Estimating NBA Playoff Success Probabilities

Daniel J. Bell

Faculty Mentors: Dr. Joseph Nolan and Dr. David Agard

NKU Northern Kentucky University

Introduction

In the National Basketball Association (NBA), team success is measured by playoff appearances and championships. Since the merger between the NBA and the American Basketball Association (ABA) in 1976, there have been 35 NBA Champions starting with the 1976-1977 NBA season, as well as many more teams that have made the playoffs, but did not win the championship that year. The goal of this research is to elucidate relationships between measures of individual player success and that of the team. The main research question:

Does the number of All Stars on a team make it more likely to win an NBA championship?

Data[1] were collected from each of the 30 current NBA franchises back to the 1976-1977 season. The following variables were recorded:

- Playoffs – Did the team make the playoffs? (1=yes)
- Playoff Result – Coded 0 to 5.
  - 0 = did not make the playoffs
  - 1 = lost in the first round
  - 2 = lost in the conference semifinals
  - 3 = lost in the conference finals
  - 4 = lost in the NBA finals
  - 5 = NBA Champion
- Wins – The number of times a team outscores their opponent.

Measuring Team AllStarNess

For this project, an All Star is defined as a player who was selected to play in the NBA All Star Game in a season. Being an All Star is an individual honor for the players, but these players play on teams. Therefore a team count of All Stars can be determined. Five metrics (schemes) were developed to obtain the counts of "All Star-ness" for each team each year. This variable is referred to as the team All Star total for a specified scheme.

- Scheme 1 counts only the current year's All Star game.
- Scheme 2 counts only the All Star game from the previous year.
- Scheme 3 uses the total number of previous All Star game appearances in each player's career.
- Scheme 4 employs a linearly weighted decay to the value of prior All Star game appearances.
- Scheme 5 delays the linear decay from Scheme 4 for five years (i.e. the most recent five years have full weight).

[1] Data were collected from http://www.basketball-reference.com

Example

Consider Michael Jordan in the Year 1993. The following table contains his contribution to his Team's All-star Total under the various schemes.

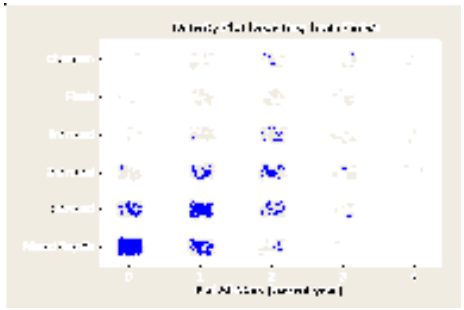
Year	All Star	S_1	S_2	S_3	S_4	S_5
1993	Yes	1	0	0	0	0
1992	Yes	0	1	1	1	1
1991	Yes	0	0	1	0.9	1
1990	Yes	0	0	1	0.8	1
1989	Yes	0	0	1	0.7	1
1988	Yes	0	0	1	0.6	1
1987	Yes	0	0	1	0.5	0.8
1986	Yes	0	0	1	0.4	0.6
1985	Yes	0	0	1	0.3	0.4
1984	No	0	0	0	0	0
1983	No	0	0	0	0	0
Total	**N/A**	**1**	**1**	**8**	**5.2**	**6.8**

The following table represents the contribution Michael Jordan **would have** made to his team for each year of his career:

Year	All Star	S1	S2	S3	S4	S5
2003	Yes	1	1	13	3.4	4
2002	Yes	1	0	12	2.8	3.8
2001	No	0	0	12	3.3	4.4
2000	No	0	0	12	3.9	4.8
1999	No	0	1	11	4.6	5.6
1998	Yes	1	1	10	4.3	5.2
1997	Yes	1	1	9	4	5
1996	Yes	1	0	9	3.7	5
1995	No	0	0	9	4.5	6
1994	No	0	1	9	5.4	7
1993	Yes	**1**	**1**	**8**	**5.2**	**6.8**
1992	Yes	1	1	7	4.9	6.4
1991	Yes	1	1	6	4.5	5.8
1990	Yes	1	1	5	4	5
1989	Yes	1	1	4	3.4	4
1988	Yes	1	1	3	2.7	3
1987	Yes	1	1	2	1.9	2
1986	Yes	1	1	1	1	1
1985	Yes	1	0	0	0	0

**Note: In some of these years, he was not actually on a team.

Test for an Association



Analysis #1: Chi-Square Test shows an association between "Finish" and Team All-Star Total (Scheme 1).

Statistic	DF	Value	Prob
Chi-Square	20	392.1369	<.0001

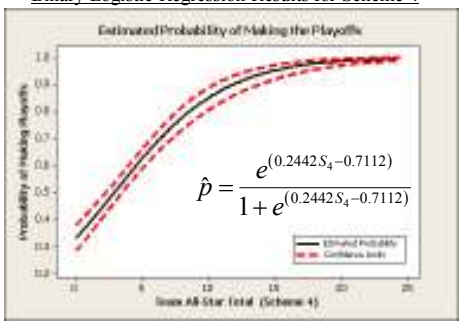
Analysis #2: Gamma Test shows positive ordinal correlation.

Statistic	Value	ASE	Confidence Limits
Gamma	0.6934	0.0228	0.6487-0.7381

Estimating Probabilities

Binary and Ordinal Logistic Regression Models are employed to predict the probability of a Team Finish based on Team All-Star Total using the various schemes. Scheme 4 is used to illustrate results.

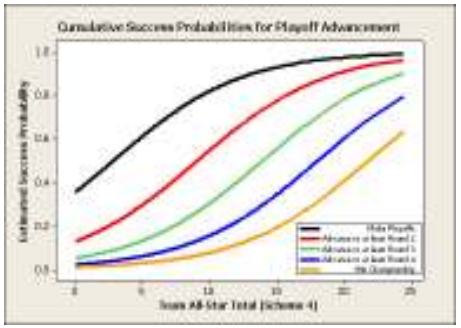
Binary Logistic Regression Results for Scheme 4



Parameter	DF	Estimate	StdErr	Wald χ^2	P-value
Intercept	1	0.7112	0.1096	42.07	<.0001
Sum4	1	-0.2442	0.0229	113.52	<.0001

Estimated Probabilities

Ordinal Logistic Regression Results for Scheme 4



Statistical analysis indicated a significant relationship between Team All-Star Total for Scheme 4 and the estimated success probability (Wald $X^2 = 177.76$, p-value <0.0001). The plot above shows estimated probability curves for success at each round of the playoffs. For example, a team with an AST score of 16 would have roughly a 20% chance of winning the championship.

A comparison of the measurement schemes is also interesting:

Scheme	R-squared Logistic	R-squared Ordinal
1	0.2744	0.3308
2	0.1467	0.1745
3	0.1228	0.1471
4	0.1522	0.1851
5	0.1415	0.1738

Greater R^2 suggests greater association. As expected, current year all-star status has the greatest association. It is interesting to note that the schemes employing decay seem to do better than Scheme 3 (which simply totals the number of historical All-Star appearances for a team).

Future Ideas

Future research may focus on any of the following areas:

- There are infinitely many weighting schemes available. Further investigation could attempt to identify an overall "best" scheme.
- Some salary data was collected and the relationships to team success very briefly explored. Additional data could make it possible to employ team salary as a predictor in a similar model.

Acknowledgement: This research was a product of a UR-STEM summer undergraduate research experience supported by *NKU FORCE: Focus on Occupations, Recruiting, Community, and Engagement*