Conditional Probability Adjustments for NFL Kickers

1. Without adjusting for distance, the highest FG% belonged to Matt Bryant, who had a 93.1 FG% in 2011.
2. Shaun Suisham had the lowest FG% at just 74.19%.
3. The kicker that improved the most when adjusting for distance is Sebastian Janikowski, whose 91.68% adjusted FG% is a 3.11% increase over his actual FG%. Ryan Succop also improved, with his 82.61% adjusted FG% being 2.61% than his actual FG%. The 3rd-highest improvement comes from Nick Folk, whose 78.25% adjusted FG% is 2.25% better than his actual FG%. These improvements occurred because we adjusted for distance of kicks using the league-wide probability of kicking from each specific distance range. Take Sebastian Janikowski, one of the strongest kickers in NFL history, as an example. He had the most field goal attempts from 50 and 59 yards, and he was near the top for field goals attempted between 40 and 49 yards. His actual FG% does not factor in that these kicks are more difficult than shorter field goals. This means that his actual FG% punishes him for attempting more difficult field goals, which is a result of him having a strong left leg. The adjusted FG% takes into consideration the league-wide probability that a field goal attempted is from these far ranges, which helps value kickers who attempt difficult field goals.
4. I believe that the best kicker would be someone who had both a strong adjusted FG% and actual FG%. With this in mind, Janikowski is a great option, as he finished near the top in both metrics in 2011. He also had an exceptionally strong leg that opened up more scoring opportunities for his teams. Rob Bironas, Connor Barth, Matt Bryant, and Josh Scobee were also very strong in both of these metrics.
5. For other data I wish I had, I would want as much context as possible for the kicking conditions of each kick. Two great examples would be the weather during the kick and the site of the game. It is generally understood that kicking indoors is easier than kicking outside because wind and weather play less of a factor. Additionally, kicking outside in extreme weather like wind and rain is harder than kicking outside in normal conditions. More data can also be given for the true kicking distance of each field goal attempted. The metrics we have calculated consider all field goals within a 10-yard range to be of equal difficulty. For example, a 21-yard field goal is considered just as tough as a 29-yard field goal in the current calculation of adjusted FG%, which is obviously not true. Further subclasses of distance can be created to solve this problem. Therefore, the metrics we have calculated don't accurately account for certain kicking factors.

Binomial Distribution for Nick Anderson

1. From both methods, the probability of Nick Anderson missing 4 free throws in a row was 1.2%.
2. The big assumption that we are making with the binomial distribution is that each free throw is supposedly independent from the next free throw. We know this is not necessarily true, as there is somewhat of a connection between consecutive free throws. For example, if you miss the first of two free throws, your probability of making the next free throw may change. On top of this, the 67% free throw mark for Anderson comes from a much larger sample of shots. These four free throws occurred at the end of a game in the NBA Finals. The pressure was at an all time high, and Anderson may have been extremely tired from the rest of the game. If these factors affected any of the four free throws, then we are making assumptions by using the binomial distribution.

Normal Distribution for Star NFL Receivers

1. I believe that Calvin Johnson in 2012 had the best single season of the bunch. Every receiver

in the data set had a phenomenal season, but Johnson's was the best when adjusting for era through the Z-score calculation. Additionally, Calvin Johnson was great at other receiving statistics outside of just yards: receptions, touchdowns, blocking downfield, etc.

2. The Z-score does tell us more than the raw yardage totals because it shows how elite a season is relative to the rest of the league. If the average (and standard deviation) receiving yards in a season is really high, then more receiving yards are needed to have an exceptional Z-score. This is directly tied to the era that a receiver plays in. Take John Jefferson and Otis Taylor from the data set as an example. Jefferson had over 200 more receiving yards in his season than Taylor. However, Taylor finished with a higher Z-score because the 1971 average receiving yards and standard deviation were lower than those same metrics from the 1980 season. This is because the 1980 season was more pass-heavy than the 1971 season.
3. There is a lot of additional data I would want to have. First, I would want to know the quarterback that is throwing the football for each of these receivers. Calvin Johnson and Marvin Harrison were fortunate enough to have Matthew Stafford and Peyton Manning throwing them the football during these seasons, which certainly helped them accumulate receiving yards. Second, I would want to know the skill level of the defenses of the teams that these receivers play for. If a receiver's team has a bad defense, then they will most likely be passing a lot to keep up, which helps a receiver boost their stats. For example, the 2012 Lions had a bad defense, which forced Detroit's offense to be passing a lot. I'd also want to know the skill level of the opposing defenses that these receivers face. If a receiver faces a lot of bad defenses, then it is easier to accumulate receiving yards.

Standard Error for Kevin Durant

1. According to the data, Kevin Durant had a <u>0.05% chance</u> of scoring at least 30 points per game during the 2011-2012 NBA season. This is the probability of averaging 30 or more for the entire season. If we were looking at 30 points or more for just one game, the probability would be much higher.
2. The key assumption that we are making is that the fact we are saying the 2011-2012 can be predicted using the 2010-2011 season. In reality, too many factors are changing between these seasons that make it difficult to predict the second season. For example, the Thunder's roster and schedule both changed from the first season to the second season. Some players on the team got better and some got worse. Additionally, Durant's role itself may have changed within the Thunder offense.
3. Changing the sample size makes our standard error smaller or larger, depending on the direction of the change. If we increase the sample size, the denominator in the calculation of standard error increases. Therefore, an increase in sample size results in a decrease in standard error. If we decrease the sample size, the denominator for standard error becomes smaller, which makes the standard error larger. Therefore, sample size and standard error have an <u>inverse relationship</u>.
4. Counting stats, especially points, vary the most game to game and season to season. This is the case for games because there is so much variability in a given player's performance for a game. For some games, a player might have the hot hand and score a lot of points. For others, they may struggle with their shooting and take less shots, which results in fewer points. This high variability can be seen in Durant's 2010-2011 season, where the standard deviation of points was 6.84. This is the case for seasons because a player's role and opportunity to perform change depending on roster transactions, coaching changes, scheme changes, etc. These factors would inform my decision-making as a GM because I would value players who can succeed in multiple facets of the game. Because a player's role might change all the time, I would value players who have proven that they can fill the stat sheet across the board. Instead of just a strong scorer, I would look for players who can score,

assist, rebound, steal the ball, block shots, etc.