*Probability in the Engineering and Informational Sciences* (2022), **36**:4 988–998
doi:10.1017/S026996482100019X

CAMBRIDGE UNIVERSITY PRESS

**RESEARCH ARTICLE**

# Team's seasonal win probabilities

Sheldon M. Ross

Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089, USA.
E-mail: smross@usc.edu.

**Keywords:** Applied probability, Simulation, Stochastic modeling

**Abstract**
We consider a competition involving $r$ teams, where each individual game involves two teams, and where each game between teams $i$ and $j$ is won by $i$ with probability $P_{i,j} = 1 - P_{j,i}$. We suppose that $i$ and $j$ are scheduled to play $n(i, j)$ games and say that the team that wins the most games is the winner of the competition. We show that the conditional probability that $i$ is the winner, given that $i$ wins $k$ games, is increasing in $k$. We bound the tail probability of the number of wins of the winning team. We consider the special case where $P_{i,j} = v_i/(v_i + v_j)$, and obtain structural results on the probability that team $i$ is the winner. We give efficient simulation approaches for computing the probability that team $i$ is the winner, and the conditional probability given the number of wins of $i$.

## 1. Introduction and summary

We consider a competition involving $r$ teams, where each individual game involves two teams, and where each game between teams $i$ and $j$ is, independent of results of other games, won by $i$ with probability $P_{i,j} = 1 - P_{j,i}$. We suppose that some games may have already been played, that team $i$ currently has a total of $b_i$ wins, $i = 1, \ldots, r$, and that $i$ and $j \neq i$ are scheduled to play an additional $n(i, j) = n(j, i)$ games. We let $W(i, j)$ be the number of these $n(i, j)$ games in which $i$ beats $j$. With $W(i, i) = 0$ for notational convenience, also let $W_i = \sum_j W(i, j)$ be the total number of additional games won by team $i$.

In Section 2, we show for any constants $a_t, t \neq i$, that $P(W_i \geq \max_{t\neq i}(W_t + a_t) \mid W_i = k)$ is an increasing function of $k$, and that $W_i|W_j = k$ is, for $i \neq j$, stochastically decreasing in $k$.

In Section 3, we consider the random variable $\max_i W_i$ and obtain bounds on its tail distribution.

In Section 4, we consider the Bradley–Terry special case model, which assumes that team $i$ has a positive value $v_i$, $i = 1, \ldots, r$, and that $P_{i,j} = v_i/(v_i + v_j)$. We show that $P(W_i \geq k)$ is an increasing function of $v_i$ and a decreasing function of $v_j, j \neq i$, and that whereas $P(W_i \geq \max_{t\neq i}(W_t + a_t))$ is an increasing function of $v_i$, it need not be true, even when all $n(i, j) = n$, that it is a decreasing function of $v_j, j \neq i$ or that $P(W_i \geq \max_{t\neq i}(W_t + a_t)|W_j = k)$ is a decreasing function of $k$ when $j \neq i$. We then focus on the case where all $n(i, j) = n$, and show in this case that $v_i > v_j$ implies that $P(W_i \geq \max_{t\neq i} W_t + a) > P(W_j \geq \max_{t\neq j} W_t + a)$.

In Section 5, we present efficient simulation techniques for estimating all the win probabilities $P(W_i \geq \max_{t\neq i}(W_t + a_t))$ as well as $P(W_i \geq \max_{t\neq i}(W_t + a_t) \mid W_i = k)$.

Most of the literature involving competitions between teams assumes that a fixed number of losses eliminates a team from further competition. Indeed the majority consider the knockout format, in which a single loss eliminates the team. Many knockout models assume that each of the teams is given a position, and then specify a structure defining the sequence of games to be played. For instance, a structure with $r = 6$ players might in round 1 have a match between those in positions 1 and 2 and one between those in positions 3 and 4; round 2 might involve a game between the round 1 winners as well as a game between the teams in positions 5 and 6; the two winners in round 2 games would then play

© The Author(s), 2021. Published by Cambridge University Press

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

in round 3 to determine the winner of the tournament. Many of the models assume a Bradley–Terry model for team strengths, and also that the team positions are randomly and uniformly determined. Such knockout models are said to have a random structured format.

Chung and Hwang [3] proved, for any random structured format, that if $v_i > v_j$, then the probability that team $i$ wins the knockout tournament is greater than the probability that team $j$ wins. Maurer [12] was interested in determining the random structured format that maximized the probability that the best team (i.e., the one with a maximal value of $v_i$) wins the tournament and gave the result when $v_1 > v = v_2 = v_3 = \cdots = v_n$. Cao and Ross [2] considered the format where the teams are randomly permuted, the first two then playing a game, with the winner playing the next in line, and so on, with the overall winner being the team that wins the last game. Assuming a Bradley–Terry model, bounds on the win probabilities of each team were given in [2]. A different type of random knockout tournament was considered in Adler *et al.* [1], where it was assumed that the number of matches in each round was fixed, and that the matchup of teams in any round was totally random. That is, if there are $m_i$ games in round $i$, $\sum_i m_i = r - 1$, then the $2m_1$ teams that play in round 1 are randomly chosen from all $r$ teams and then randomly paired up. The winners of this round, along with the $r - 2m_1$ teams that did not play in round 1, then move to round 2, where $2m_2$ of them are randomly chosen and randomly paired to play round 2 games, and so on. Assuming a Bradley–Terry model, it was shown that if $v_1 = \max_i v_i$, that $P(1 \text{ wins the tournament}) \geq v_1 / \sum_{i=1}^{r} v_i$. Bounds on the win probabilities of the other teams is also given in [1]. Some other papers studying knockout tournaments are David [4], Glenn [6], Hennessy and Glickman [7], Horen and Riezman [8], Hwang [9], and Marchand [10].

Double elimination tournaments are similar to knockout tournaments, with the difference being that teams are eliminated not after their first loss but after they have lost twice. Edwards [5] discusses such tournaments. Round robin type tournaments, where each team plays each of the other teams once, have also been studied (see, e.g., [13,17,18]), but they typically assume only a few teams.

## 2. General model: structural results

We consider the general $P_{i,j}$ model, with our first objective being to show that a team's conditional probability of having the most wins is an increasing function of the number of wins of that team. To establish this, we will use Efron's theorem.

**Lemma 1** (Efron's Theorem)**.** *If $X_1, \ldots, X_r$ are independent log-concave random variables, then $(X_1, \ldots, X_r) \mid \sum_{t=1}^{r} X_t = k$ is stochastically increasing in $k$. That is, for any component-wise increasing function $g(x_1, \ldots, x_r)$*

$$E\left[ g(X_1, \ldots, X_r) \,\middle|\, \sum_{t=1}^{r} X_t = k \right] \text{ is increasing in } k$$

*Let $\mathbf{W}_{-i} = (W_1, \ldots, W_{i-1}, W_{i+1}, \ldots, W_r)$.*

**Proposition 1.** *$\mathbf{W}_{-i} \mid W_i = k$ is stochastically decreasing in $k$.*

*Proof.* Fix $i$. For $s \neq i$, let $W_s^* = n(s,i) + \sum_{t \neq i} W(s,t)$ be the number of wins of team $s$ against all teams aside from $i$ plus the number of times team $s$ plays team $i$. (Thus, $W_s = W_s^* - W(i,s)$.) Then, if $g$ is an increasing function,

$$\begin{aligned} &E[g(\mathbf{W}_{-i}) \mid W_i = k, W_s^* = w_s^*, s \neq i] \\ &\quad = E[g(w_1^* - W(i,1), \ldots, w_r^* - W(i,r)) \mid W_i = k, W_s^* = w_s^*, s \neq i] \\ &\quad = E[g(w_1^* - W(i,1), \ldots, w_r^* - W(i,r)) \mid W_i = k] \end{aligned} \tag{1}$$

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

where (1) follows because $(W_s^*, s \neq i)$ is independent of $(W(i,1), \ldots, W(i,r))$. Because $W(i,s), s \neq i$ are independent binomial, and thus log-concave, random variables, it follows from (1) and Efron's theorem that $E[g(\mathbf{W}_{-i}) \mid W_i = k, W_s^* = w_s^*, s \neq i]$ is a decreasing function of $k$, and thus that $E[g(\mathbf{W}_{-i}) \mid W_i = k, W_s^*, s \neq i]$ is too. Taking expectations, given $W_i = k$, now yields that $E[g(\mathbf{W}_{-i}) \mid W_i = k] = E[E[g(\mathbf{W}_{-i}) \mid W_i = k, W_s^*, s \neq i] \mid W_i = k]$ is a decreasing function of $k$. □

**Corollary 1.** *For given constants $a_t, t \neq i$,*

$$P\left(W_i \geq \max_{t \neq i}(W_t + a_t) \mid W_i = k\right) \quad \textit{is an increasing function of } k.$$

*Proof.*

$$\begin{aligned} P\left(W_i \geq \max_{t \neq i}(W_t + a_t) \mid W_i = k\right) &= P\left(\max_{t \neq i}(W_t + a_t) \leq k \mid W_i = k\right) \\ &\leq P\left(\max_{t \neq i}(W_t + a_t) \leq k+1 \mid W_i = k\right) \\ &\leq P\left(\max_{t \neq i}(W_t + a_t) \leq k+1 \mid W_i = k+1\right) \quad \text{by Proposition 1} \\ &= P\left(W_i \geq \max_{t \neq i}(W_t + a_t) \mid W_i = k+1\right) \end{aligned}$$

□

**Remark.** *Because team $j$ ends up with a total of $b_j + W_j$ wins, Proposition 1 shows that the probabilities that no team wins more games than team $i$ and that team $i$ uniquely has the most wins both increase in the number of wins of $i$. (Set $a_t = b_t - b_i$ for the former, and $a_t = b_t - b_i + 1$, for the latter result.)*

**Corollary 2.**

*(a)* $\mathbf{W}_{-i} \mid W_i \geq k$ *and* $\mathbf{W}_{-i} \mid W_i \leq k$ *are both stochastically decreasing in* $k$.
*(b)* $\mathbf{W}_{-i} \mid W_i \geq k \leq_{\text{st}} \mathbf{W}_{-i} \leq_{\text{st}} \mathbf{W}_{-i} \mid W_i \leq k$

*Proof.* Let $g$ be an increasing function of $r-1$ components. Also, let $f(s) = E[g(\mathbf{W}_{-i}) \mid W_i = s]$, and note that $f$ is, by Proposition 1, a decreasing function. Now,

$$\begin{aligned} E[g(\mathbf{W}_{-i}) \mid W_i \geq k] &= \sum_{s=0}^{k} E[g(\mathbf{W}_{-i}) \mid W_i = s] P(W_i = s \mid W_i \geq k) \\ &= \sum_{s=0}^{k} f(s) P(W_i = s \mid W_i \geq k) \\ &= E[f(W_i(k))] \end{aligned}$$

where $W_i(k)$ is a random variable distributed as $W_i \mid W_i \geq k$. Because $W_i(k)$ stochastically increases in $k$ (i.e., $W_i(k+1) \geq_{\text{st}} W_i(k)$,) and $f$ is decreasing, it follows that $E[f(W_i(k))]$ decreases in $k$, showing that $\mathbf{W}_{-i} \mid W_i \geq k$ is stochastically decreasing in $k$. The proof that $\mathbf{W}_{-i} \mid W_i \leq k$ is stochastically decreasing in $k$ is similar. Letting $n_i$ denote the total number of games played by team $i$, part (b) follows from part (a) since

$$\mathbf{W}_{-i} | W_i \geq 0 =_{\text{st}} \mathbf{W}_{-i} =_{\text{st}} \mathbf{W}_{-i} | W_i \leq n_i$$

□

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

## 3. The maximum number of wins

In this section, we show how to bound $P(\max_i W_i > k)$. To begin, note that $W_i$ is the sum of $n_i = \sum_{j:j\neq i} n(i,j)$ independent Bernoulli random variables. Now, in general, the probability mass function of the sum of independent Bernoulli random variables $X_1, \ldots, X_m$, with parameters $p_1, \ldots, p_m$, can be recursively computed by letting

$$P_i(j) = P(X_1 + \cdots + X_i = j), \quad j \leq i \tag{2}$$

Then, starting with $P_1(0) = 1 - p_1$, $P_1(1) = p_1$, recursively compute using

$$P_i(j) = P_{i-1}(j-1)p_i + P_{i-1}(j)(1-p_i), \quad j \leq i,\ i = 2, \ldots, m$$

This enables us to compute the probability mass function of $W_i$ for each $i = 1, \ldots, r$.

**Remark.** *Computing the mass function of $W_i$ by using that it is the sum of $n_i$ independent Bernoulli random variables and then using the preceding recursion appears to be computationally faster than computing it by using that $W_i$ is the sum of $r-1$ independent binomial random variables and then using an analogous recursion for such a sum.*

We can also compute the joint mass function of $W_i$ and $W_j$. For instance, for the joint mass function of $W_1$ and $W_2$ first compute $P_1^*(j), j = 0, \ldots, n_1 - n(1,2)$ and $P_2^*(j), j = 0, \ldots, n_2 - n(1,2)$, the probability mass functions of the number of additional wins of team 1 against all teams aside from 2, and of the number of additional wins of team 2 against all teams aside from 1. Then use

$$P(W_1 = u, W_2 = v) = \sum_{j=0}^{n(1,2)} \binom{n(1,2)}{j} P_{1,2}^j P_{2,1}^{n(1,2)-j} P_1^*(u-j) P_2^*(v - n(1,2) + j)$$

To bound $P(\max_i W_i > k)$, use that $P(\max_i W_i > k) = P(\cup_i \{W_i > k\})$. From Boole's inequality and the conditional expectation inequality, we obtain the bounds

$$\sum_{i=1}^{r} \frac{P(W_i > k)}{1 + \sum_{j\neq i} P(W_j > k \mid W_i > k)} \leq P\left(\max_i W_i > k\right) \leq \sum_{i=1}^{r} P(W_i > k) \tag{3}$$

where the conditional expectation bound (see [14] for details on this bound) can be computed by using that

$$P(W_j > k \mid W_i > k) = \sum_{u>k} \sum_{v>k} P(W_i = u, W_j = v)/P(W_i > k).$$

A computationally simpler lower bound can be obtained by letting, for $j = 1, \ldots, r$, $X_j$ be the indicator of the event that $W_j > k$, and defining $\mathbf{X}_{-i} = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_r)$. Now, (b) of Corollary 2 implies that $\mathbf{X}_{-i} \mid X_i = 1$ is stochastically smaller than $\mathbf{X}_{-i}$. But, from Proposition 2 of [16], this implies that $P(\sum_{i=1}^r X_i = 0) \leq \prod_{i=1}^r P(X_i = 0)$. Consequently,

$$P\left(\max_i W_i \leq k\right) \leq \prod_{i=1}^{r} P(W_i \leq k)$$

Thus, we obtain

**Proposition 2.**

$$1 - \prod_{i=1}^{r} P(W_i \leq k) \leq P\left(\max_i W_i > k\right) \leq \sum_{i=1}^{r} P(W_i > k)$$

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

**Example 1.** *Suppose that* $n(i,j) = r = 10$, *and that all* $p_{i,j} = 1/2$. *That is, there are* 10 *teams with each pair of teams playing* 10 *times, and with each game being equally likely to be won by either team. Thus,* $W_i$ *is binomial with parameters* 90 *and* $1/2$. *Using that*

$$p_k \equiv P(W_i > k) = \begin{cases} 0.000210, & \text{if } k = 60 \\ 0.013151, & \text{if } k = 55 \\ 0.036275, & \text{if } k = 53 \\ 0.123053, & \text{if } k = 50 \end{cases}$$

*and that Proposition 2 yields that* $1 - (1 - p_k)^{10} \le P(\max_i W_i > k) \le 10p_k$, *we see that*

$$0.002098 \le P\left(\max_i W_i > 60\right) \le 0.00210$$
$$0.123994 \le P\left(\max_i W_i > 55\right) \le 0.13151$$
$$0.308915 \le P\left(\max_i W_i > 53\right) \le 0.36275$$
$$0.731012 \le P\left(\max_i W_i > 50\right) \le 1.23053$$

## 4. The Bradley–Terry special case

Suppose now that $P_{i,j} = v_i/(v_i + v_j)$, $i \ne j$. That is, consider the Bradley–Terry model.

**Proposition 3.** *(a)* $P(W_i \ge k)$ *is a strictly increasing function of* $v_i$, *and a strictly decreasing function of* $v_t, t \ne i$.
*(b)* $P(W_i \ge \max_{t \ne i}(W_t + a_t))$ *is an increasing function of* $v_i$.

*Proof.* (a) Note that $W_i$ is the sum of $\sum_j n(i,j)$ independent Bernoulli random variables, of which $n(i,j)$ have parameter $v_i/(v_i + v_t)$, $t \ne i$. Because $v/(v + v_t)$ increases in $v$ and decreases in $v_t$, it thus follows that $W_i$ stochastically increases in $v_i$ and stochastically decreases in $v_t$, $t \ne i$.

The proof of (b) follows by a coupling argument. Consider two scenarios: the first in which team $i$ has value $v_i$ and the second in which team $i$ has value $v_i'$, where $v_i' < v_i$. In both scenarios player $j$, $j \ne i$, has value $v_j$. Couple the results of all games not involving team $i$ to be equal in the two scenarios, and couple the results of games involving team $i$ so that a win of team $i$ in the second scenario implies a win in the first scenario. Consequently, if $W_i \ge \max_{t \ne i}(W_t + a_t)$ in the second scenario, then it is also true in the first scenario, thus proving (b). □

Let us suppose for the rest of this section that all $n(i,j) = n$. That is, each pair of teams play the same number of games.

**Proposition 4.** *Suppose* $n(i,j) = n$. *If* $v_i > v_j$, *then* $P(W_i \ge k) > P(W_j \ge k)$, $k > 0$.

*Proof.* This follows by a simple coupling argument. □

For a given player $i$, let $\mathbf{w} = (w_1, \ldots, w_{r-1})$ be the vector of its opponent's values in the decreasing order; that is, the $r - 1$ opponents of $i$ have values $w_1 \ge w_2 \ge \cdots \ge w_{r-1}$. Say that the opponent value vector $\mathbf{w}$ majorizes the opponent value vector $\mathbf{w}'$, written $\mathbf{w} \ge_m \mathbf{w}'$, if $w_1 + \cdots + w_j \ge w_1' + \cdots + w_j'$, $j = 1, \ldots, r-2$ and $w_1 + \cdots + w_{r-1} = w_1' + \cdots + w_{r-1}'$. We now show that the number of wins of a given player stochastically increases as the values of its opponents become more spread in the sense of majorization.

**Proposition 5.** *Suppose each pair plays* $n$ *games. For fixed* $v_i$, *if* $N_i(\mathbf{w})$ *is the total number of wins of* $i$ *when the opponent vector is* $\mathbf{w}$, *then* $N_i(\mathbf{w}) \ge_{\text{st}} N_i(\mathbf{w}')$ *when* $\mathbf{w} \ge_m \mathbf{w}'$.

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

*Proof.* It follows by the Schur–Ostrowski theorem (see [11]) that it suffices to prove the result when $i$ has only two opponents. Suppose first that $n = 1$ so that $i$ plays exactly one game with each of the other two players. Suppose the opponent values are $w_1 = a + x$, $w_2 = a - x$, $0 \le x \le a$. Then,

$$P(N_i(a+x, a-x) \ge 1) = 1 - \frac{a+x}{a+x+v_i}\frac{a-x}{a-x+v_i} = 1 - \frac{a^2-x^2}{(a+v_i)^2-x^2},$$

is an increasing function of $x$. Furthermore,

$$P(N_i(a+x, a-x) \ge 2) = \frac{v_i}{a+x+v_i}\frac{v_i}{a-x+v_i} = \frac{v_i^2}{(a+v_i)^2-x^2}$$

is also an increasing function of $x$, proving the result when $n = 1$. When $n > 1$, consider $i's$ games against 1 and 2 in $n$ pairs. That is, for each $j = 1, \ldots, n$, consider the number of wins of $i$ in its $j$th games against 1 and 2. Because stochastic monotonicity is preserved under convolution, it follows that the total number of wins of $i$ is stochastically increasing in $x$. □

**Corollary 3.** *Suppose each pair plays n games. Let Bin$(m, p)$ be a binomial random variable with parameters $m, p$. If $\bar{w} = \sum_{j=1}^{r-1} w_j/(r-1)$,*

$$N_i(\mathbf{w}) \ge_{\text{st}} Bin(n(r-1), \bar{w})$$

*Proof.* Follows from Proposition 5 because $\mathbf{w} \ge_m (\bar{w}, \ldots, \bar{w})$. □

**Remarks.** *1. A function $g(\mathbf{w})$ having the property that $g(\mathbf{w}) \ge g(\mathbf{w}')$ whenever $\mathbf{w} \ge_m \mathbf{w}'$ is said to be a Schur convex function. If the inequality is reversed, then it said to be a Schur concave function.*

2. *There is no analog of Proposition 5 when it comes to the probability that i has the most number of wins. That is, if $P_i^*(\mathbf{w})$ is the probability that i wins at least as many games as any other player, then $P_i^*(\mathbf{w})$ is not necessarily either a Schur convex or a Schur concave function. For instance, suppose $r = 4$, $v_i = n = 1$, $\sum_{j=1}^{3} w_j = 6$. Then, $P_i^*(2,2,2) = 1/27 + 3(1/3)^2(2/3)(3/4) = 11/54$, whereas $P_i^*(3,3,0) = 1/16 + 2(1/4)(3/4)(1/2) = 1/4$, and $P_i^*(6,0,0) = 1/7$. (It is also true that the probability that i wins strictly more games than any other player need not be either Schur convex or concave.)*

The following example shows that, for $r > 2$, the probability that $i$ wins the most, or the unique most number of games is not necessarily a decreasing function either of $v_j$ or of the given value of $W_j$, for $j \ne i$.

**Example 2.** *Suppose $r = 3$, and that $v_1 = 9$, $v_2 = 4$. Then, when $n = 1$, it is easy to check that*

$$P(W_1 \ge \max(W_2, W_3)) = \begin{cases} 0.7338, & \text{if } v_3 = 1 \\ 0.7343, & \text{if } v_3 = 2 \end{cases}$$

*When $n = 2$, condition on $W(1,2), W(1,3)$ to obtain*

$$\begin{aligned} &P(W_1 > \max(W_2, W_3)) \\ &\quad = P(W(1,2) = W(1,3) = 2) + P(W(1,2) = 1, W(1,3) = 2)\left(1 - \left(\frac{v_2}{v_2+v_3}\right)^2\right) \\ &\qquad + P(W(1,2) = 2, W(1,3) = 1)\left(1 - \left(\frac{v_3}{v_2+v_3}\right)^2\right) \end{aligned}$$

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

$$= \left(\frac{v_1}{v_1+v_2}\right)^2 \left(\frac{v_1}{v_1+v_3}\right)^2 + \frac{2v_1^3 v_2}{(v_1+v_2)^2(v_1+v_3)^2}\left(1-\left(\frac{v_2}{v_2+v_3}\right)^2\right)$$
$$+ \frac{2v_1^3 v_3}{(v_1+v_2)^2(v_1+v_3)^2}\left(1-\left(\frac{v_3}{v_2+v_3}\right)^2\right)$$

*This gives, when* $v_1 = 9$, $v_2 = 4$, *that*

$$P(W_1 > \max(W_2, W_3)) = \begin{cases} 0.5953, & \text{if } v_3 = 1 \\ 0.6060, & \text{if } v_3 = 2 \end{cases}$$

*Before giving an example illustrating that* $P(W_1 > \max(W_2, W_3) \mid W_3 = k)$ *need not be a decreasing function of* $k$, *note first the interesting result, when all* $n(i, j) = n$, *that conditional on* $W_i = 1$, *the probability that the lone win of* $i$ *was against* $j$ *is proportional to* $1/v_j$. *To verify this, note that*

$$P(W(i,j) = 1, W_i = 1) = n\frac{v_i}{v_i+v_j}\left(\frac{v_j}{v_i+v_j}\right)^{n-1}\prod_{k\neq i,j}\left(\frac{v_k}{v_i+v_k}\right)^n$$
$$= \frac{nv_i}{v_j}\prod_{k\neq i}\left(\frac{v_k}{(v_k+v_i)}\right)^n$$

*which gives that*

$$P(W(i,j) = 1 \mid W_i = 1) = \frac{1/v_j}{\sum_{k\neq i} 1/v_k}$$

*Now, when* $n = 1$, $v_1 = 9$, $v_2 = 4$, $v_3 = 1$

$$P(W_1 \geq \max(W_2, W_3) | W_3 = k) = \begin{cases} 0.6923, & \text{if } k = 0 \\ 0.9053, & \text{if } k = 1 \end{cases}$$

*where the* $W_3 = 1$ *case is computed by conditioning on who* 3 *beat. Since this will be* 2 *with probability* $v_1/(v_1 + v_2)$, *we obtain*

$$P(W_1 \geq \max(W_2, W_3) \mid W_3 = 1) = \frac{v_1}{v_1+v_2} + \frac{v_2}{v_1+v_2}\frac{v_1}{v_1+v_2}$$

*When* $n = 2$, $v_1 = 9$, $v_2 = 4$, $v_3 = 1$

$$P(W_1 > \max(W_2, W_3) | W_3 = k) = \begin{cases} 0.4793, & \text{if } k = 0 \\ 0.7742, & \text{if } k = 1 \end{cases}$$

*where the* $W_3 = 1$ *case is computed by conditioning on who* 3 *beat, giving*

$$P(W_1 > \max(W_2, W_3) \mid W_3 = 1) = \frac{v_2}{v_1+v_2}\left(\frac{v_1}{v_1+v_2}\right)^2 + \frac{v_1}{v_1+v_2}\left(1-\left(\frac{v_2}{v_1+v_2}\right)^2\right)$$

**Remark.** *In the case where* $v_i = v_i$, $v_j = \alpha w_j$, $j \neq i$, *the probability that* $i$ *has the most number of wins is a decreasing function of* $\alpha$. *To prove this, consider two scenarios, the first where* $\alpha = \alpha_1$ *and the second where* $\alpha = \alpha_2$, *where* $\alpha_1 > \alpha_2$. *Because, for* $i \notin \{k, j\}$, *the probability that* $k$ *wins when playing* $j$ *is the same in both scenarios, we can couple the results of all games not involving* $i$ *so that their results are identical in both scenarios. We can also couple all games involving* $i$ *so that if* $i$ *wins a game in scenario two, then it also wins that game in scenario one. In this way, we see that* $i$*'s probability of having the most wins is larger in scenario one than in scenario two.*

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

**Proposition 6.** *Suppose $n(i,j) = n$. If $v_i > v_j$, then for any integer $a \le nr/2$,*

$$P\left(W_i \ge \max_{t \ne i} W_t + a\right) > P\left(W_j \ge \max_{t \ne j} W_t + a\right).$$

*Proof.* Fix $i$ and $j$, and suppose that $v_i > v_j$. Let $A_s = \{W_s \ge \max_{t \ne s} W_t + a\}$, and let $N_{i,j}$ denote the number of events $A_i, A_j$ that occur. Then, for $k = i, j$

$$P(A_k) = P(N_{i,j} = 2) + P(A_k \mid N_{i,j} = 1)P(N_{i,j} = 1)$$

Consequently,

$$P(A_i) > P(A_j) \Leftrightarrow P(A_i \mid N_{i,j} = 1) > P(A_j \mid N_{i,j} = 1)$$

Suppose that $N_{i,j} = 1$, and let $I = i, J = j$ if $A_i$ occurs and let $I = j, J = i$ if $A_j$ occurs. For a given vector $\mathbf{w} = (w, w_0(t), w_1(t), t \ne i, j)$, let $H(\mathbf{w}) = \{W(I,J) = w, W(I,t) = w_0(t), W(J,t) = w_1(t), t \ne i, j\}$. Also, let $w_I = w + \sum_{t \ne i,j} w_0(t)$ and $w_J = n - w + \sum_{t \ne i,j} w_1(t)$.

Suppose, from now, that $\mathbf{w}$ is such that $P(H(\mathbf{w})|N_{i,j} = 1) > 0$, and note that this implies that $w_I > w_J$. With $E_{i,j} = \{W(i,j) = w, W(i,t) = w_0(t), W(j,t) = w_1(t), t \ne i, j\}$, we have, upon letting $C = \binom{n}{w} \prod_{t \ne i,j} \binom{n}{w_0(t)} \binom{n}{w_1(t)}$, that

$$\begin{aligned} P(E_{i,j}) &= C \left(\frac{v_i}{v_i + v_j}\right)^w \left(\frac{v_j}{v_i + v_j}\right)^{n-w} \\ &\quad \times \prod_{t \ne i,j} \left(\frac{v_i}{v_i + v_t}\right)^{w_0(t)} \left(\frac{v_t}{v_i + v_t}\right)^{n - w_0(t)} \left(\frac{v_j}{v_j + v_t}\right)^{w_1(t)} \left(\frac{v_t}{v_j + v_t}\right)^{n - w_1(t)} \\ &= CK(\mathbf{w}) v_i^{w + \sum_{t \ne i,j} w_0(t)} v_j^{n - w + \sum_{t \ne i,j} w_1(t)} \\ &= CK(\mathbf{w}) v_i^{w_I} v_j^{w_J} \end{aligned}$$

where

$$K(\mathbf{w}) = \frac{\prod_{t \ne i,j} v_t^{2n - w_0(t) - w_1(t)}}{(v_i + v_j)^n \prod_{t \ne i,j} (v_t + v_i)^n (v_t + v_j)^n}$$

Because, given $N_{i,j} = 1$, $A_i$ and $A_j$ are mutually exclusive, we have that

$$P(A_i \mid H(\mathbf{w}), N_{i,j} = 1) = \frac{P(E_{i,j})}{P(E_{i,j}) + P(E_{j,i})}$$

Because $v_i > v_j$ and $w_I > w_J$, we see that

$$\frac{P(E_{i,j})}{P(E_{j,i})} = \left(\frac{v_i}{v_j}\right)^{w_I - w_J} > 1$$

Hence, for any $\mathbf{w}$ for which $P(H(\mathbf{w})|N_{i,j} = 1) > 0$, we have that $P(A_i \mid H(\mathbf{w}), N_{i,j} = 1) > P(A_j \mid H(\mathbf{w}), N_{i,j} = 1)$. Taking expectations, conditional on $N_{i,j} = 1$, gives that $P(A_i|N_{i,j} = 1) > P(A_j|N_{i,j} = 1)$, which proves the result. □

**Remarks.** *1. In Proposition 6, the condition on $a$ is included because if $a > nr/2$, then $P(W_i \ge \max_{t \ne i} W_t + a) = 0$ for all $i$.*

*2. One might at first think that Proposition 6 is a simple consequence of Proposition 3(b) by the following argument. Suppose two scenarios, where in scenario one $v_i = v_j = v$ and in scenario two that $v_i = v'$, $v_j = v$, where $v' > v$. In both scenarios, $v_k, k \ne i, j$, are the same. Now, letting the*

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

*superscript indicate the scenario, we have that* $P(W_i^1 \geq \max_{t\neq i} W_t^1 + a) = P(W_j^1 \geq \max_{t\neq j} W_t^1 + a)$ *and, by Proposition 3(b), that* $P(W_i^2 \geq \max_{t\neq i} W_t^2 + a) > P(W_i^1 \geq \max_{t\neq i} W_t^1 + a)$, *enabling us to conclude that* $P(W_i^2 \geq \max_{t\neq i} W_t^2 + a) > P(W_j^1 \geq \max_{t\neq j} W_t^1 + a)$. *However because, as seen in Example 2, it is not necessarily true that* $P(W_j^1 \geq \max_{t\neq j} W_t^1 + a) \geq P(W_j^2 \geq \max_{t\neq j} W_t^2 + a)$, *we cannot conclude (as would be needed to prove Proposition 6) that* $P(W_i^2 \geq \max_{t\neq i} W_t^2 + a) > P(W_j^2 \geq \max_{t\neq j} W_t^2 + a)$.

## 5. Efficient simulation

In this section, we show how to effectively simulate some of the probabilities of interest.

### 5.1. *Simulation estimation of* $P(W_j \geq \max_{i\neq j}(W_i + a_i)|W_j = k)$ *for fixed* $j$ *and all* $k$

For notational convenience, let $j = 1$ and suppose that we are interested in using simulation to estimate, for all $k = 0, \ldots, n_1$,

$$P_k \equiv P\left(W_1 \geq \max_{i=2,\ldots,r}(W_i + a_i) \mid W_1 = k\right)$$

To begin (the $n_1 = \sum_{j\neq 1} n(1, j)$ games of player 1, for example, say that game $i$ is the $i$th game of player 1 against player 2 for $i = 1, \ldots, n(1, 2)$; game $n(1, 2) + i$ is the $i$th game of player 1 against player 3 for $i = 1, \ldots, n(1, 3)$; and so on. Let $X_i$ be the indicator for the event that player 1 wins game $i, i = 1, \ldots, n_1$, and let $p_i = P_{1,o_i}$, where $o_i$ is the opponent of player 1 in game $i$. Now, for $S_i \equiv X_1 + \cdots + X_i$, use the recursion of Section 3 to determine the values

$$P_i(j) = P(S_i = j), \quad j \leq i,\ i = 1, \ldots, n_1.$$

Also, let $C_i$ be the set of indices of games between 1 and $i$, $i = 2, \ldots, r$.

Each simulation run yields an unbiased estimate of all values $P_1, \ldots, P_{n_1}$. A simulation run is as follows:

- Generate the values of the $\binom{r-1}{2}$ independent binomial random variables $W(i, j)$ for $1 < i < j$, and, for $i = 2, \ldots, r$, let $W_i^* = \sum_{j:1<i<j} W(i, j) + \sum_{j:1<j<i}(n(i, j) - W(j, i))$. Thus, $W_i^*$ is the simulated value of the number of games that $i$ wins against all teams aside from team 1.
- Generate $n_1$ independent uniform $(0, 1)$ random variables $U_1, \ldots, U_{n_1}$.
- For each $k, k \leq n_1$, we now show how to generate the values of $X_1^k, \ldots, X_{n_1}^k$, random variables whose distribution is the conditional distribution of $X_1, \ldots, X_{n_1}$ given that $S_{n_1} = k$. Using that

$$P(X_{n_1} = 1 \mid S_{n_1} = k) = \frac{p_{n_1} P_{n_1-1}(k-1)}{P_{n_1}(k)}$$

set

$$X_{n_1}^k = \begin{cases} 1, \text{ if } U_{n_1} < \dfrac{p_{n_1} P_{n_1-1}(k-1)}{P_{n_1}(k)} \\ 0, \text{ if otherwise} \end{cases}$$

Let $x_{n_1}^k$ be the generated value of $X_{n_1}^k$.

Suppose that the values $X_{n_1}^k = x_{n_1}^k, \ldots, X_{j+1}^k = x_{j+1}^k$ have been generated. Let $s_{j+1}^* = x_{n_1}^k + \cdots + x_{j+1}^k$, and use that

$$P(X_j = 1 \mid X_{n_1}^k = x_{n_1}^k, \ldots, X_{j+1}^k = x_{j+1}^k, S_{n_1} = k) = P(X_j = 1 \mid S_j = k - s_{j+1}^*)$$
$$= \frac{p_j P_{j-1}(k - s_{j+1}^* - 1)}{P_j(k - s_{j+1}^*)}$$

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

to generate the value of $X_j^k$, by setting

$$X_j^k = \begin{cases} 1, \text{ if } U_j < \dfrac{p_j P_{j-1}(k - s_{j+1}^* - 1)}{P_j(k - s_{j+1}^*)} \\ 0, \text{ if otherwise} \end{cases}$$

Let $x_j^k$ be the generated value of $X_j^k$. Continue until all values $x_1^k, \ldots, x_{n_1}^k$ have been determined.

- Recall that $C_i$ is the set of indices of the games that 1 plays against $i$, and let $W^k(1,i) = \sum_{j \in C_i} x_j^k$, and $W_i^k = W_i^* + n(i,1) - W^k(1,i),\ i > 1$. Now let

$$I_k = I\left\{k \geq \max_{i>1}(W_i^k + a_i)\right\}$$

It is easy to see that $I_k$ is an unbiased estimate of $P_k$.

- It is not necessary to repeat the preceding for all $k$. Using that $P(X_i = 1 \mid S_i = t)$ is, by Efron's theorem, an increasing function of $t$, it follows from the fact that we are using the same values of $W_i^*, i > 1$, and the same uniforms $U_1, \ldots, U_{n_1}$ to estimate all quantities $P_2, \ldots, P_r$ that $X_j^k$ will, for all $j$, be increasing in $k$ and thus that $I_k$ will be increasing in $k$. Hence, if it results that $I_k = 1$, then we should set $I_j = 1$ for all $j \geq k$, or if it results that $I_k = 0$, then we should set $I_j = 0$ for all $j \leq k$. Doing this greatly reduces the computation necessary to obtain unbiased estimates for all $P_k, k = 0, \ldots, n_1$. Multiple simulation runs should be performed and the estimate of $P_k$ is the average value of $I_k$ obtained.

### *5.2. Simulation estimation of* $P(W_1 \geq \max_{i \neq 1}(W_i + a_i))$

To estimate $P \equiv P(W_1 \geq \max_{i \neq 1}(W_i + a_i))$, we suggest using the same procedure as in the previous subsection. That is, first do the preliminary work of computing the values $P_i(j),\ j \leq i,\ i = 1, \ldots, n_1$. Then, in each simulation run generate the values $I_k, 0 \leq k \leq n_1$. Letting $k^* = \min\{k : I_k = 1\}$, and using that

$$P = \sum_{k=0}^{n_1} P\left(W_1 \geq \max_{i \neq 1}(W_i + a_i) \mid W_1 = k\right) P_{n_1}(k)$$

it follows that $P(W_1 \geq k^*) = \sum_{k=k^*}^{n_1} P_{n_1}(k)$ is an unbiased estimate of $P$. Averaging this value over many simulation runs gives the estimate.

### *5.3. Simulation estimation of* $P(W_i \geq \max_{t \neq i}(W_t + a_t))$ ***for all*** $i$

In this case, we suggest generating the $\binom{r}{2}$ independent binomials $W(i,j), i < j$, and setting $W_i = \sum_{i<j} W(i,j) + \sum_{j<i}(n(i,j) - W(j,i)),\ i = 1, \ldots, r$. Then use the estimator $I_i = I\{W_i \geq \max_{t \neq i}(W_t + a_t))\}$ along with $W_i$ as a control variable. That is, the estimate of $P(W_i \geq \max_{t \neq i}(W_t + a_t))$ from a given run is $I_i + c_i(W_i - \sum_{j \neq i} n(i,j)P_{i,j})$, where the appropriate value of $c_i$ is determined from the simulation data (see [15]).

**Funding statement.** This material is based on work supported by, or in part, by the National Science Foundation under contract/grant CMMI1662442.

**Conflict of interest.** The authors declare no conflict of interest.

## References

[1] Adler, I., Cao, Y., Karp, R.M., Pekoz, E., & Ross, S.M. (2017). Random knockout tournaments. *Operations Research* 65(6): 1589–1596.

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press

[2] Cao, Y. & Ross, S.M. (2019). Winner plays structure in random knockout tournaments. *Probability in the Engineering and Informational Sciences* 33(4): 500–510.
[3] Chung, F.R.K. & Hwang, F.K. (1978). Do stronger players win more knockout tournaments? *Journal of the American Statistical Association* 73(363): 593–596.
[4] David, H.A. (1959). Tournaments and paired comparisons. *Biometrika* 46(1/2): 139–149.
[5] Edwards, T.E. (1996). Double-elimination tournaments: counting and calculating. *The American Statistician* 50(1): 27–33.
[6] Glenn, W.A. (1960). A comparison of the effectiveness of tournaments. *Biometrika* 47(3/4): 253–262.
[7] Hennessy, J. & Glickman, M. (2016). Bayesian optimal design of fixed knockout tournament brackets. *Journal of Quantitative Analysis in Sports* 12(1): 1–15.
[8] Horen, J. & Riezman, R. (1985). Comparing draws for single elimination tournaments. *Operations Research* 33(2): 249–262.
[9] Hwang, F.K. (1982). New concepts in seeding knockout tournaments. *The American Mathematical Monthly* 89(4): 235–239.
[10] Marchand, E. (2002). On the comparison between standard and random knockout tournaments. *Journal of the Royal Statistical Society: Series D (The Statistician)* 51(2): 169–178.
[11] Marshall, A. & Olkin, I. (1979). *Inequalities: theory of majorization and its applications*. New York, USA: Academic Press.
[12] Maurer, W. (1975). On most effective tournament plans with fewer games than competitors. *The Annals of Statistics* 3(3): 717–727.
[13] Moser, L. & Harary, F. (1966). The theory of round robin tournaments. *The American Mathematical Monthly* 73(3): 231–246.
[14] Pekoz, E. & Ross, S.M. (2007). *A second course in probability*. Berkeley, USA: Probability/Bookstore.com.
[15] Ross, S.M. (2013). *Simulation*, 5th ed. San Diego, USA: Academic Press.
[16] Ross, S.M. (2016). Improved Chen-Stein bounds on the probability of a union. *Journal of Applied Probability* 53(4): 1265–1270.
[17] Saarinen, S., Tovey, C.A., & Goldsmith, J. (2014). A model for intransitive preferences. Presented at *Meeting of Association for Advancement of Artificial Intelligence*.
[18] Searls, D.T. (1963). On the probability of winning with different tournament procedures. *Journal of the American Statistical Association* 58(304): 1064–1081.

**Cite this article:** Ross SM (2022). Team's seasonal win probabilities. *Probability in the Engineering and Informational Sciences* **36**, 988–998. https://doi.org/10.1017/S026996482100019X

https://doi.org/10.1017/S026996482100019X Published online by Cambridge University Press