Tampa Bayes Buccaneers: Predicting a Rookie Quarterback's True Completion Percentage

Iris Gupta, Arthur Hu, Shariar Vaez-Ghaemi, James Zhou

Abstract:

Predicting the future performance of athletes plays a big role in sports. We set out to find a method to predict completion percentages for quarterbacks in their rookie NFL using college performance and their performance in NFL games. We collected college stats and NFL stats for 30 rookie quarterbacks between 2012 and 2020. Using this data, we used Bayes Theorem to predict what a quarterback's "true" completion percentage was. We then evaluated the model in two ways to determine whether the prior, the college data, or the posterior, the NFL data, was more important in our model. The priors were not a good predictor of the end of season predictions, indicating that our model put a lot of emphasis on more recent data. We then analyzed how our model progressed over the season; our models predictions moved gradually through the season, indicating that our model was not easily swayed by a game that was particularly good or bad. Rookie quarterbacks often progress in their playing: the average quarterback from our sample had a 7% difference between their completion percentage at the end of the season compared to their college numbers. Our model predicting at halfway through the season was on average 2.97% off from the end of season average. While our model relies on a few large assumptions about quarterbacks, we were able to create a model that adjusted for the improvement of a player over the course of a season without being moved by the most recent games. This shows that there may be some fundamental difference between college and NFL football leading to a difference in performance.

Introduction:

For the purposes of this paper, we focused on football with passing completion percentage as a performance indicator due to the robust amount of data available. Particularly in football though, there surrounds an audience culture where bets are placed on player performance, which influences bets placed on team performance. Questions about football player performance are also important for coaches and team decision-makers because it influences the probability of keeping and trading players. Beyond the sport, insight about player performance can inform the processes of aging, muscle physiology, and the cardiovascular system.

As consumers of football, there has long been speculation and debate about how to make statements about player performance. In the past, projections were based solely on college playing performance and used to determine outcomes on draft day. However, for intuitive reasons, these projects were not generally accurate of future player performance. Too much reliance on college football data poses a problem because college football schemes are often different from NFL schemes. Therefore, it is important to account for more recent data which can provide us with more holistic projections of the football player's performance. On the flipside, it is problematic to focus solely on rookie data because rookies have an extremely small sample of performances in their first year.

We want to determine a method to reconcile the two data sets of prior and posterior completion percentage for several football players. More specifically, we plan to predict an NFL quarterback's "true" game completion percentage given their college completion percentage (prior) and their NFL completion percentage (posterior) for one year. In order to do this, we collected week-by-week data for 30 football players, accounting for the college stats and their NFL stats for the year they were drafted. We focused on players who individually have robust data sets and used data spanning from the years 2012 to 2020. We included 2-5 players for each year.

Methodology:

We created our Bayesian classifier in R.

We first created a two tailed-probability density function to determine the likelihood of seeing an observation *as crazy* as a given observation, given a normal distribution of expected outcomes. Here, we make the assumption that an NFL quarterback's game completion percentages follow a normal distribution. This is a flawed assumption to make about rookie quarterbacks because their number of pass attempts vary by game. For example, Lamar Jackson was 1-1 (100%) in Week 9 of his rookie season and 13-19 (68%) in Week 10. If he had thrown 40 pass attempts per game, variation would condense into a normal distribution, but that was not the case. Regardless, using a normal distribution was the easiest way of implementing Bayes.

In order to estimate a parameter with Bayes' Theorem, we needed to estimate the probability of seeing a *series* of data given a normal distribution. Our version of this function simply multiplies the probability densities of each observation in a set of data. So, if P(A% or crazier) = 0.4, P(B% or crazier) = 0.1, and P(C% or crazier) = 0.8, then P(A,B,C over three games) = 0.4*0.1*0.8 = 0.032.

In using the interpretation of Bayes Theorem below, P(A|B) is the probability of a player's "true" completion percentage being A% given the observations in set B. P(B|A) is the probability of seeing the observations in set B given a completion percentage of A%, while P(A) is the probability of a player's true completion percentage being A% given our prior knowledge from that player's college data. P(B) is the probability of seeing observations in B given our prior information; because P(B) is constant, it is not important for maximum likelihood estimation.

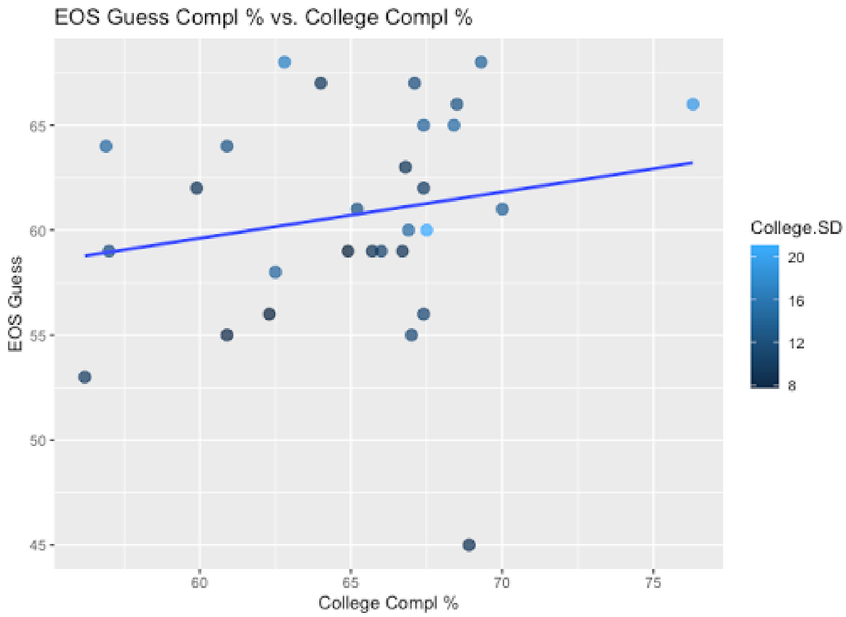
$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Analysis:

We decided to evaluate our model by identifying what information our model deemed important and what information it put low emphasis on. We chose this framework of analysis because it is similar to how people on NFL Twitter evaluate each other's analysis. Anyone who has ever seen the words "Y'all sleeping on Johnny Manziel for one bad NFL game!" understands the importance of weighting new information appropriately.

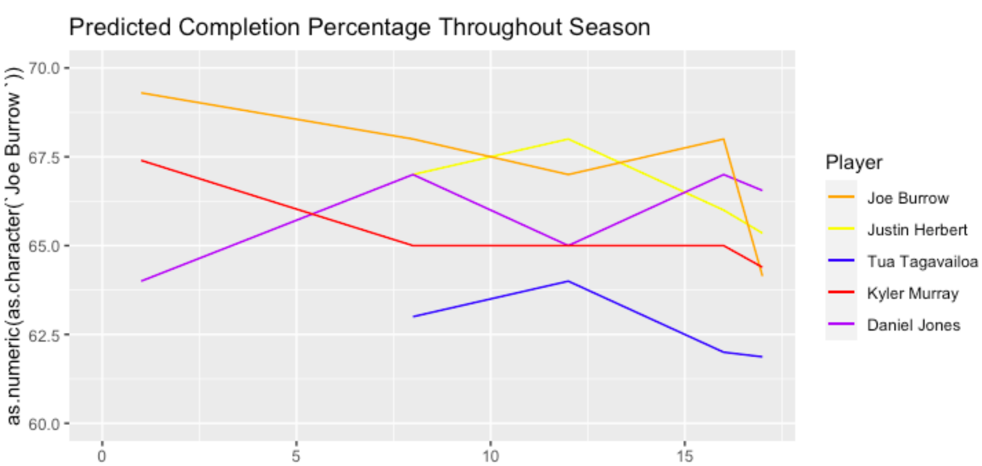
Figure 1 below shows how closely our model "hugs" the prior. The x-axis displays the player's college completion percentage, and the y-axis displays the estimated completion percentage outputted by our model after the end of the player's rookie season. The color of each dot indicates the player's standard deviation in college. The relationship between college completion percentage and end-of-season guess is very weak, which shows that our model is somewhat naive – meaning that it treats new observations with more weight than priors.

Figure 1**:** Relationship Between Prior and End-of-Season Output



Next, we chose to analyze how the trajectory of our model's guess changes throughout the season. Nobody likes an NFL analyst who constantly changes their perception of a player, so we wanted to make sure that our model wasn't a digital Max Kellerman. Figure 2 focuses on five of the quarterbacks taken in 2019 and 2020. Each player's estimated completion percentage at the start of the season, after their first half of rookie starts, after their first ¾ of rookie starts, and after all of their rookie starts is shown. Our model doesn't make big leaps in guesses, even when a player's completion percentage drastically jumps week-by-week. For example, when Tua Tagovailoa threw a 100% completion percentage in Week 6, our model resisted the urge to go bullish on the rookie QB.

Figure 2: Model's Perceptions of Players Over Time



Among the 30 quarterbacks in our dataset, the average quarterback's college completion percentage was 7% off from the completion percentage at the end of the rookie season. 7% is a sizable jump (or fall) for any quarterback, which is why this Bayesian estimation was necessary in the first place. The average quarterback in our list has a 2.97% difference between the Bayesian output halfway through the season and their completion percentage at the end of their rookie season. We consider this to be very impressive because it shows that our Bayesian model can make almost the same conclusion about a player's accuracy half-way through their rookie season as a stat-watcher could make at *the end* of the player's rookie season.

In the last paragraph, we said the model was strong because its estimates for completion percentage early in the season were similar to the cumulative NFL completion percentage at the end of the player's season. We want to note that it is extremely flawed to treat the cumulative end-of-season completion percentage as the "true parameter." Firstly, players improve throughout their rookie season, and cumulative percentages don't account for improvement. Second, there are so few games in an NFL season that by the end of a rookie season, there is still a lot of variation.

Discussion:

We can determine from figure 1 that college completion alone is not a good predictor of a quarterback's predicted completion percentage by the end of their rookie season, which suggests that their performance can change dramatically over the course of the season.

Figure 2 shows that the model is not too sensitive to new data points, as we notice that the predicted completion percentage of five selected quarterbacks did not change drastically over one game, but rather gradually over the course of the season. This means that our model is able to weigh recent data more heavily, but still views the quarterback's performance holistically when making its prediction. The model does occasionally make significant adjustments, as seen with Joe Burrow on the last game of the season, where Burrow's predicted completion percentage went from the highest to second lowest of the five quarterbacks listed.

Our model hinges on the assumption that the game completion percentages can be organized in a normal distribution. This is not the case, as quarterbacks simply do not have enough pass attempts per game, making certain data points susceptible to existing on the extremes. The number of games that we can include is also limited, as the rookie season is at most 16 games, with many players missing games. For instance, Tua Tagavailoa missed the first half of the season, which makes the season's end prediction less accurate compared to other players.

Conclusion:

Our Bayesian classifier attempts to serve as a metric of predicting a rookie quarterback's future prospects by answering the question of how accurate college completion percentage is at foreshadowing the predicted completion percentage of that player at the end of their rookie season. By weighing more recent performance more heavily, we are able to adjust for the improvement of a player over the course of a season without over sensationalizing on the most recent data point. While our model struggles from various flaws, it is able to uncover a relationship that is weak, but present between college and NFL completion percentage, suggesting an underlying difference between college and NFL football that causes certain college players to not live up to their Draft Day hype. This serves as a warning that college performances can be deceiving, but also as a sign that players can be late bloomers who only unleash their full potential on the biggest stage. It is important to note that our model has to work with very few games, making it difficult to parse the variation in the data and gain a clear picture between the prior and posterior. Moreover, rookie quarterbacks will sometimes have very few attempts in any given game, skewing the normal distribution our model relies on to its extremes. It is likely that our model would attain a more concrete picture of a player's performance among other sports that give us more data to work with, such as cricket.