $n-1$ in the same graph. So one of these numbers is not realized and all the other numbers in N' are realized exactly once-and the h is realized twice. This implies the required conclusion. □

In view of Assertion 2 we may partition the set J_n into two subsets J_n' and J_n'' consisting of those graphs G_n that satisfy conditions (a) and (b), respectively. We note that H_2' and H_2'', as defined earlier, belong to J_n' and J_n'', respectively. If $G_n \in J_n$, let G_{n-1} denote the graph determined by the vertices in G_n other than the vertex of degree 0 in case (a) or degree $n-1$ in case (b).

Assertion 3. *Suppose* $n \geq 3$.

If $G_n \in J_n'$ *then*

$$(a1) \quad G_n = G_{n-1} \cup K_1, \qquad \textit{where} \quad (a2) \quad G_{n-1} \in J_{n-1}'';$$

if $G_n \in J_n''$ *then*

$$(b1) \quad G_n = G_{n-1} + K_1, \qquad \textit{where} \quad (b2) \quad G_{n-1} \in J_{n-1}'.$$

Proof. Relations $(a1)$ and $(b1)$ follow from Assertion 2 and the definitions of the $\cup$ and $+$ operations. It follows from Assertions 1 and 2 that the unique pair of special vertices u and v of G_n must belong to the subgraph G_{n-1} in both cases (a) and (b). When we consider the relation between the degrees of u and v in G_{n-1} and their degrees in G_n, for cases (a) and (b), we see that u and v must be the unique pair of special vertices in the subgraph G_{n-1}; consequently $G_{n-1} \in J_{n-1}$. In case (a), G_{n-1} cannot have a vertex of degree 0 because if it did there would be two vertices of degree 0 in the graph G_n, and this would contradict Assertion 1. But if G_{n-1} does not have a vertex of degree 0, then alternative (b) must hold when Assertion 2 is applied to G_{n-1} and this implies relation $(a2)$. Similarly, in case (b), G_{n-1} cannot have a vertex of degree $n-2$ because if it did there would be two vertices of degree $n-1$ in the graph G_n and this would contradict Assertion 1. But if G_{n-1} does not have a vertex of degree $(n-1)-1 = n-2$, then alternative (a) must hold when Assertion 2 is applied to G_{n-1} and this implies relation $(b2)$, as required. □

When we combine these Assertions we find that Theorem 3 follows by induction. □

The formal relations (13) describe a straightforward iterative procedure for constructing the extremal graphs H_n' and H_n''. After $n-2$ iterations we obtain expressions for H_n' and H_n'' in terms of $H_2' = K_1 \cup K_1$ and $H_2'' = K_1 + K_1$, respectively, when n is even; or in terms of $K_1 + K_1$ and $K_1 \cup K_1$, when n is odd. The following examples illustrate the pattern by which these graphs develop.

$$H'_2 = K_1 \cup K_1 \qquad H''_2 = K_1 + K_1$$
$$H'_3 = (K_1 + K_1) \cup K_1 \qquad H''_3 = (K_1 \cup K_1) + K_1$$
$$H'_4 = ((K_1 \cup K_1) + K_1) \cup K_1 \qquad H''_4 = ((K_1 + K_1) \cup K_1) + K_1$$
$$H'_5 = (((K_1 + K_1) \cup K_1) + K_1) \cup K_1 \qquad H''_5 = (((K_1 \cup K_1) + K_1) \cup K_1) + K_1.$$

Each symbol K_1 in these expressions represents a vertex in the corresponding graph H'_n or H''_n. If w and z denote the vertices represented by two particular copies of K_1, where the copy representing w is to the left of the copy representing z, then w and z are joined in the corresponding graph if and only if the operation immediately preceding the copy representing z is the sum symbol. The two vertices that were in the initial graphs $H'_2 = K_1 \cup K_1$ and $H''_2 = K_1 + K_1$ are the pair of special vertices u and v in all the graphs H'_n and H''_n derived from them.

Corollary 1. *If $n \geq 2$ and $G_n \in J_n$, let $D(G_n)$ denote the common value of the degrees d_u and d_v of the unique pair of special vertices u and v of G_n.*

(C1) If $n = 2k$ or $2k + 1$, where $k \geq 1$, then $D(H''_n) = k$.

(C2) If $n = 2$, then $D(H'_2) = 0$; and if $n = 2k+1$ or $2k+2$, where $k \geq 1$, then $D(H'_n) = k$.

Proof. One way to prove this is to apply the defining relations for H''_n and H'_{n-1} to conclude that $H''_n = H'_{n-1} + K_1 = (H''_{n-2} \cup K_1) + K_1$, for $n \geq 4$. This implies that $D(H''_n) = D(H''_{n-2}) + 1$ for $n \geq 4$, upon appealing to the definitions of the graph union and graph sum operations. Since $D(H''_2) = D(H''_3) = 1$, conclusion (C1) now follows by induction. Since $H'_n = H''_{n-1} \cup K_1$, by definition, it follows that $D(H'_n) = D(H''_{n-1})$ for $n \geq 3$. So conclusion (C2) now follows from conclusion (C1). This suffices to complete the proof of the Corollary since the result for $D(H'_2)$ is obviously true.

$\square$

We may also note that it follows from the definition of these graphs that these two special vertices are both joined to the same vertices other than themselves; and these other vertices all have different degrees. It follows, therefore, that the number of different ways to label the vertices of H'_n and H''_n is $n!/2$, for each of them. Consequently, the total number of labeled graphs G_n with exactly one pair of special vertices is $n!$, for $n \geq 2$. By coincidence this the same as the number of labeled tournaments T_n with no two vertices with the same score.

Acknowledgements

We thank Noga Alon for referring us to the work of Paul Erdős and Robin J. Wilson. We also thank Boris Alemi for executing the Maple program on a powerful computer and obtaining the data in the required format. The research of YM was supported by grant no. 2020063 from the United States–Israel Binational Science Foundation (BSF).

References

Aldous, D. J., Kolesnik, B. (2022). To stay discovered: On tournament mean score sequences and the Bradley-Terry model. Stochastic Process. Appl. **1**50, 844–852.

Alon, N. (2022). Personal comunication.

Bollobás, B. (1981). Degree sequences of random graphs. Discrete Math. **3**3, 1–19.

Bollobás, B. (2001). Random graphs. Second edition. Cambridge University Press, Cambridge.

David, H. A. (1959). Tournaments and Paired Comparisons Biometrika **4**6, 139–149.

David, H. A. (1988). The method of paired comparisons. Second edition. Charles Griffin & Co., Ltd., London; The Clarendon Press, Oxford University Press, New York.

David, H. A., Edwards, A. W. F. (2001).The Evaluation of Tournament Outcomes: Comments on Zermelo (1929). Annotated readings in the history of statistics. Springer-Verlag, New York, 161–166.

Epstein, R. A. (1967). The theory of gambling and statistical logic. First edition. Academic Press, New York.

Epstein, R. A. (2013). The theory of gambling and statistical logic. Special second edition. Elsevier/Academic Press, Amsterdam.

Erdős, P., Wilson, R. J. (1977). On the chromatic index of almost all graphs. J. Combinatorial Theory Ser. B **2**3, 255–257.

Feller, W. (1968). An introduction to probabilty theory and its applications. Vol. I. Third edition. New York-London-Sydney: Wiley.

Frieze, A., Karoński, M. (2016). Introduction to random graphs. Cambridge University Press, Cambridge.

Guy, R. K. (1984). A pentagonal pot-pourri of perplexing problems, primarily probabilistic. Amer. Math. Monthly **9**1, 559–563.

Hamming, R. W. (1968). Book review. Games of Chance: The theory of gambling and statistical logic. Richard A. Epstein. Academic Press, New York, 1967. xvi + 492 pp. Science**1**61, Issue 3844, 878.

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. **3**4, 92–94.

MacMahon, P. A. (1923). An American tournament treated by the calculus of symmetric functions. Quart. J. Pure Appl. Math. **v**ol. XLIX, No. 193, 1–36. Reprinted in Percy Alexander MacMahon Collected Papers, Vol. I, George E. Andrews, ed., MIT Press, 1978, pp. 308–343.

Malinovsky, Y., Moon, J. W. (2022). On the negative dependence inequalities and maximal score in round-robin tournament. Statist. Probab. Lett. **1**85, 109432.

Malinovsky, Y., Rinott, Y. (2022). On tournaments and negative dependence. J. Appl. Probab. Accepted. `https://arxiv.org/abs/2206.08461`

Metropolis, N., Ulam, S. (1949). The Monte Carlo method. J. Amer. Statist. Assoc. **4**4, 335–341.

Moon, J. W. (1968). Topics on Tournaments. Holt, Rinehart and Winston, New York. [A version (2013) with some corrections is publicly available at `https://www.gutenberg.org/ebooks/42833`].

Stockmeyer, P. K. (2022). Counting Various Classes of Tournament Score Sequences. arXiv:2202.05238v1 [math.CO] 10 Feb 2022.

Zeilberger, D. (2016). On the Most Commonly-Occuring Score Vectors of American Tournaments of n-players, and their Corresponding Records. `https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/Percy.html`

Zermelo, E. (1929). Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. Math. Z. **2**9, 436–460.