

## Reaching the NFL Playoffs Based on Week One Results: A Probability Model with Simulation

Peter Boyd

*Purdue University*, [peterboyd03@gmail.com](mailto:peterboyd03@gmail.com)

Brian Kidd

*Purdue University*

Christopher Vincent

*Purdue University*

Follow this and additional works at: <https://scholar.rose-hulman.edu/rhumj>

---

### Recommended Citation

Boyd, Peter; Kidd, Brian; and Vincent, Christopher (2019) "Reaching the NFL Playoffs Based on Week One Results: A Probability Model with Simulation," *Rose-Hulman Undergraduate Mathematics Journal*: Vol. 20 : Iss. 1 , Article 1.

Available at: <https://scholar.rose-hulman.edu/rhumj/vol20/iss1/1>

## Reaching the NFL Playoffs Based on Week One Results: A Probability Model with Simulation

---

By *Brian Kidd, Peter Boyd, and Christopher Vincent*

**Abstract.** We consider the question of how important winning the first game of the season is to making the playoffs for an NFL team. We analyze this question both statistically and probabilistically. First we examine historical data from past NFL seasons to consider whether the first week of the season is any more important than other weeks of the season. Secondly, we attempt to explain probabilistically how winning in any given week changes the probability of that team making the playoffs.

The purpose of the research in this paper was to determine if the first week of the NFL season is more significant than other weeks in terms of whether or not the team in question makes the playoffs. We used data of games from past seasons to analyze each week across all seasons. Week One was shown to be no more significant than any other week.

### 1 Introduction

The National Football League has been providing a source of entertainment to the United States and other countries since the early 1900s. Although people are drawn to the NFL for exciting games and highlights, many fans are also interested in the statistics that govern the predictions and outcomes of the games. In 2012, NFL Media Senior Analyst Gil Brandt stated,

“How important is Week 1? Since 1978, when the NFL went to the 16-game schedule, teams that are victorious on Kickoff Weekend are more than twice as likely to reach the playoffs than losers of the opening game.” [1]

The presentation of this statistic seems to imply that winning in week one is *more important* than winning in other weeks with respect to making the playoffs. This implication raises the following question: is week one truly unique, or is this statistic something we could find across all weeks? Does winning in week one build “momentum” that can

---

*Mathematics Subject Classification.* 62P99

*Keywords.* probability, simulation, sports analytics, NFL

help push the team to a successful season, or does the NFL season operate more like a sequence of independent trials?

In this paper, we address week one of the NFL season both statistically and probabilistically.

There are two intuitive reasons why teams that win in a given week are more likely to make the playoffs, and in the second part of the paper we attempt to give a probabilistic analysis of these factors.

**A head start** Teams that win in week one (or any other week) have a one game advantage over teams that lose in that week. This gives them an advantage to finish the season with a better record than the teams that lost in that week.

**An updated opinion** Prior to the season starting, no one knows for sure how good a team will be. Given the information on the results of week one, however, we can form an updated opinion on the strength of the team for the season. Teams that win in week one are now more likely to be one of the strong teams than those who lost in week one.

In this paper we build a simplified probabilistic model for an NFL season in which we can examine how these two factors affect a team's chances of making the playoffs. Our focus in this paper is not only to explain the model, but to show the steps that one goes through in building a probabilistic model: starting with a simplified model that is easy to analyze and gradually building more realistic features into the model.

## 2 Statistical analysis of weekly win importance

### 2.1 Description of the Data

The data we use in our statistical analysis was scraped from the NFL & Pro Football League Encyclopedia at [pro-football-reference.com](http://pro-football-reference.com) and analyzed using R Studio. We use the win/loss from the season results from 1978 to the 2014 season. For our first analysis, we start with 1978 in order to coincide with Brandt,s statistics.

In this time frame, the structure of NFL seasons has undergone change several times. In 1982, there was a player strike which limited the season, with weeks 3-10 not being played and a 17th week being added. In 1987, another player strike led to a cancellation of week three. Beginning in 1990, a 17th week was implemented to provide teams with a bye week. Despite minor changes and differences among some seasons, analyses will not be impacted dramatically. Our goal is to understand how each weekly outcome affects a team's probability of making the playoffs. Thus, missing weeks due to the circumstances listed above does not impede our objective, rather, there are simply a nonconstant number of total games played in each week. In 1993 only, the NFL experimented with providing teams two bye weeks, causing there to be an 18th week in this season, which

we ignored in our analysis to maintain a consistent maximum of 17 weeks. Although games that end in a tied score are quite uncommon, there were a few occurrences, which were removed from the data set. New teams being added to the league, along with the previously discussed changes, led to inconsistencies in the total number of games played each week. Similarly, the number of teams that made the playoffs varied as well.

In 2002, the NFL was restructured into two conferences comprised of four divisions per conference. The playoffs of 2002 also established a new system where six teams advanced to the playoffs from each conference: four division winners and two wild-card teams, teams with the best record that were not division winners [3].

## 2.2 Past Analysis

We begin by focusing on how history unfolded in regards to winning or losing a specific game and reaching the playoffs. We look at results from 1978 through 2014 and compare these results to Brandt's statement that teams winning in week one are more likely to make the playoffs. Table 1 lists all of the results combined by week giving the number of games, number of playoff teams, and the probabilities of making the playoffs given a win or loss.

As seen in Table 1, given that a team wins in any one week, the probability that the selected winning team advances to the playoffs appears to be roughly equal among all 16 weeks.

From our analyses thus far, we observe that winning in any week yields a probability of 0.523 of making it to the playoffs. Alternatively, a loser in any given week has a probability of 0.242 of making it to the playoffs.

With an understanding of the data, we now shift towards attempting to model seasons probabilistically.

# 3 Probabilistic Analysis

## 3.1 First Probabilistic Model

Our objective in building a probabilistic model is to show how the two intuitive effects of winning in week one ("head start" and "updated opinion") mathematically increase the chances of that team making the playoffs. In the end we wish to have a model that is somewhat realistic and also closely matches the statistical data from the previous section. However, we will first present a very simple model that demonstrates how winning changes our opinion of a team's strength, and we will then add more realistic features that allow us to more closely model reality. While the initial models are simple and may not be realistic, these models are used as preliminary steps that gradually increase in accuracy.

Week	Winners Making Playoffs	Losers Making Playoffs	Total Games	P(Playoffs Win)	P(Playoffs Loss)
1	318	143	597	0.5327	0.2395
2	303	155	597	0.5075	0.2596
3	298	113	500	0.5438	0.2062
4	268	129	537	0.4991	0.2402
5	269	136	530	0.5075	0.2566
6	274	137	528	0.5189	0.2595
7	279	115	528	0.5284	0.2178
8	268	126	523	0.5124	0.2409
9	280	128	531	0.5273	0.2411
10	294	124	548	0.5355	0.2263
11	304	145	583	0.5214	0.2487
12	319	138	595	0.5361	0.2319
13	307	153	598	0.5134	0.2559
14	309	152	596	0.5185	0.2550
15	322	140	599	0.5376	0.2337
16	315	145	599	0.5259	0.2421
17	237	117	446	0.5314	0.2623
Average	292	135.0588	557.8235	0.5234	0.2422

Table 1: Results by week of the 1978 through 2014 NFL seasons. Beginning in 1990, a 17th week was implemented to provide teams with a bye week. In 1982, a seventeenth week was added due to a player strike during weeks 3-10. In 1987, a player strike led to no games in Week Three. Week 18 only occurred in 1993 to allow for two bye weeks [3].

In our simplified model, we begin by assuming that the 32 teams are ranked from 1-32, with 1 representing the best team and 32 representing the worst team.

Rather than asking if a given team makes the playoffs, we will instead ask the simpler question of if the team is one of the top 12 ranked teams. Suppose we pick one of the teams at random and let  $\mathcal{R}$  be the rank of the team chosen. Since  $\mathcal{R}$  is uniformly distributed on  $\{1, 2, \dots, 32\}$  it follows that  $P(\mathcal{R} \leq 12) = \frac{12}{32} = 0.375$ . However, if we play one week of an NFL season, knowledge of the outcome may change our opinion of the distribution of  $\mathcal{R}$ . To model this probabilistically, we make the following two assumptions:

- The NFL schedule for week 1 is randomly chosen; that is, the 32 teams are randomly paired together in 16 games with each pairing being equally likely.

- In each game the team with the higher ranking *always* wins.

If  $\mathcal{W}$  denotes the event that our previously randomly chosen team is the winner in its matchup in this random scheduling, we are interested in the probability  $P(\mathcal{R} \leq 12 | \mathcal{W})$ .

Teams are still uniformly distributed to the rankings, so  $P(\mathcal{R} = k) = \frac{1}{32}$ . For any week, since we assume no ties, there is a winner and a loser, so by symmetry,  $P(\mathcal{W}) = \frac{1}{2}$ . Furthermore, by applying Bayes' rule, we find

$$\begin{aligned} P(\mathcal{R} = k | \mathcal{W}) &= \frac{P(\mathcal{R} = k, \mathcal{W})}{P(\mathcal{W})} \\ &= \frac{P(\mathcal{R} = k) \cdot P(\mathcal{W} | \mathcal{R} = k)}{P(\mathcal{W})} \\ &= \frac{P(\mathcal{W} | \mathcal{R} = k)}{16} \end{aligned} \tag{1}$$

The probability that a team wins given their rank seems much more intuitive. In this deterministic model, a team ranked  $\mathcal{R}$  will beat all teams with a lower ranking. For example, take the team ranked 5<sup>th</sup>. This team could be paired against any of the 31 other teams in the league. Of these 31 possible opponents, the team will beat the teams ranked 6<sup>th</sup> through 32<sup>nd</sup>, which is  $32-5=27$  teams. Therefore,

$$P_k = P(\mathcal{W} | \mathcal{R} = k) = \frac{32 - k}{31}, \quad k = 1, 2, \dots, 32. \tag{2}$$

By combining this knowledge into Equation 1, and summing over the top 12 ranks we obtain

$$P(\mathcal{R} \leq 12 | \mathcal{W}) = \sum_{k=1}^{12} P(\mathcal{R} = k | \mathcal{W}) = \frac{1}{16} \sum_{k=1}^{12} \left( \frac{32 - k}{31} \right) = 0.6169.$$

This value is not the same as what history has shown, as seen in Table 1, but the inconsistency can be attributed in part to the fact that the higher ranked team does not always win in actuality, along with the simplicity of the model.

### 3.2 Win Probability Function

To make the model more realistic, let  $P_k(j)$  be the probability that team ranked  $k$  beats team ranked  $j$ . In our previous analyses, the better team always won, which does not reflect what happens in reality. A more accurate win probability function will include the following:

- The better team (lowering ranking) will be more likely, but not guaranteed, to win.
- $P_k(j)$  will be monotone increasing in  $j$  and monotone decreasing in  $k$ .

- The sum of  $P_j(k)$  and  $P_k(j)$  must equal 1 because one of the two teams must win.
- When teams are closely ranked, the win probability should be close to one half for each team.

By accounting for the list above, we utilize the difference between the two teams ranks to create our win probability function; many different functions can be employed in this scenario, but, based on the conditions listed above, we chose the following function:

$$P_k(j) = \frac{j-k}{64} + \frac{1}{2}. \quad (3)$$

For all following analyses, we no longer define  $P_k$  as seen earlier in Equation (2). To find a new  $P_k$ , we calculate the expected value of the probability that  $P_k(j)$  over all  $j \neq k$ , leading to

$$P_k = \sum_{\substack{j=1 \\ j \neq k}}^{32} \frac{P_k(j)}{31} = \frac{1}{31} \left( \sum_{j=1}^{32} \left( \frac{j-k}{64} + \frac{1}{2} \right) - \frac{1}{2} \right) = \frac{95-2k}{124}. \quad (4)$$

Equation (1) leads to

$$P(\mathcal{R} = k | \mathcal{W}) = \frac{1}{16} \frac{95-2k}{124}.$$

Because 12 teams advance to the playoffs, for simplicity, we assume that the 12 lowest ranked teams advance. Now, by combining Equations (1) and (4) and summing over ranks 1 through 12, we find

$$P(\mathcal{R} \leq 12 | \mathcal{W}) = \frac{1}{16} \sum_{k=1}^{12} \left( \frac{95-2k}{124} \right) = 0.4960.$$

This probability is much closer to the historical value seen in Table 1, indicating improvement in our model.

### 3.3 Head Start Theory

As noted in the beginning, there are two intuitive reasons why winning in week one increases the likelihood of a team making the playoffs. Having already explained the updated opinion theory, we now turn to the head start theory.

As a first step, since making the playoffs is primarily based on the number of wins a team has, we first find the probability distribution of the number of games, denoted  $N$ , that a team wins in a season. In the NFL, teams are first designated playoff spots based on division placement, followed by conference placement, etc. The NFL has complicated set tiebreaking procedures that are followed if teams end with the same record.[4] In the simulation component of our report, if teams end with a tied record, we will randomly select one of the tied teams as a tiebreaker.

We begin by conditioning not only winning the first week, but also on the team's rank. This conditional probability will follow the Binomial distribution with  $n = 15$ , the number of remaining games in the season, and  $p = P_k$ , the average win probability function for unknown opponents.

$$P(N = w \mid \mathcal{W} \cap \{\mathcal{R} = k\}) = \binom{15}{w-1} P_k^{w-1} (1 - P_k)^{16-w}.$$

Note that in the above calculation, we assume that each week consists of 16 randomly assigned pairings of teams. Theoretically, teams could play each other multiple times, which maintains independence. However, we want the probability to be conditioned only on winning week one because we hope to generalize a standard probability independent of rank. To accomplish this task, we must first remove the conditioning on rank shown below:

$$\begin{aligned} P(\{N = w\} \cap \{\mathcal{R} = k\} \mid \mathcal{W}) &= P(N = w \mid \mathcal{W} \cap \{\mathcal{R} = k\}) \cdot P(\mathcal{R} = k \mid \mathcal{W}) \\ &= \frac{P_k}{16} \left( \binom{15}{w-1} P_k^{w-1} (1 - P_k)^{16-w} \right). \end{aligned} \tag{5}$$

Then, because each team has a unique rank, we can sum the above probability over all ranks to obtain the probability of having  $w$  wins given a win in week one

$$\begin{aligned} P(N = w \mid \mathcal{W}) &= \sum_{k=1}^{32} P(\{N = w\} \cap \{\mathcal{R} = k\} \mid \mathcal{W}) \\ &= \sum_{k=1}^{32} \frac{P_k}{16} \left( \binom{15}{w-1} P_k^{w-1} (1 - P_k)^{16-w} \right). \end{aligned} \tag{6}$$

For reference, we compare the distribution from Equation (6) to the distribution of  $N$  prior to the start of the season. Following the same procedure without conditioning on winning week one leads to

$$P(N = w) = \sum_{k=1}^{32} \frac{1}{32} \left( \binom{16}{w} P_k^w (1 - P_k)^{16-w} \right). \tag{7}$$

In Figure 1, we see that the Head Start Theory can add a shift to the distribution and significantly change probabilities of number of games won after we already know that a



team won in the first week. When it is known that a team won in week one, the probability that the team wins fewer than 8 games is less than the probability calculated before week one knowledge of winning the same number of games; once we know that a team has already won the first game, the probability of winning a low number of games has decreased. Probabilities for winning fewer than eight games is less than the assumed probability of winning the same number of games prior to the start of the season. However, the probability of winning more than nine games increases considerably from the initial probabilities once we have already observed one win for a team.

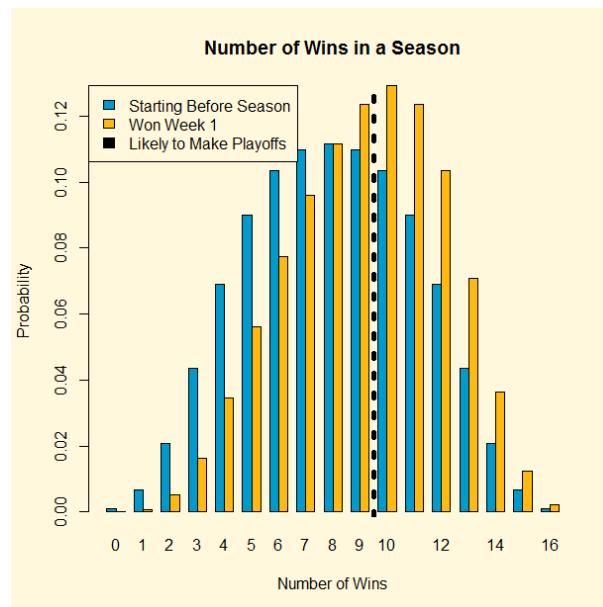


Figure 1: This graph indicates the probability of achieving a number of wins in a given season. The black dashed line represents the average wins needed to make the playoffs. We can note the bell-shaped curve for both conditions with emphasis on the shift to the right for winning week one. A team that wins week one is more likely to be a good team since they have already won one game, leading to our beliefs of the Head Start and Updated Option Theories.

Knowing that 12 of 32 teams make the playoffs, we use Equation 7 to estimate the number of wins needed to reach the playoffs. We sum the probability of winning  $w$  games until we can bound the probability  $\frac{12}{32} = 0.375$ , resulting in

$$0.334 \approx P(N \geq 10) < \frac{12}{32} < P(N \geq 9) \approx 0.444.$$

Using this inequality and weighted averaging, we can find the estimated number of wins needed to make the playoffs to be 9.63. Thus, 10 wins are typically needed to make the

playoffs, which is represented in Figure 1 as the dashed line. To get an estimate of the probability of making the playoffs after conditioning on winning week one, we apply the same weights to obtain

$$P(\text{Playoffs} | \mathcal{W}) \approx 0.63P(N \geq 10 | \mathcal{W}) + 0.37P(N \geq 9 | \mathcal{W}) = 0.5236$$

which is very close to the historical value, suggesting that our function is fairly accurate.

Although this approximation has value, it was calculated using the distribution of number of wins for only one team. To calculate  $P(\text{Playoffs} | \mathcal{W})$  exactly, we must know the joint distribution for number of wins for all 32 teams in order to find the top 12 teams. The joint distributions quickly become intractable, so to circumvent this issue we rely on simulation to address these issues without explicitly calculating the values.

## 4 Simulation

We simulate one thousand seasons in order to find the probability that a team makes the playoffs given a certain number of wins. The simulation utilizes the win probability function expressed in Equation 3. In order to perform the simulation, we make several assumptions and use win probability function seen in Equation 3.

- Assume no ties, all games result in one winner and one loser.
- Teams are randomly paired for each game each week.
- For teams the same number of wins at the end of the season, tiebreakers are randomly chosen, assigning one team to advance to the playoffs.

With these assumptions, we began the simulation by performing one week of games. The winners of these games were recorded. We simulate the remaining 15 games of the season in the same manner. Using the combined 16 games, the 12 teams with the most wins were selected as the playoff teams, loosely following the structure of the NFL. We simulated 1000 seasons and found the following Monte-Carlo estimate.

$$P(\text{Playoff} | \{N = w\} \cap \mathcal{W}) \approx \frac{\text{Number of Playoff Teams That Won Week 1 with } w \text{ Wins}}{\text{Number of Teams That Won Week 1 with } w \text{ Wins}}. \quad (8)$$

The results of 1000 iterations in the simulation are seen below in Table 2.

We continue the analysis by finding the probability that a team makes it to the playoffs, given they won week 1. Although we have found Equation (8), we will need to

Wins	0-7	8	9	10	11-16
Emperical Probability	0	0.0107	0.4194	0.9574	1.0

Table 2: Emperical probability of making the playoffs, given week one win and fixed number of total wins.

multiply it by Equation (6):

$$P(\text{Playoffs} \cap \{N = w\} | \mathcal{W}) = P(N = w | \mathcal{W}) \cdot P(\text{Playoff} | \{N = w\} \cap \mathcal{W}).$$

By then summing the right side over all  $w$  wins, we can find the final probability of making the playoffs given a week one win through simulation.

By replicating the above process, we found the probability that a team makes it to the playoffs, given that they won week one, to be

$$P(\text{Playoffs} | \mathcal{W}) \approx 0.5254258. \quad (9)$$

This simulation based probability closely represents the probability of making the play-offs given a week one win (0.5327), observed from NFL data, as seen in Table 1, thus indicating strong predictive accuracy of our model.

## 5 Discussion

As we initially speculated in our introduction, week one is no more indicative than any other week during the season in terms of increasing the likelihood of making the playoffs. This notion is intuitive because making the playoffs is based on the amount of wins a team has in comparison to other teams; winning a game in any week improves a teams' chances of making the playoffs.

Next, we created a probabilistic model that found probabilities close to the “real” probabilities found from past data. As we added complexity to our model, the probabilities improved to better estimates of the probabilities from the data. While the end result is not perfect, we believe that it allowed for adequate analyses without adding too many complexities. Because equation (9) was close to the historical value, history supports the plausibility of our model and provides further evidence in favor of our hypothesis. In conclusion, we found that while Brandt was not incorrect in his statistical findings, his findings can be misinterpreted as week one is more important than other weeks.

### 5.1 Future Work

This paper provides an introduction for statistical modeling of NFL seasons and evaluating past data. However, there are many possible considerations for future work. The

model can become more realistic by:

- Incorporating conferences and divisions
- Create schedules that are true to NFL scheduling parameters

We simulated seasons considering the above facts but found results very similar to our simpler simulations, so we did not elaborate on our findings. Additionally, our win probability function is quite simple. Many other factors may be valuable in making the model more realistic, such as

- Home Advantage
- Weather
- Winning and losing streaks
- Recent trades

Our simulations are based largely on our assumption of a constant, well-defined ranking system throughout the NFL. Because such a system does not explicitly exist, our simulations currently outline possible procedures to infer results as they pertain to weekly wins. In order to create more realistic simulations, one may consider the implementation of pre-season rankings and predictions published by NFL analysts. By incorporating conferences and divisions, as previously stated, a more realistic modeling approach could be reached. Additionally, it is quite common for National Collegiate Athletic Association (NCAA) football teams to be ranked in a manner that we have implemented in our simulations. Similar analyses could be performed at the collegiate level to consider a broader perspective and extrapolate beyond Gil Brandt's statement.

It should first be determined if these factors, as well as others, do indeed affect a team's record throughout the season. If so, can they be modeled probabilistically and incorporated into a win probability function? Our modeling attempts have laid the foundation for adding such complexities to better model the NFL.

## References

- [1] Brandt, G. (2012, August 3). Season Starts with Some Great Quarterback Duels. *NFL.com* Retrieved from <http://www.nfl.com/news/story/09000d5d81a5d175/article/season-starts-with-some-great-quarterback-duels>.
- [2] J. P. Keener, The Perron-Frobenius theorem and the ranking of football teams, *SIAM Rev.* **35** (1993), no. 1, 80–93. MR1207798

- [3] National Football League *History*. Accessed September 18, 2015. Retrieved from <http://www.nfl.com/history/chronology/1869-1910>.
- [4] National Football League *Standings*. Retrieved from <http://www.nfl.com/standings/tiebreakingprocedures>.
- [5] Sports Reference LLC. (2015) *NFL & Pro Football League Encyclopedia*. Retrieved from <http://www.pro-football-reference.com>

**Brian Kidd**

Purdue University, Texas A&M University

**Peter Boyd**

Purdue University, Oregon State University

**Christopher Vincent**

Purdue University, Indiana University