STAT 165 Final Project:

When will the NBA 8th seed win the championship?

By Leo Villani and Jonathan Pei

Introduction

In the NBA, playoff seeding plays a major role in a team's likelihood of winning the championship. Since the league expanded to a 16-team playoff format in 1984, no 8th seed has ever won the NBA Finals. Only once has an 8th seed even reached the Finals (the 1999 New York Knicks, during a lockout-shortened season), and only twice has an 8th seed defeated a 1st seed in a 7-game series since 2007.

Yet the possibility remains: as parity increases and league structure evolves (e.g. the introduction of the play-in tournament, load management strategies, and increased three-point variance), it seems plausible that a low-seeded underdog could eventually run the table.

We are interested in when this will happen.

Forecasting Question

By what year will an NBA 8th seed first win a championship?

Relevant Resolution Criteria

- The 8th seed refers to a team that enters the playoffs as the 8th seed in either conference at the start of the first round.
- We will include teams that qualify through the play-in tournament if they enter the 8-team bracket as the 8th seed.
- An 8th seed that wins the Finals by any method (in a full or shortened season) resolves the forecast.
- If not resolved by December 31, 2500, the question resolves as "never by 2500."
- League restructuring that significantly alters playoff seeding or format may lead us to revise the resolution criteria or halt the forecast depending on the severity of the change.

Background Research

Playoffs have heavily favored higher seeds especially in the early rounds. 1 and 2 seeds win the majority of their first round matchups, and no 8th seed has ever won the NBA championship. Despite this, several recent trends suggest that this may not continue to hold true.

History:

Only one 8th seed has ever reached the NBA Finals: the 1999 New York Knicks, in a season shortened to 50 games due to a lockout. Also, only six 8th seeds have advanced past the first round in a best-of-7 series: 1994 Nuggets, 1999 Knicks, 2007 Warriors, 2011 Grizzlies, 2012 76ers, and 2023 Heat.

The 2023 Miami Heat came closest to winning the championship, reaching the Finals as an 8th seed before losing to the Denver Nuggets.

Play-In Tournament:

Introduced in 2020 and fully implemented in 2021, this adds volatility and allows teams with late season momentum to enter the playoffs in strong form.

Load Management:

Star players increasingly rest during the regular season, sometimes suppressing a team's seeding without reflecting its full playoff capability.

Three-Point Shooting Variance:

The increased reliance on three-point shots introduces more randomness, especially in short series, potentially benefiting underdogs.

Increased League Parity:

Salary cap structures, international talent pools, and advanced player development have improved team quality across the board, narrowing the gap between 1st and 8th seeds.

Traditional models like ELO and FiveThirtyEight's RAPTOR have typically given 8th seeds less than a 5% chance of winning a first-round series against 1st seeds. However, betting markets and simulations in recent years have shown increasing respect for well-coached, healthy lower seeds.

Over a long time horizon like the next 500 years, even rare events become increasingly likely given repeated chance.

Other Notable Examples

- Looking at the NHL, where 8th seeds have won championships, such as the 2012 Los Angeles Kings, increased randomness due to goaltending variance makes such outcomes more common.
- MLB wild cards have won multiple World Series titles since playoff expansion.

Forecasting Strategy Overview

We used three main forecasting principles:

- Fermi estimates (we approximated large scale outcomes by breaking down the forecast into estimating success probabilities and estimating a distribution for the finals).
- Combining forecasts (combining empirical probabilities, smoothing and a MC forecast).
- Identifying whether a distribution should be normal, log-normal, power law, or another type of distribution (geometric).

Our general forecasting strategy is thus as follows:

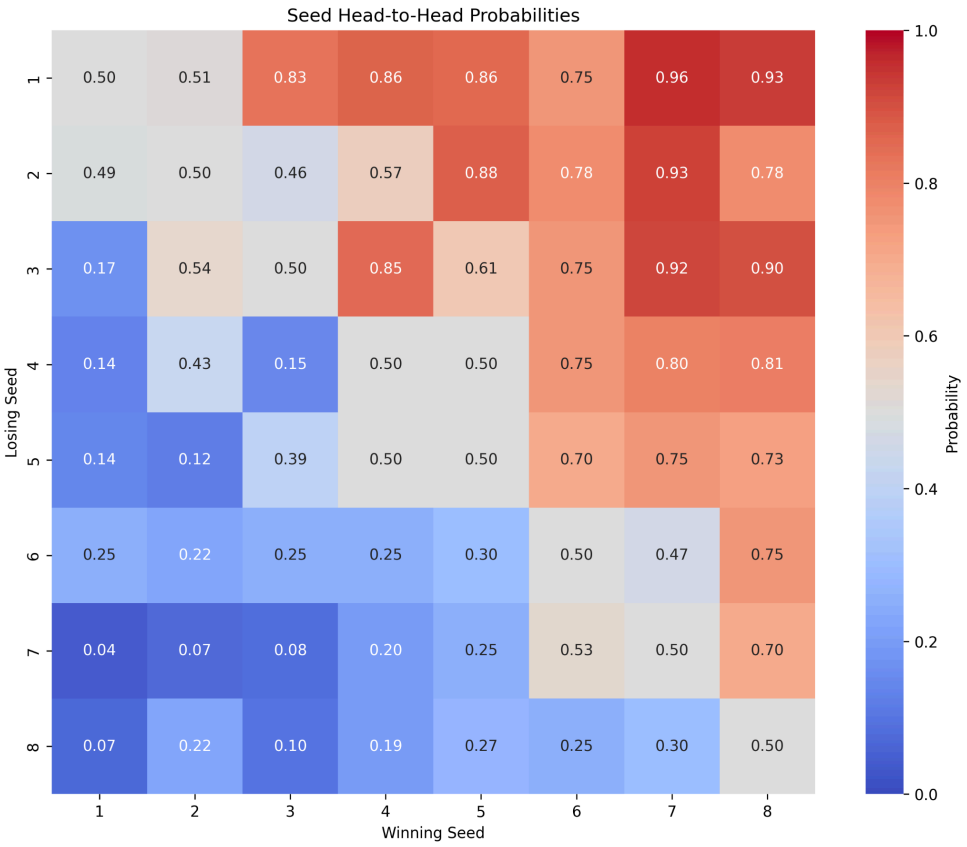
1. Collect data to determine historical probabilities regarding game wins and reaching the finals.
2. Estimate the probability of the 8th seed winning the NBA championship in a given year.
3. Use the probability from step (2) to predict the number of years until the next time the 8th seed wins the NBA championship.
4. Add the result from step (3) to the current year (i.e. 2025) to obtain our desired final forecast.

Data Collection

We first downloaded a csv file with all games played in the playoffs since 1984 when the modern seed championship started. We then cleaned the data into a list of all matches played format, and then iterated through all possible matchups and got probability estimates for a general winning rate.

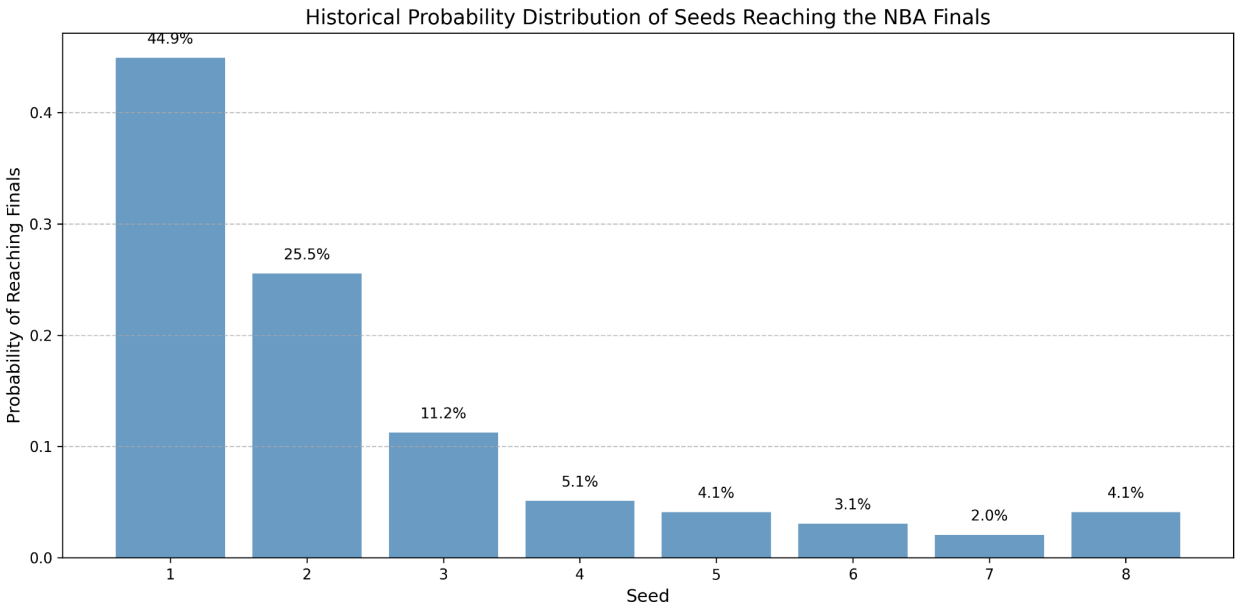
To handle low sampling issues as we only have 40 years of data we applied some smoothing based on seed difference. In particular, we applied weighted smoothing, where we computed a linear combination of the original probability and a scaled factor of seed difference.

We display the computed head-to-head probabilities below:



Using the same dataset, we computed the probability of a given seed reaching the NBA finals. Similarly, we account for low sampling issues by applying laplace smoothing.

We display the distribution of NBA qualification by seed below:



Success Probability Per Year

Using the head-to-head probabilities and tournament bracket simulations, we computed the annual probability that the 8th seed emerges as the champion:

1. Bracket generation: For each year, we reconstructed the tournament tree based on seeding. Note that we ran the bracket simulation on one half of the bracket but for the other half we just got a distribution for a seed going to the finals based on number of times reached since 1984.

2. Monte Carlo simulation: We ran 1,000,000 simulated tournaments, sampling match outcomes according to our estimated probabilities. We tried to run as many simulations as possible to reduce variance.

3. Probability extraction: The fraction of simulations where the 8th seed wins yields that year's success probability.

Hence, for a given year, we computed that the probability that the 8th seed wins the championship is approximately 0.0025.

Final Forecast

We assume that the probability distribution of our desired forecast is geometric, with parameter equal to the probability computed in the previous part. This assumption is due to the relative independence of winning probabilities between each year.

Hence, we forecast that it will take around 354 years until the 8th seed wins the NBA championship, which would occur in the year of 2420.

Based on our simulations, we also compute an 80% confidence interval of [2068, 2952]. Note that this confidence interval is relatively large due to the nontrivial probability of upsets in the NBA playoffs, as well as the unpredictability of NBA matches in general.

You can find our code in the following repo: https://github.com/LeoV974/nba_montecarlo.