An unexpected characteristic of tournament predictive power

LEONARD M. WAPNER

Introduction

A tournament is a series of games (matches) among competitors for an overall prize. Beyond sporting events, tournament structure can relate to elections, applicants competing for an employment position, etc. Tournament structure also exists within certain species of birds and mammals where dominance relations develop, forming a tournament-like pecking structure [1]. The format of the tournament depends on the tournament's objective giving consideration to time, financial, geographical, and other constraints. The two most common formats where players compete pairwise are the *single elimination* (knockout, sudden death) tournament and the *round robin* tournament. Our discussion includes both. A third popular format is that of a contest, where all players perform simultaneously only once. Examples of this type include most golf tournaments as well as track and field athletic events. A detailed analysis of contests as they relate to topics discussed here is given in [2].

A simple probability model, first proposed by Zermelo [3], reveals a counterintuitive characteristic of a tournament's *predictive power*—the likelihood of the strongest player winning the tournament. Tournaments considered here involve only four players, leaving open possibilities for generalisation to larger tournaments and other formats.

The model

The relative strengths of players in head-to-head competition with each other are represented by positive numbers. Assume players X and Y have respective strengths $x > 0$ and $y > 0$. We denote X being at least as strong as Y by $X \geqslant Y$ in which case $x \geqslant y$. We then approximate the probability of X beating Y competing head-to-head by

$$P_{XY} = \frac{x}{x + y}.$$

The following inequalities should be clear.

$$X \geqslant Y \implies P_{XY} \geqslant P_{YX} = 1 - P_{XY} \tag{1}$$

$$X \geqslant Y \geqslant Z \Rightarrow P_{XZ} \geqslant P_{XY}, P_{XZ} \geqslant P_{YZ}, P_{ZY} \geqslant P_{ZX} \text{ and } P_{YX} \geqslant P_{ZX}. \tag{2}$$

Fairness relating to monotonicity

A *monotonic function* is a function which is either entirely non-increasing or non-decreasing. Let $T(X)$ represent the probability of X winning a tournament under consideration. Its value depends on the

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press

tournament's structure and the strengths of the players. With fairness in mind, a desirable monotonicity property of any tournament would be

$$A \succcurlyeq B \implies T(A) \geqslant T(B) \tag{3}$$

indicating that a stronger player has no less chance of winning the tournament than a weaker player.

Single elimination tournaments

A single elimination tournament ideally consists of n (a power of 2) players paired up for the first round of play. Winners advance to the next round and losers are eliminated from competition. Ultimately the two surviving winners play each other in the final round determining the tournament winner. The full tournament involves $\log_2 n$ rounds consisting of $n - 1$ games in total. Compared to other tournament formats, single elimination tournaments involve a relatively small number of games which can be completed in a shorter time involving lesser cost. The tournament will produce a unique winner with no draws to resolve.

But depending on how the players are initially paired, the tournament may or may not possess the desired monotonicity as defined in (3) above. Consider players A, B, C and D having strengths $a = 8$, $b = 7$, $c = 6$ and $d = 1$. Assume A plays B and C plays D in the first round, with the winners advancing to the second round. The probabilities of each player winning the single elimination tournament are denoted below by T_S.

$$\begin{aligned}
T_S(A) &= P_{AB}P_{CD}P_{AC} + P_{AB}P_{DC}P_{AD} \\
&= \frac{8}{8+7}\cdot\frac{6}{6+1}\cdot\frac{8}{8+6} + \frac{8}{8+7}\cdot\frac{1}{1+6}\cdot\frac{8}{8+1} \approx 0.33 \\
T_S(B) &= P_{BA}P_{CD}P_{BC} + P_{BA}P_{DC}P_{BD} \\
&= \frac{7}{7+8}\cdot\frac{6}{6+1}\cdot\frac{7}{7+6} + \frac{7}{7+8}\cdot\frac{1}{1+6}\cdot\frac{7}{7+1} \approx 0.27 \\
T_S(C) &= P_{CD}P_{AB}P_{CA} + P_{CD}P_{BA}P_{CB} \\
&= \frac{6}{6+1}\cdot\frac{8}{8+7}\cdot\frac{6}{6+8} + \frac{6}{6+1}\cdot\frac{7}{7+8}\cdot\frac{6}{6+7} \approx 0.38 \\
T_S(D) &= P_{DC}P_{AB}P_{DA} + P_{DC}P_{BA}P_{DB} \\
&= \frac{1}{1+6}\cdot\frac{8}{8+7}\cdot\frac{1}{1+8} + \frac{1}{1+6}\cdot\frac{7}{7+8}\cdot\frac{1}{1+7} \approx 0.02.
\end{aligned}$$

Paired this way, third ranked C has a better chance of winning the tournament than higher ranked A and B. This may be undesirable as monotonicity (3) is violated by C being granted an unjustified advantage. For this and other reasons, pairing the two strongest players in the first round should be avoided, which is possible only if we can approximate the

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press

strengths of the players. Past win-loss history is one way of approximating these strengths. A *seeding* process can then be used to assign all players a preliminary rank and then pairing them to minimise the chance of such non-monotonicity and increase the likelihood that the strongest players will proceed to the later round. Detailed discussions of optimal seeding are given in [4] and [5].

But it is possible that we have no knowledge of the players' strengths in which case the primary objective of the tournament may be to determine the strongest player. In this case, first round pairing should be considered random relative to the players' strengths. Fortunately, random first round pairing preserves monotonicity and, in this sense, a tournament so designed should be considered fair. The proof of monotonicity follows.

Since each of the three possible first round pairings occurs with probability 1/3, the probability of A winning the tournament is given by

$$T_S(A) = \frac{1}{3}(P_{AD}P_{BC}P_{AB} + P_{AD}P_{CB}P_{AC}) + \frac{1}{3}(P_{AC}P_{BD}P_{AB} + P_{AC}P_{DB}P_{AD}) + \frac{1}{3}(P_{AB}P_{CD}P_{AC} + P_{AB}P_{DC}P_{AD}). \tag{4}$$

Similarly, the probability of B winning is

$$T_S(B) = \frac{1}{3}(P_{BC}P_{AD}P_{BA} + P_{BC}P_{DA}P_{BD}) + \frac{1}{3}(P_{BD}P_{AC}P_{BA} + P_{BD}P_{CA}P_{BC}) + \frac{1}{3}(P_{BA}P_{CD}P_{BC} + P_{BA}P_{DC}P_{BD}).$$

Showing $A \geqslant B \implies T_S(A) \geqslant T_S(B)$ is equivalent to showing $A \geqslant B$ implies the following inequality:

$$\frac{1}{3}\left[\left(P_{AD}P_{BC}P_{AB} - P_{BC}P_{AD}P_{BA}\right) + \left(P_{AD}P_{CB}P_{AC} - P_{BD}P_{CA}P_{BC}\right) + \left(P_{AC}P_{BD}P_{AB} - P_{BD}P_{AC}P_{BA}\right) + \left(P_{AC}P_{DB}P_{AD} - P_{BC}P_{DA}P_{BD}\right) + \left(P_{AB}P_{CD}P_{AC} - P_{BA}P_{CD}P_{BC}\right) + \left(P_{AB}P_{DC}P_{AD} - P_{BA}P_{DC}P_{BD}\right)\right] \geqslant 0.$$

The expressions within each pair of parentheses are shown to be non-negative by (1) and (2), establishing the desired inequality.

Bear in mind that after the pairing is selected and play begins, the tournament may no longer be monotonic as there is the chance that a less than ideal pairing (such as the two strongest players facing each other in the first round) is selected. But if we consider the random pairing as part of the tournament then the combination of random pairing and actual play is, as a whole, monotonic and fair.

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press

An unexpected occurrence of non-monotonicity

Investigating monotonicity in a different context unearths an unexpected characteristic of single elimination tournaments. One might expect $T_S(A)$ to decrease monotonically as the strength of B increases, assuming the strengths of C and D remain constant. Oddly, this is not necessarily so. Let the strengths of A, C and D be $a = 2$, $c = 1$ and $d = 1$ respectively. Allow b, the strength of B, to increase from 0. By random first round pairing and (4)

$$\begin{aligned} T_S(A) = f(b) &= \frac{1}{3}\left(\frac{2}{2+1}\cdot\frac{b}{b+1}\cdot\frac{2}{2+b} + \frac{2}{2+1}\cdot\frac{1}{1+b}\cdot\frac{2}{2+1}\right) \\ &+ \frac{1}{3}\left(\frac{2}{2+1}\cdot\frac{b}{b+1}\cdot\frac{2}{2+b} + \frac{2}{2+1}\cdot\frac{1}{1+b}\cdot\frac{2}{2+1}\right) \\ &+ \frac{1}{3}\left(\frac{2}{2+b}\cdot\frac{1}{1+1}\cdot\frac{2}{2+1} + \frac{2}{2+b}\cdot\frac{1}{1+1}\cdot\frac{2}{2+1}\right) \\ &= \frac{44b+28}{27(b^2+3b+2)}. \end{aligned}$$

As expected, $T_S(A) \to \infty$ as $b \to 0$. But note the non-monotonic behaviour of as shown in Figure 1.

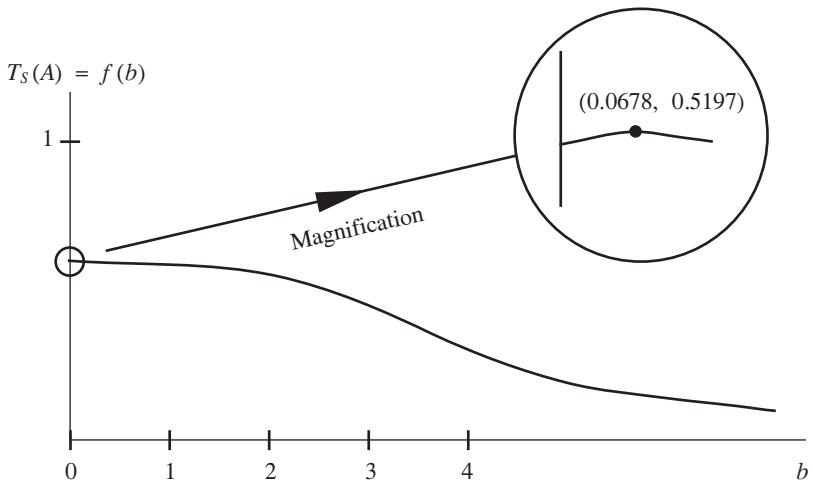


FIGURE 1: Non-monotonic behviour of $T_S(A) = f(b)$

As the strength of B increases from 0, the probability of A winning the tournament initially increases! To see this we can evaluate $f'(b)$,

$$f'(b) = -\frac{4(11b^2+14b-1)}{27(b^2+3b+2)^2} = 0 \Rightarrow b = \frac{2\sqrt{15}-7}{11} \approx 0.0678.$$

(The negative root is meaningless and discarded.) So $T_S(A) = f(b)$ achieves its maximum at a relatively small value of b when compared to

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press

strengths of the other players. For larger values of b, $T_S(A) = f(b)$ approaches 0 monotonically.

An intuitive explanation of the non-monotonic behaviour comes from considering what happens when B, the weakest player, initially gains strength. With probability 1/3, B is paired with C in the first round of play and the chance of the weakest player advancing to the final round increases, giving A (and D) a better chance of winning the overall tournament. Hence the predictive power of the tournament initially increases.

The phenomenon exists in non-sporting competitions as well. As of January 2020, it was likely that Republican incumbent United States President Donald Trump would run against one of 12 major Democratic candidates in the November general election. To become the Democratic nominee, a candidate must first win enough state primary elections to gain support for the nomination. Polls indicated that former Democratic Vice-President Joe Biden was the Democrat most likely to beat Donald Trump in the general election. President Trump's chances of winning a second term may have increased when weak Democratic candidates gained popularity, as this increased the likelihood of a weaker than presumed Democratic candidate being nominated.

Round robin tournaments

In a round robin tournament each competitor plays every other competitor an equal number of times. The winner of the tournament is the player having the most wins, assuming there are no draws. With n competitors playing each other once, there are

$$\binom{n}{2} = \frac{n!}{2!\,(n-2)!}$$

games in total. This tournament format avoids the fairness issues related to the first round pairings of single elimination tournaments. And, depending on how draws are resolved, the predictive power of this format may exceed that of single elimination tournaments. (This will be further addressed at this Article's conclusion.) The primary disadvantage is the significantly greater number of games required when compared to single elimination, even if each competitor plays others only once.

Given the balanced nature of round robin tournaments and the complexity of resolving draws, it would be natural to employ the *method of paired comparisons*, where items (in this case players) are compared two at a time to develop an overall ranking of all. This technique having origins dating back to the mid-1800s was first applied to round robin tournaments by Kendall [6]. However, in keeping with the focus of this article, we use the probability model applied earlier in this article to single elimination tournaments.

With first round pairings being irrelevant, one might not expect the odd nonmonotonic behaviour of predictive power associated with single

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press

elimination tournaments. Surprisingly, this behaviour may still occur, depending on how draws are resolved. To show this, assume competitors A, B, C and D play each other once, where a total of six games would be required, with each competitor playing three games. A player wins the tournament in one of three ways, with draws being resolved as indicated. For each case the probability of player A winning the tournament is given.

Case 1) Any player achieving a win-loss record of 3-0 wins the tournament. The probability of A winning this way is $P_{AB}P_{AC}P_{AD}$. The outcomes of the tournament's other three games are irrelevant.

Case 2) If a two-way draw occurs by the top two players achieving 2-1 records, then the winner of the tournament would be the winner of the game played when the two players met head-to-head. Player A can achieve this in six possible ways as indicated by the sum of six products of probabilities

$$P_{AB}P_{BC}P_{BD}P_{AC}P_{CD}P_{DA} + P_{AB}P_{BC}P_{BD}P_{CA}P_{DC}P_{AD} + P_{AC}P_{CB}P_{CD}P_{AB}P_{BD}P_{DA}$$
$$+P_{AC}P_{CB}P_{CD}P_{BA}P_{DB}P_{AD} + P_{AD}P_{DB}P_{DC}P_{AB}P_{BC}P_{CA} + P_{AD}P_{DB}P_{DC}P_{BA}P_{CB}P_{AC}.$$

Case 3) A three-way draw resulting from three players achieving 2-1 records is more problematic. If time allows, a playoff could determine the tournament winner. Otherwise, the winner could be selected randomly from the three tied players. For simplicity we use the latter method here. There are six possible three-way draws which include player A. If included, player A has a probability of 1/3 of winning the tournament. So the probability of A winning the tournament via a three-way draw is

$$\tfrac{1}{3}\big(P_{AB}P_{CB}P_{DB}P_{AC}P_{CD}P_{DA} + P_{AB}P_{CB}P_{DB}P_{CA}P_{DC}P_{AD} + P_{AC}P_{BC}P_{DC}P_{AB}P_{BD}P_{DA}$$
$$+P_{AC}P_{BC}P_{DC}P_{BA}P_{DB}P_{AD} + P_{AD}P_{BD}P_{CD}P_{AB}P_{BC}P_{CA} + P_{AD}P_{BD}P_{CD}P_{BA}P_{CB}P_{AC}\big).$$

Overall, the probability of A winning the round robin tournament, denoted by $T_R(A)$, is the sum of the three probabilities given in Cases 1, 2 and 3.

$$T_R(A) =$$
$$P_{AB}P_{AC}P_{AD} + P_{AB}P_{BC}P_{BD}P_{AC}P_{CD}P_{DA} + P_{AB}P_{BC}P_{BD}P_{CA}P_{DC}P_{AD}$$
$$+P_{AC}P_{CB}P_{CD}P_{AB}P_{BD}P_{DA} + P_{AC}P_{CB}P_{CD}P_{BA}P_{DB}P_{AD}$$
$$+P_{AD}P_{DB}P_{DC}P_{AB}P_{BC}P_{CA} + P_{AD}P_{DB}P_{DC}P_{BA}P_{CB}P_{AC}$$
$$+\tfrac{1}{3}\big(P_{AB}P_{CB}P_{DB}P_{AC}P_{CD}P_{DA} + P_{AB}P_{CB}P_{DB}P_{CA}P_{DC}P_{AD} + P_{AC}P_{BC}P_{DC}P_{AB}P_{BD}P_{DA}$$
$$+P_{AC}P_{BC}P_{DC}P_{BA}P_{DB}P_{AD} + P_{AD}P_{BD}P_{CD}P_{AB}P_{BC}P_{CA} + P_{AD}P_{BD}P_{CD}P_{BA}P_{CB}P_{AC}\big). \quad (5)$$

An unexpected relationship between the two formats

If the expressions for $T_S(A)$ and $T_R(A)$ as given in (4) and (5) are written

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press

in terms of a, b, c and d, they are seen to be equal.

$$T_S(A) = T_R(A)$$

$$= \frac{a^2\left[2c^2d^2 + 6bcd(c+d) + 3a(b+c)(b+d)(c+d) + 2b^2\left(c^2 + 3cd + d^2\right)\right]}{3(a+b)(a+c)(a+d)(b+c)(b+d)(c+d)}.$$

This equality is highly unexpected. A similar equality holds for players B, C and D, implying the predictive powers of the two tournament formats are equal and the non-monotonic behaviour of predictive power noted for single elimination tournaments also occurs for the round robin format given here.

Closing remarks

In some cases, a player stands a greater chance of winning a tournament when an opponent's strength slightly increases. Reasoning inversely, it is not always in one's best interest when a competitor weakens. This applies to both sporting and non-sporting competitions.

The equality of predictive powers for the two formats is highly unexpected and is almost certainly not coincidental. The author is unable to provide an intuitive explanation for this result, which relies on round robin three-way draws being here resolved by random selection. If such draws were resolved by a playoff requiring additional games, then the strongest of the three tied players would be most likely to win the playoff. As a result, the predictive power of the round robin format would then exceed that of a single elimination tournament.

Finally, others have noted counterintuitive non-monotonicities of a different sort associated with all three tournament types (single elimination, round robin, contests) under discussion. For example, for Pareto distributions of player strengths, the predictive power of all three formats is non-monotonic with respect to N, the number of players. For contests with Pareto distributions of strengths, the predictive power approaches 1 as n approaches infinity [2].

Acknowledgement

The author gratefully acknowledges the anonymous referee for direction and constructive suggestions.

References

1. F. Harary and L. Moser, The theory of round robin tournaments, *The American Mathematical Monthly* **73**(3) (1966) pp. 231-246.
2. D. Ryvkin and A. Ortmann, The predictive power of three prominent tournament formats, *Management Science* **54**(3) (2008) pp. 492-504.
3. E. Zermelo, Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung, *Mathematische Zeitschrift* **29** (1929) pp. 436-460.

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press

4. A. Schwenk, What is the correct way to seed a knockout tournament?, *The American Mathematical Monthly* **107**(2) (2000) pp. 140-150.
5. C. Groh, B. Moldovanu, A. Sela, and U. Sunde, Optimal seedings in elimination tournaments, *Economic Theory* **49**(1) (2012) pp. 59-80.
6. M. Kendall, Further contributions to the theory of paired comparisons, *Biometrics* **11**(1) (1955) pp. 43-62.

10.1017/mag.2021.47
LEONARD M. WAPNER
© The Mathematical Association 2021
Division of Mathematical Sciences,
El Camino College, 16007 Crenshaw Boulevard, Torrance, CA 90506, USA
e-mail: *lwapner@elcamino.edu*

The answers to the *Nemo* page from March 2021 on parallelograms were:

1.	Joseph Conrad	The Inheritors	Chapter 8
2.	Anthony Trollope	Can you Forgive Her?	Chapter 2
3.	HG Wells	The Invisible Man	Chapter 26
4.	Thomas Hardy	A Pair of Blue Eyes	Chapter 11
5.	Sheridan Le Fanu	Guy Deverell	Chapter 2
6.	GK Chesterton	Tremendous Trifles	The Perfect Game

Congratulations to Ian Anderson and Martin Lukarevski on tracking all of these down. In this issue we return to a familiar theme: mathematics and everyday life. The quotations are to be identified by reference to author and work. Solutions are invited to the Editor by 23rd September 2021

1. Histories make men wise; poets witty; the mathematics subtle; natural philosophy deep; moral grave; logic and rhetoric able to contend.

2. Every man who is not a monster, a mathematician, or a mad philosopher, is the slave of some woman or other.

3. As for my father, I never knew whether he believed God was a mathematician but he certainly believed God could count and that only by picking up God's rhythms were we able to regain power and beauty.

4. I would by no means wish a daughter of mine to be a progeny of learning; I don't think so much learning becomes a young woman: for instance, I would never let her meddle with Greek, or Hebrew, or algebra, or simony, or fluxions, or paradoxes, or such inflammatory branches of learning; nor will it be necessary for her to handle any of your mathematical, astronomical, diabolical instruments.

Continued on page 281.

https://doi.org/10.1017/mag.2021.47 Published online by Cambridge University Press