*Journal of Probability
and Statistical Science
24(1), 1-15 Feb. 2026*

Winning Probability—in a Game, in a Series, and Afterwards

Jyotirmoy Sarkar
Department of Mathematical Sciences
Indiana University, Indianapolis, Indiana, USA

ABSTRACT

Two teams, A and B, are facing off in a best-of-seven series. What is the probability P(A) that Team A will win a game? What is the probability $P_7(A)$ that Team A will win the series? Different people will answer the questions differently based on many subjective criteria. Random variables with support [0, 1] are used to model the answers. How are the distributions of P(A) and $P_7(A)$ related? After the number of games played and the series winner are known, *a posteriori*, what is the revised distribution of P(A)?

Keywords: Teaching statistics; prior distribution; posterior distribution; Bayes' theorem.

PREAMBLE

In 2025, the National Basketball Association (NBA) Playoff Final best-of-seven series of games was played between the Western Conference champion Oklahoma City Thunder and the Eastern Conference champion Indiana Pacers. Everyone was asking: Which team will win each game? Win the series? How many games will be played to determine the winner? See Figure 1.

Figure 1: 2025 NBA Playoff Final tied at 3-3 as of June 19. Which team will win the series?
[Courtesy: nba.com/news/2025-nba-playoffs-schedule]

1. Introduction

Two teams, A and B, are facing off in a best-of-seven (or any odd) game series. The answer to which team will win in any one game depends on whom you ask. For diehard fans, the answer is obvious. On the other hand, if you collected the opinions of third-party observers or even game experts, you would realize that no one knows the answer for sure. Rather, they would guess the

- Received August 18, 2025, in final form February 2026.
- Jyotirmoy Sarkar (corresponding author) is affiliated with the Department of Mathematical Sciences, Indiana University, Indianapolis, Indiana, USA
jsarkar@iu.edu

probability P(A) that Team A will win in any game, and the probability 1-P(A) that Team B will win, as best they can. They would likely agree that their guess for P(A) is not set in stone. They are willing to allow P(A) to be a random variable taking values around their best guess. By collecting the opinions of many dispassionate experts, you may formulate a (subjective) prior distribution of P(A) over [0, 1].

At the start of the NBA Playoffs Final, the Thunder were the favorite (84%), and the Pacers were a long shot (16%). But the Pacers won games 1, 3, and 6, necessitating Game 7! The experts promptly changed the ratings to Thunder 68%, Pacers 32% for Game 7.

According to the frequentist interpretation, P(A) is the long-run proportion of times event A occurs in an infinite sequence of experiments. As such, in sports, P(A) cannot be computed empirically since the teams do not play an endless sequence of games. Even if they did, each team's performance is likely to change over time. This is another reason for treating P(A) as a random variable that captures the degree of belief people assign to its values. Likewise, when the series is over, the updated P(A) cannot be verified empirically since the teams may not play each other, or if they do, their players/skills may change. Hence, the updated P(A) after the series is over is also a random variable with a posterior distribution.

This paper addresses how to compute (1) the probability $P_{2m+1}(A)$ of winning the best-of-(2m+1) game series as a function of the probability P(A) of winning each game, (2) the distribution of $P_{2m+1}(A)$ based on the prior distribution of P(A), and (3) the posterior distribution of P(A) after the series is over and we know how many games were played and who won the series.

All computations and illustrations are made with the freely available software R.

2. If P(A) Were a Known Constant

Assume that in any game, the probability that A wins is P(A)=p, a *known constant*. To approximate $P_7(A)$, we can simulate the series of 7 games a large number of times, say 10^4 times, so that the approximation is correct to two decimal places. See R codes in the Appendix.

When p=.6, we estimate $P_7(A)=.7076$. Moreover, we estimate the probability mass function for the number of games N needed to determine the winner: $P_7(N=4)=.153$, $P_7(N=5)=.268$, $P_7(N=6)=.299$, $P_7(N=7)=.280$. Furthermore, given the number of games, the conditional probability that A is the winner is: $P_7(A|N=4)= 0.83$, $P_7(A|N=5)=.76$, $P_7(A|N=6)=.70$, $P_7(A|N=7)=.60$. Note that the weighted average of these conditional probabilities is

$$.153\times.83 + .268\times.76 + .299\times.70 + .280\times.60 = .7080 \sim P_7(A).$$

A simulation study is a good guide to estimate results. Nonetheless, we must derive the answers using probability theory. To develop the solution, we begin with simpler problems.

Example 1: In the Women's Australian Open Final, Player A is favored over Player B 55:45. What is the probability that Player A will win the best-of-three-sets?

Assume that the successive sets are independent. (In reality, this assumption may be somewhat questionable because of physical, psychological, and environmental conditions.) The joint probability of independent events is the product of the individual probabilities. See, for example, Wackerly *et al*. The game can end in 2 sets, with A winning with probability (wp) $.55^2$ and B winning wp $.45^2$. So, the game can end in N=2 sets wp $P_3(N=2)=.55^2+.45^2=.505$. If the game ends in two sets, then Player A wins the series wp $P_3(A|N=2)=.55^2/.505=.599$.

The game moves to set 3, only if in the first two sets, each player wins one set, which happens wp $P_3(N=3)=2\times.55\times.45 = 0.495$. Player A wins the deciding third set wp $P_3(A|N=3)=P(A)=.55$. Therefore, conditioning on the number of sets, Player A wins the series wp $P_3(A)=(.55^2/.505)\times.505+(.55)\times.495=.57475$. Likewise, Player B wins the series wp

$P_3(B)=(.45^2/.505)\times.505+(.45)\times.495=.42525=1 - .57475$. Thus, compared to a single-set game, a best-of-three-sets game gives the better player an added advantage.

Example 2: In the Men's Wimbledon Semi-Final, Player A is favored over Player B 52:48. What is the probability that a best-of-five-set will end in (a) 3 matches? (b) 4 matches, (c) 5 matches? What is the probability that A will win the series?

The number N of matches required to determine the series winner is a discrete random variable. To describe it completely, we must list its values together with the associated probabilities that add to 1. The game can end in 3 matches with A winning wp $.52^3$ and with B winning wp $.48^3$. So, the game can end in N=3 matches wp $P_5(N=3)=.52^3+.48^3=.2512$. The game can end in 4 matches if the winners' sequence is one of the following six sequences of outcomes: AABA, ABAA, BAAA, BBAB, BABB, ABBB. So, the game ends in N=4 matches wp $P_5(N=4)=3\times.52^3\times.48 + 3\times.52\times.48^3 = .374999$. Finally, the game ends in N=5 matches if in the first four matches, each player wins two matches, which happens wp ${}^4C_2\times.52^2\times.48^2 = .373801$. Here ${}^4C_2=4!/[(4-2)!*2!]$ is the binomial coefficient counting the number of ways 2 items can be chosen from among 4 indistinguishable items. Check that

$$P_5(N=3) + P_5(N=4) + P_5(N=5) = .251200 + .374999 + .373801 = 1.$$

More generally, in a best-of-five matches, if Player A is favored over Player B p:q in each game, where q=1-p, the game ends in k=3, 4, 5 matches, wp $P_5(N=3)=(p^3+q^3)$, $P_5(N=4)=3(p^3q+pq^3)$ and $P_5(N=5)=6p^2q^2$, respectively. Check that the sum is

$$\begin{aligned}&(p^3+q^3) + 3pq\,(p^2+q^2) + 6p^2q^2\\&= (p^3+q^3) + 3pq\,(1-2pq) + 6p^2q^2\\&= (p^3+q^3) + 3pq = (p+q)^3 = 1.\end{aligned}$$

What is the probability that Player A will be the series winner? First, we record the conditional probability of A winning, given the number of games needed to determine the series winner. Then we use the law of total probability (or the weighted average of conditional probabilities) to compute the overall probability of A winning the series. See Wiki1.

If the series ends in k=3 matches, A wins wp $P_5(A|N=3)= p^3/(p^3+q^3)$. If k=4 matches, then A wins wp $P_5(A|N=4)= p^2/(p^2+q^2)$. If k=5 matches, then A wins wp $P_5(A|N=5)=p$. Overall, A wins the best-of-five-sets wp

$$P_5(A) = p^3+3p^3q+6p^3q^2 = p^3\,[1+3q+6q^2].$$

If p=.52, then $P_5(A)=.52^3\times3.8224=.53746$.

What if the two players were equally favored? Then $P_5(N=3)=1/4$, $P_5(N=4)=3/8$, $P_5(N=5)=3/8$; $P_5(A|N=3)=.5$, $P_5(A|N=4)=.5$, $P_5(A|N=5)=.5$; and $P_5(A)=.5$. On the other hand, if Player A is favored over Player B 6:4, then $P_5(N=3)=.28$, $P_5(N=4)=.3744$, $P_5(N=5)=.3456$; $P_5(A|N=3)=.7714$, $P_5(A|N=4)=.6923$, $P_5(A|N=5)=.6$; and $P_5(A)=.68256$.

The more alike the players are (in terms of winning probability), the **stochastically larger** N is (as larger values of N have higher probabilities). To reiterate, the more disparate the players are, the **stochastically smaller** N is (as smaller values of N have higher probabilities). That is, the series winner will be decided stochastically sooner. See Wiki2.

Example 3: Returning to the NBA Playoff Final, suppose that in a best-of-seven-games, Player A is favored over Player B p:q in each game, then the series ends in k=4, 5, 6, 7 games, wp $P_7(N=4)=(p^4+q^4)$, $P_7(N=5)=4(p^4q+pq^4)$, $P_7(N=6)=10(p^4q^2+p^2q^4)$, and $P_7(N=7)=20p^3q^3$, respectively. Check that (substituting q=1-p) the sum is

$$\begin{aligned}&(p^4+q^4) + 4(p^4q+pq^4) + 10(p^4q^2+p^2q^4) + 20p^3q^3\\&= p^3(p+4pq+10q^2) + q^3(q+4pq+10p^2)\end{aligned}$$

$= p^3(10\text{-}15p+6p^2)+ (1\text{-}p)^3(1+3p+6p^2) = 1$

What's the probability that Player A will win the series? If the series ends in k=4 games, A wins wp $P_7(A|N=4)=p^4/(p^4+q^4)$. If k=5 games, then A wins wp $P_7(A|N=5)= p^3/(p^3+q^3)$. If k=6 games, then A wins wp $P_7(A|N=6)= p^2/(p^2+q^2)$. If k=7 games, then A wins wp $P_7(A|N=7)=p$. Overall, A wins the best-of-seven-games series wp (given by the weighted average)

$$P_7(A) = p^4+4\ p^4q+10\ p^4q^2+20\ p^4q^3 = p^4\ [1+4q+10q^2+20q^3].$$

In fact, if we imagine that the series does not end as soon as a team wins 4 games, but it continues till all 7 games are played out (even though the series winner is already determined), then Team A wins the series if and only if it wins **four or more** games out of 7. Hence,

$$P_7(A) = \text{sum}^7_{i=4}\ {}^7C_i\ p^i\ q^{7-i}.$$

Remark 1: We leave to the reader to prove the equivalence of the two expressions for $P_7(A)$. Hint: Multiply the four terms inside [ ] by $(p+q)^3$, $(p+q)^2$, $(p+q)$, 1 respectively, and expand.

If Player A is favored over Player B 84:16, as the Thunder were against the Pacers, then $P_7(N=4)=.4985$, $P_7(N=5)=.3208$, $P_7(N=6)=.1321$, $P_7(N=7)=.0486$; $P_7(A|N=4)=.9987$, $P_7(A|N=5)=.9931$, $P_7(A|N=6)=.9650$, $P_7(A|N=7)=.8400$; and $P_7(A)=.98475$.

Remark 2: If Player A is favored over Player B 6:4, then $P_3(A)=.6480$, $P_5(A)=.6826$, and $P_7(A)=.7102$. Thus, the more sets in the series, the greater the advantage to the favorite team and the tougher the challenge to the underdog to win the series.

Remark 3: If experts give you their best guesses for $p=P(A)$ and $P_7(A)$ that do not satisfy $P_7(A) = \text{sum}^7_{i=4}\ {}^7C_i\ p^i\ q^{7-i}$, at least approximately, then rest assured they are not experts in probability.

Figure 2 shows the probability of winning $P_{2m+1}(A)$ as a function of $p=P(A)$ in a best-of-3, 5, 7 games series. Note that $P_{2m+1}(A)$ is an increasing function of p, and it increases with m>=1 if and only if p>1/2. We shall prove these results in Theorem 1.

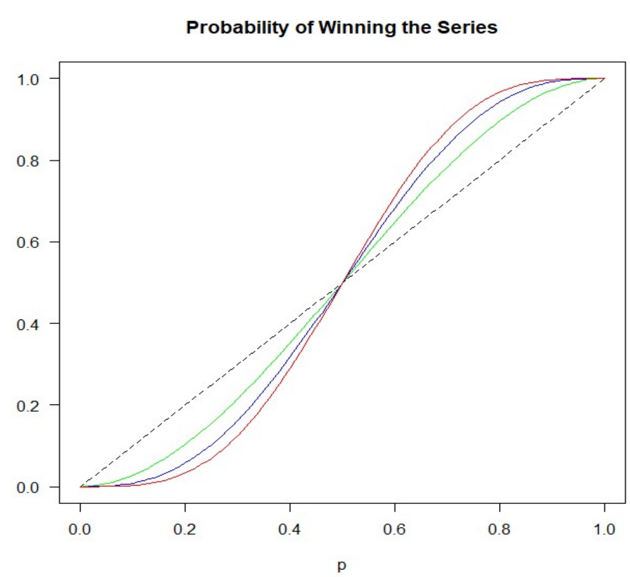

Figure 2: The probability $P_{2m+1}(A)$ of winning in a best-of-(2m+1) games series for selected values of m: m=1 (green), m=2 (blue), and m=3 (red).

Having studied the best-of-3, 5, or 7 games, we can generalize further. In a best-of-(2m+1)-games series, the number of games N has a PMF

$$P_{2m+1}(N=k) = {}^{k-1}C_m\, p^{m+1}\, q^{k-m-1}, \text{ for } k=m+1, m+2, \ldots, 2m+1.$$

The conditional probability of Player A winning the series, given N=k, is

$$P_{2m+1}(A|N=k) = p^{2m+2-k} / (p^{2m+2-k} + q^{2m+2-k})$$

The overall probability of Player A winning the series is

$$P_{2m+1}(A) = p^{m+1}\, [1 + {}^{m+1}C_1\, q + {}^{m+2}C_2\, q^2 + {}^{m+3}C_3\, q^3 + \ldots + {}^{2m}C_m\, q^m]$$
$$= \text{sum}^{2m+1}_{i=m+1}\ {}^{2m+1}C_i\, p^i\, q^{2m+1-i}.$$

Theorem 1: (1) $P_{2m+1}(A)$ increases in p, and (2) $P_{2m+1}(A)$ increases in m>=1 for p>1/2 and decreases in m>=1 for p<1/2. When p=1/2, $P_{2m+1}(A)$=1/2 for all m>=1.

Proof: (1) It suffices to show that the derivative of $P_{2m+1}(A)$ with respect to p is positive. Indeed,

$$p^{-m}\, (d/dp)\, P_{2m+1}(A)$$
$$= (m+1)\, [1 + {}^{m+1}C_1\, q + {}^{m+2}C_2\, q^2 + {}^{m+3}C_3\, q^3 + \ldots + {}^{2m-1}C_{m-1}\, q^{m-1} + {}^{2m}C_m\, q^m]$$
$$- (1-q)\, [\ {}^{m+1}C_1 + 2\ {}^{m+2}C_2\, q + 3\ {}^{m+3}C_3\, q^2 + \ldots + (m-1)\ {}^{2m-1}C_{m-1}\, q^{m-2} + m\ {}^{2m}C_m\, q^{m-1}]$$
$$= [(m+1) + (m+2)\ {}^{m+1}C_1\, q + (m+3)\ {}^{m+2}C_2\, q^2 + \ldots + 2m\ {}^{2m-1}C_{m-1}\, q^{m-1} + (2m+1)\ {}^{2m}C_m\, q^m]$$
$$- [{}^{m+1}C_1 + 2\ {}^{m+2}C_2\, q + 3\ {}^{m+3}C_3\, q^2 + \ldots + m\ {}^{2m}C_m\, q^{m-1}]$$
$$= (2m+1)\ {}^{2m}C_m\, q^m > 0.$$

(2) We shall simplify the expression for $P_{2m+1}(A) - P_{2m-1}(A)$ for any m>=1. Indeed,

$$P_{2m+1}(A) - P_{2m-1}(A)$$
$$= p^{m+1}\, [1 + {}^{m+1}C_1\, q + {}^{m+2}C_2\, q^2 + {}^{m+3}C_3\, q^3 + \ldots + {}^{2m-1}C_{m-1}\, q^{m-1} + {}^{2m}C_m\, q^m]$$
$$- p^m\, [1 + {}^{m}C_1\, q + {}^{m+1}C_2\, q^2 + {}^{m+2}C_3\, q^3 + \ldots + {}^{2m-3}C_{m-2}\, q^{m-2} + {}^{2m-2}C_{m-1}\, q^{m-1}]$$
$$= p^m \{ (1-q)\, [1 + {}^{m+1}C_1\, q + {}^{m+2}C_2\, q^2 + {}^{m+3}C_3\, q^3 + \ldots + {}^{2m-1}C_{m-1}\, q^{m-1} + {}^{2m}C_m\, q^m]$$
$$- [1 + {}^{m}C_1\, q + {}^{m+1}C_2\, q^2 + {}^{m+2}C_3\, q^3 + \ldots + {}^{2m-3}C_{m-2}\, q^{m-2} + {}^{2m-2}C_{m-1}\, q^{m-1}] \}$$
$$= p^m \{\ 1 + {}^{m+1}C_1\, q + {}^{m+2}C_2\, q^2 + {}^{m+3}C_3\, q^3 + \ldots + {}^{2m-1}C_{m-1}\, q^{m-1} + {}^{2m}C_m\, q^m$$
$$- q - {}^{m+1}C_1\, q^2 - {}^{m+2}C_2\, q^3 - \ldots - {}^{2m-2}C_{m-2}\, q^{m-1} - {}^{2m-1}C_{m-1}\, q^m - {}^{2m}C_m\, q^{m+1}$$
$$- 1 - {}^{m}C_1\, q - {}^{m+1}C_2\, q^2 - {}^{m+2}C_3\, q^3 - \ldots - {}^{2m-2}C_{m-1}\, q^{m-1}\ \}$$
$$= p^m\, q^m\, ({}^{2m-1}C_m - {}^{2m}C_m\, q) = p^m\, q^m\ {}^{2m-1}C_m\, (1-2q) = (p-q)\ {}^{2m-1}C_m\, (pq)^m$$

which is positive for p>1/2, zero for p=1/2, and negative for p<1/2. This completes the proof.

Remark 4: The proof of Theorem1, Part (2) gives us yet another expression for $P_{2m+1}(A)$, namely $P_{2m+1}(A) = p + (p - q)\ \text{sum}^m_{i=1}\ {}^{2i-1}C_i\, (pq)^i$, where the second term expresses the advantage to the favorite contestant when p>1/2. The advantage increases with m. Moreover, if p=1/2, then there is no advantage.

Remark 5: In gambling, each play is slightly tilted in favor of the gambling house. That is, p = P(gambling house wins) > 1/2. If a gambler plays a large number of times 2m+1, betting one dollar to win one dollar in each play, then the house eventually wins with probability $P_{2m+1}(A)$, which converges to 1 as m tends to infinity. The addictive gambler will surely be ruined. See Wiki3.

3. If P(A) is Random with a Prior Distribution

Suppose P(A)=p is not a fixed number but a random variable. Let $\pi_0(p)$ denote the prior probability density function (PDF) of P(A) with support [0, 1]. How can we calculate Team A's probability of winning the series consisting of (2m+1) games?

The series-winning probability $Q(A)=P_{2m+1}(A)=g(p)$, say, is no longer a single number. As a function of a random variable p, Q(A) is also a random variable. Its PDF is given by

$$s(q) = \pi_0(g^{-1}(q)) / (dq/dp).$$

See, for example, Wackerly *et al*.

To compute the PDF s(q) of Q(A), we proceed as follows: Whereas it suffices to evaluate s(q) at grid points G=seq(0, 1, .01)=0.00, 0.01, 0.02, 0.03, …, 1.00, we begin by evaluating $\pi_0(p)$ at a much finer set of grid points G0=seq(0, 1, .0001). Then we evaluate the CDF of P(A) at points in G0 by taking the cumulative sum of $\pi_0(p)$. Next, we evaluate q=g(p) on G0. The CDF of Q(A) at q=g(p) equals the CDF of P(A) at p. However, we restrict the CDF of Q(A) on G only. Finally, its successive differences form the PDF of Q(A) on G. See the R codes in the Appendix. Figure 3 shows the PDF of the game-winning probability P(A)=p (when m=0) and the series-winning probability $Q(A)=P_{2m+1}(A)$ in the best-of-3, 5, or 7 games (when m=1, 2, 3).

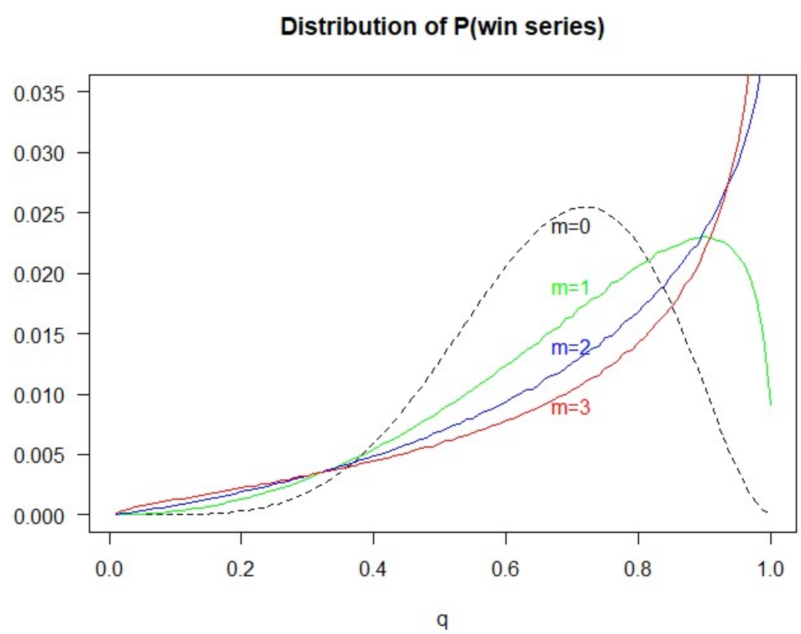

Figure 3: The PDF s(q) of $Q(A)=P_{2m+1}(A)=g(p)$, obtained from the prior PDF $\pi_0(p) \propto p^5q^2$ (corresponding to m=0)

4. Posterior Distribution of P(A)

After the best-of-seven series is played and we know the number of games played and the series winner, how should we determine the posterior distribution $\pi_1(p)$?

Given p, the joint probability f(n, d|p) of {N=n, I_A=d}, where d=1 if A wins and d=0 if B wins, is given in the previous section. Here, we treat that joint probability as a function of p and call it the likelihood function. Then, according to Bayes' Theorem, the posterior density is proportional to the product of the likelihood function and the prior density; that is,

$$\pi_1(p|N=n, I\_A=d) \propto f(n, d|p)\, \pi_0(p).$$

See Berger (1985).

Accordingly, the posterior density is obtained by renormalizing the product (by dividing the product by the integral of the product). For computation, we discretize the prior PDF, multiply by the likelihood, and divide each product by the total product to find the discretized posterior PMF using Bayes' Theorem. Then we continuously approximate it to obtain the posterior PDF. See the R codes in the Appendix.

The posterior PDFs, given the 8 possible values of (N, I_A), are shown in Figure 4. Note that when N is small, the posterior PDF shifts stochastically to the right of the prior PDF if A wins, but to the left if A loses. However, when N is large, the posterior PDF may actually shift stochastically to the left of the prior PDF if A wins, but it shifts further to the left if A loses. The shifts are smaller in magnitude as N becomes larger.

Remark 6: In particular, if the prior PDF of p is proportional to p^5q^2, then the prior is said to have a Beta(6, 3) distribution with mean 2/3, mode 5/7, and SD sqrt(1/45) = 0.149. Then, after the best-of-7 series is over, given (N=n, A), the posterior PDF of p is Beta(10, n-1) distribution with mean 10/(n+9), mode 9/(n+7), and SD sqrt(10(n-1)/(n+10))/(n+9). However, given (N=n, A^c), the posterior PDF of p is Beta(n+2, 7) distribution with mean (n+2)/(n+9), mode (n+1)/(n+7), and SD sqrt(7(n+2)/(n+10))/(n+9). This is why the beta family is said to be a conjugate family of distributions for Bernoulli outcomes.

5. Conclusion

To make the contest more exciting for spectators and generate more revenue, the organizers often split the contest into multiple games. Consequently, the favorite team gains an advantage while the underdog faces additional challenges to win the series. The discrepancy grows bigger as more games are added to the series. To justify these claims, we have proved that $P_{2m+1}(A)$ is an increasing function of m if and only if p>1/2. Therefore, when the underdog snatches a series victory, all the more glory ought to be theirs, and they deserve celebration with added fanfare.

Since usually there is no consensus on the value of p=P(A), it is better to model it as a random variable with a prior distribution. We have demonstrated how to compute the distribution of the probability of winning the series and the posterior distribution of p, given the number of games played and the series winner using Bayes' Theorem. Likewise, we can compute the posterior distribution of p, given any partial result of the series of games before its conclusion. This is what the experts did after Game 6 of the NBA Playoff Finals (by lowering the Thunder's winning probability from .84 to .68).

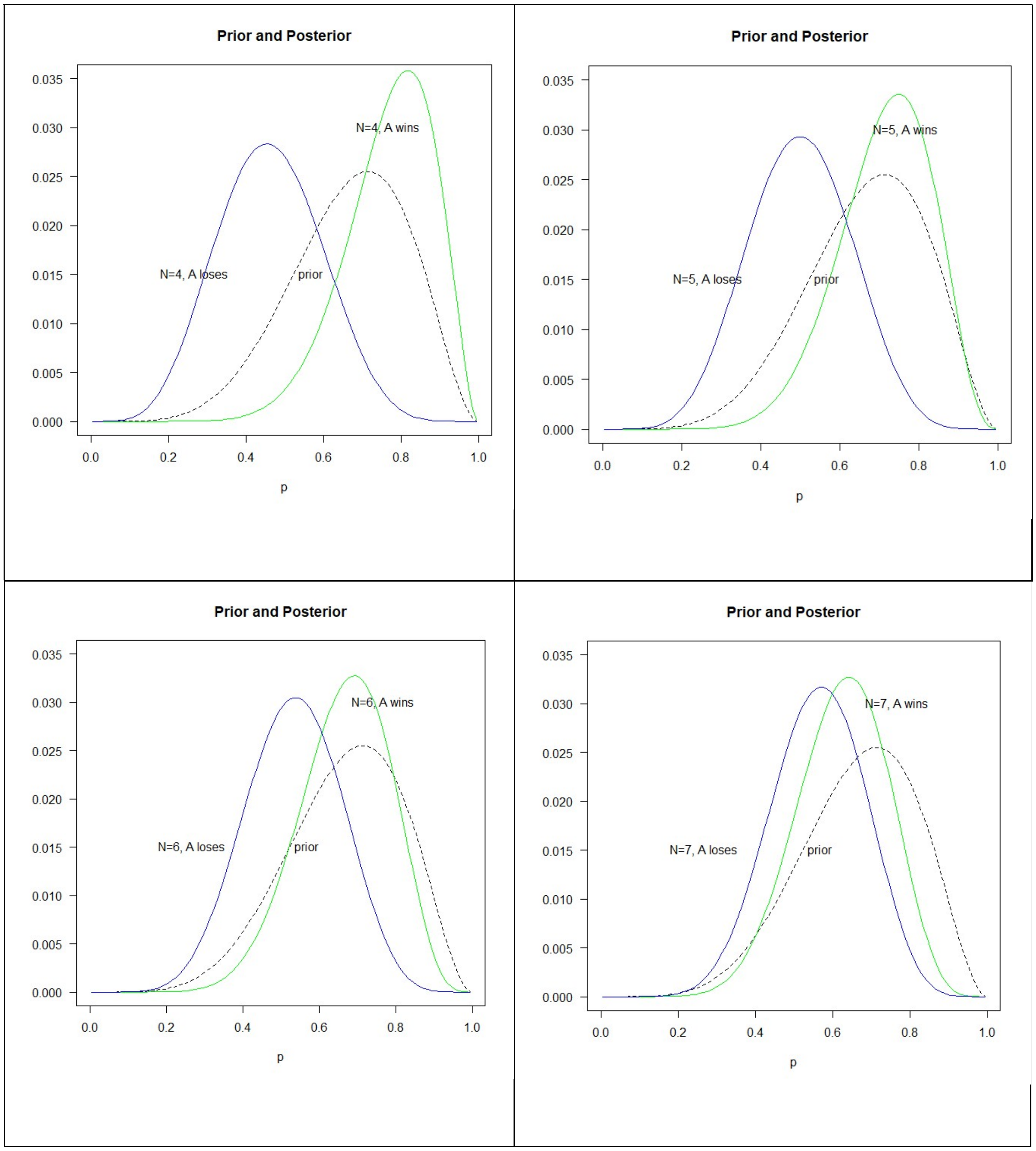

Figure 4: Given (N, I_A), the posterior PDF of p, when the prior PDF is proportional to p^5q^2

ACKNOWLEDGEMENT

The author thanks an old friend, Amar Mukherjee, whose keen interest in the problem was instrumental in writing this paper. Thanks are also due to Yu Chen for critically reading the draft.

References

[1]. R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[2]. Wackerly, D., Mendenhall, W., and Scheaffer, R. (2014). *Mathematical Statistics with Applications* (8th Ed), Brooks/Cole Cengage Learning, CA, USA.

[3]. Wiki1. Wikipedia—The Free Encyclopedia. *Law of total probability.* https://en.wikipedia.org/wiki/Law_of_total_probability

[4]. Wiki2. Wikipedia—The Free Encyclopedia. *Stochastic ordering.* https://en.wikipedia.org/wiki/Stochastic_ordering

[5]. Wiki3. Wikipedia— The Free Encyclopedia. *Gambler's ruin.* https://en.wikipedia.org/wiki/Gambler's_ruin

[6]. Berger, James, O. (1985). *Statistical Decision Theory and Bayesian Analysis* (Springer Series in Statistics) 2nd Edition, ISBN: 978-0387960982.

Appendix

Here are the R codes to compute $P_{2m+1}(A)$, its distribution, and the posterior distribution of P(A).

```
### best-of-seven games
## simulate the series
## How many games? Who is the winner?
play=function(p){x=rbinom(7,1,p)
 cx=cumsum(x); cy=cumsum(1-x)
  A=min(which((cx==4)),8)
  B=min(which((cy==4)),8)
 nb=min(A,B); win=(nb==A)
c(nb, win)
}
## Repeat the series; obtain joint & marginal distributions
R=10^4; p=.6; q=1-p
data=replicate(R, play(p))
t=table(data[1,],data[2,])
t # joint distribution

# simulated conditional prob of win, given N
t[,2]/(t[,1]+t[,2]) # conditional distribution
# theoretical conditional
c(p^4/(p^4+q^4), p^3/(p^3+q^3), p^2/(p^2+q^2), p)

# simulated PMF of N
t(t%*%c(1,1)/R)
# theoretical PMF of N
c((p^4+q^4), 4*p*q*(p^3+q^3), 10*p^2*q^2*(p^2+q^2), 20*p^3*q^3)

# simulated P7(A)
sum(t[,2])/R
# theoretical P7(A)
p^4*(1+4*q+10*q^2+20*q^3)
p^4*(p^3+7*p^2*q+21*p*q^2+35*q^3)
```

```
### Thunder vs Pacers best-of-seven games
p=.84; q=1-p
c((p^4+q^4), 4*(p^4*q+p*q^4), 10*(p^4*q^2+p^2*q^4), 20*p^3*q^3)
c(p^4/(p^4+q^4), p^3/(p^3+q^3), p^2/(p^2+q^2), p)
p^4*(1+4*q+10*q^2+20*q^3)

#### Probability of winning the series as a function of p=P(A)
p=seq(.0000, .9999, .0001)+.00005; q=1-p
pi=p^5*q^2; pi0=pi/sum(pi)
sum(pi0) # make sure this is 1
Pi0=cumsum(pi0) # CDF of p=P(A)
p3=p^2*(1+2*q);
p5=p^3*(1+3*q+6*q^2);
p7=p^4*(1+4*q+10*q^2+20*q^3);

### PDF of P(win the series) from PDF of P(win a game)
q=seq(.01,1.00,.01)
index3=rep(0,100); index5=index3; index7=index3
for (i in 1:100){
index3[i]=sum((p3<i/100))
index5[i]=sum((p5<i/100))
index7[i]=sum((p7<i/100))
}
qq=Pi0[seq(1,100,1)*100]; q0=qq-c(0,qq[-100]) # PDF of p=P(A), m=0
Q3=Pi0[index3];  q3=Q3-c(0,Q3[-100]) # PDF of p=P(A), m=1
Q5=Pi0[index5];  q5=Q5-c(0,Q5[-100]) # PDF of p=P(A), m=2
Q7=Pi0[index7];  q7=Q7-c(0,Q7[-100]) # PDF of p=P(A), m=3

plot(q, q0, type='l', lty=2, las=1, ylim=c(0, .040), ylab="",
   main="Distribution of P(win series)")
text(.7, .024, "m=0", col="black")
lines(q,q3, col="green");  text(.7, .019, "m=1", col="green")
lines(q,q5, col="blue");   text(.7, .014, "m=2", col="blue")
lines(q,q7, col="red");    text(.7, .009, "m=3", col="red")

### prior and posterior distributions on p=P(A)
p=seq(.00, .99, .01)+.005; q=1-p
pi=p^5*q^2; pi0=pi/sum(pi)
#pi0=c(0,0,0,1,2,3,6,11,12,15,15,12,11,6,3,2,1,0,0,0)/100
sum(pi0) # make sure this is 1
(m0=sum(p*pi0)) # prior mean
sqrt(sum(p^2*pi0)-m0^2) # prior sd

pip=(p^4*(1-p)^0)*pi0; pip=pip/sum(pip) # change power of (1-p)
pim=(p^0*(1-p)^4)*pi0; pim=pim/sum(pim)  # change power of (1-p)

plot(p, pi0, type='l', lty=2, las=1, ylim=c(0, .035), ylab="",
   main="Prior and Posterior")
lines(p, pip, col="green")
lines(p, pim, col="blue")
text(m0-.1, .015, "prior")
text(m0+.1, .030, "N=4, A wins")  # change value of N=4,5,6,7
text(m0-.4, .015, "N=4, A loses") # change value of N=4,5,6,7
```