Quantifying Luck in Fantasy Football

Haris Serdarevic

April 4, 2024

Contents

1 Introduction

Every fantasy football player begins the year believing this is their year. This is typically the case for the majority of every league after draft night. These feelings remain strong until one hits week one and has to sit back and watch as their opponent has the luckiest performance possible, scoring vastly more than anyone else in the league. Such managers tend to ask questions like "why does everyone go off against me?" and other meaningless questions. This pain is only exacerbated by the other players in the league reminding them that they should simply "play better defense."

A few weeks will pass and it will be apparent which teams are strong and which are weak. This is typically judged by which teams have the highest points for or average points per week. However, when considering the standings and teams records, the same teams are not always represented. In almost every league, there will be the player who is undefeated yet has a points for in the bottom five of the league, and another player who struggles to win who leads the league in points for.

The truth is, there is a lot of luck when it comes to fantasy football. As a team manager, you only control how your team could score that week, with absolutely no bearing on how your opponent will perform. The most you can do is try to maximize the points you score every week, and hope that it will be enough to beat your opponent. But some weeks, you may have a top three performance only to lose because your opponent set the league record for points in the week. That is simply unlucky, as you would have beat just about any other performance in the league that week. But you take that one unlucky loss to the chin and keep moving forward. Just for it to happen again next week, and continue to the end of the season where you feel like you have not caught a single lucky break, while there are many teams ahead of you which you feel have had some fortunate wins. It would be so nice to have a way to quantify just how unlucky you have been, and how lucky those ahead of you have been. This is what we will do in this project.

2 Methodology

To quantify luck, we will have to make some decisions on what is considered unlucky and what is lucky. Most people would agree that you would be considered unlucky if you performed well, yet received a poor result. You would be lucky if you performed poorly yet were rewarded with a victory. But there are also different levels of luck.

How unlucky you may feel depends on the severity of your loss. If you scored a little better than average yet still lost, you may feel like you got a little bit unlucky with your result. Especially considering you may have won almost all other matchups with that score. However, if you had a great week and still lost, it would make you even more sick to your stomach you receive the same result you would have if your team only scored 70 points. The same applies to getting lucky, if you barely crest 100 but squeeze out a win it would be considered a little lucky, while scoring 85 and winning is a very lucky result.

Thinking like this, we will have to find ways to define exactly what it means to having a good vs a great vs a poor performance. Then we would have to find ways to quantify each level of luck.

We have decided upon three pieces that will contribute to the luck score. They are:

(1) Opponent Luck Score

(2) Your Luck Score

(3) Close Game Luck Score

These three will be defined below in their own sections. The overall luck score comes from summing over (1), (2), and (3). Each of the three scores are determined for every team in every matchup. This is not calculated as a season whole. The overall luck score that is published every week is all the previous individual luck scores summed with the newest luck score. So it is a running total.

2.1 Opponent Luck Score

This score is the biggest weight when it comes to determining the overall luck a team had in a given week. Calculating this score comes from considering only what your opponent did in that week. Since we as players control which lineups we put out, our team performing well or poorly can to some degree be traced back to decisions we had control over (I know injuries negate what I said here, but these are way too difficult to quantify at the moment).

However we do not control who our opponent places in their lineup, or who we play whatsoever. There truly is no way to play defense in fantasy football, so we are always at the will of our opponents having a nuclear week. Because of this, we consider our opponents score as a part of our luck.

For the rest of this section, disregard your own score. It does not matter if you scored 50 or 150 when we calculate the Opponent Luck Score. If our opponent were to have a great performance, we would consider ourselves unlucky, while if our opponent had a poor performance, we would consider ourselves lucky. But how do we determine what a "good", "great", "poor" performance is?

To make this determination, we consider all the individual scores as a whole. Here is a plot of all the scores to this point in the season (week 5).

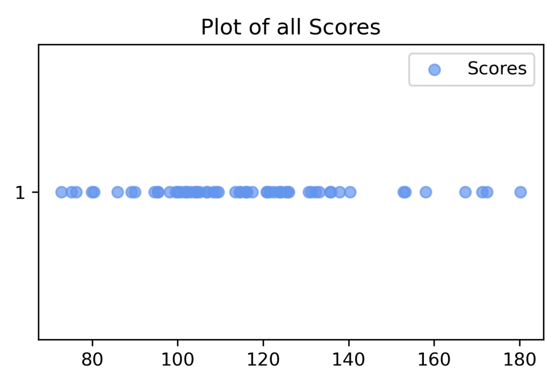

Figure 1

Now we can plot the mean (average) score to see what the most "standard" score is in our league. A naive approach would then be to say that if your opponent scores higher than average, you were unlucky and if below average you were lucky. Say the average were 115, and your opponent had 118 in a week. This difference in score is rather negligible, so we should not consider that an unlucky instance. However if your opponent had 140, that is considerably greater than the mean so it would be unlucky.

A way we can define a intervals from the mean score is by finding the standard deviation of the scores. In plain words, the standard deviation is the "average" distance each score is away from the actual mean of the scores. Typically, we have seen standard deviations of 12 in our data this season, meaning that on average the distance any given score will be from the mean is about 12 points above or below.[1]

[1] This is just attempting to define what the standard deviation is, do not critique the lack of rigor as I am not a statistician.

The reason this approach using standard deviations works is due to the fact that our data follows a normal distribution. In a normal distribution, the deviations tells the probabilities of where our data points belong. Here is a histogram of our data.

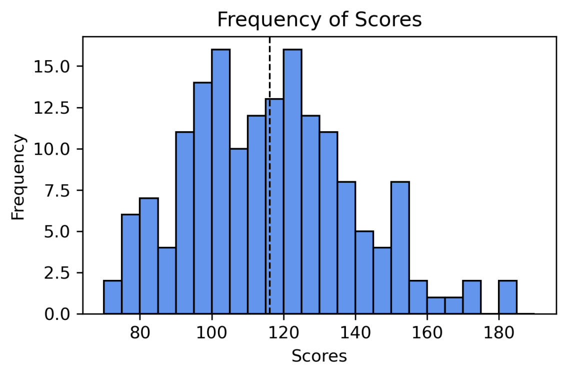

Figure 2

Now here is an example of a normal distribution, and how the standard deviations impact the probability of a score being placed.

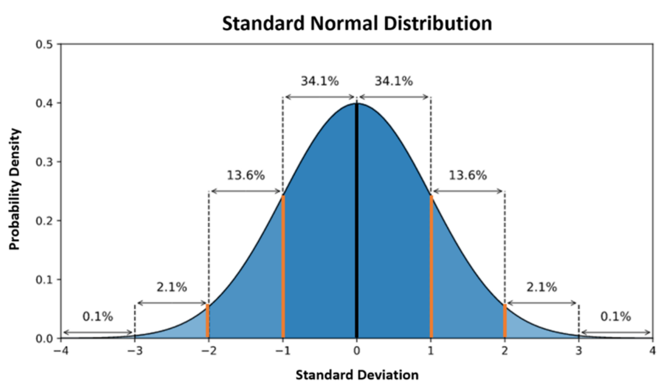

Figure 3

Here we can see that in a normal distribution of data, roughly 68% of the data will fall into the window within one standard deviation. It is important to note that in the histogram of scores we have now, the normal distribution is not quite visible. This is because we do not yet have enough data to accurately display that our data follows the normal distribution. As the season progresses it should become more visible.

Once we calculate the standard deviation, we can then plot lines which represent one, two, and three standard deviations from the mean to construct intervals or "bins" of scores. We do this in the plot here.

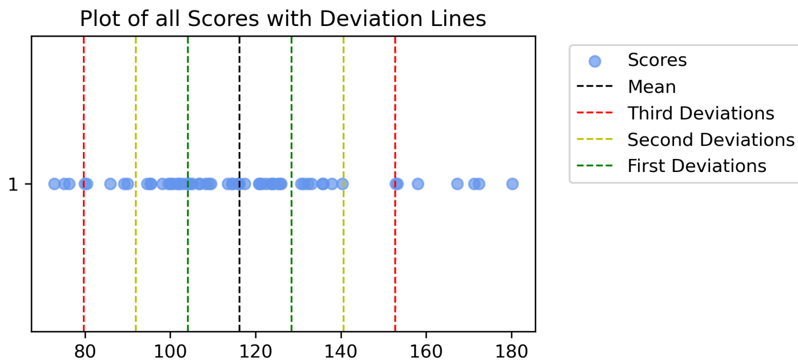

Figure 4 Scores with deviations

Here we have placed the mean line in black, this is the overall average score. Then the two green lines represent being one standard deviation above or below the mean. The yellow lines represent two standard deviations away, and red three deviations away.

Finally, we will give the function which defines how we calculate the Opponent Luck Score. Let S be the set of all scores logged this season. Let μ be the mean of S and σ be the standard deviation. Here we abstract the numbers because they change every week, but think of $\mu = 115$ and $\sigma = 12$ for a rough example. Then we can define a function $\Omega : S \to \{-3, \dots, 3\}$ defined by:

$$\Omega(x) = \begin{cases} -3 & \text{if } x \geq \mu + 3\sigma \\ -2 & \text{if } \mu + 2\sigma \leq x < \mu + 3\sigma \\ -1 & \text{if } \mu + \sigma < x < \mu + 2\sigma \\ 0 & \text{if } \mu - \sigma \leq x \leq \mu + \sigma \\ 1 & \text{if } \mu - 2\sigma < x < \mu - \sigma \\ 2 & \text{if } \mu - 3\sigma < x \leq \mu - 2\sigma \\ 3 & \text{if } x \leq \mu - 3\sigma \end{cases}$$

This defines a piecewise step function. Ignoring the troubling notation, plug in the numbers $\mu = 115$ and $\sigma = 12$, or look back at Figure 1 and imagine that between the two green lines, the function will output 0, between a yellow and green line, it will output -1 or 1, between yellow and red lines it outputs -2 or 2, and to the outside of the red lines it will output -3 or 3. This is how we determine the Opponent Luck Score.

2.2 Your Luck Score

Now that we have discussed the most difficult score, we can relax a bit and define the other two scores. This score depends a little bit on information we considered above for the Opponent Luck Score. This time however, only consider your score and your result. Do not look at how your opponent has done.

We may feel unlucky or lucky depending on the result we had based on how many points we scored. Consider the example of a situation where we score 150 points. In most weeks, one would expect to easily win with a score of 150. However, there is always the outlier chance that your opponent has an even better performance, and you end up losing despite a great score. By the Opponent Luck Score, you will already be compensated some negative luck points for getting unlucky, but you should also receive some more points because of how you performed as well. First we will show the function for Your Luck Score and then we will describe how it works. Let $S \times R$ be the set of all scores along with their corresponding result (win/loss). Points in $S \times R$ can be written $(x, 0)$ for a loss with x points or $(x, 1)$ for a win with x points. Let $\Gamma : S \times R \to \{-3, \ldots, 3\}$ be defined by:

$$\Gamma(x,0) = \begin{cases} -3 & \text{if } x \geq \mu + 3\sigma \\ -2 & \text{if } \mu + 2\sigma \leq x < \mu + 3\sigma \\ -1 & \text{if } \mu + \sigma < x < \mu + 2\sigma \\ 0 & \text{if } x \leq \mu + \sigma \end{cases}$$

$$\Gamma(x,1) = \begin{cases} 0 & \text{if } x \geq \mu - \sigma \\ 1 & \text{if } \mu - 2\sigma < x < \mu - \sigma \\ 2 & \text{if } \mu - 3\sigma < x \leq \mu - 2\sigma \\ 3 & \text{if } x \leq \mu - 3\sigma. \end{cases}$$

The first of the two equations deals with if you lost in that given week with a score of x. Referencing Figure 1, if your score belonged anywhere below the rightmost green line, you would receive a score of 0. If your score was between the green and yellow, it would be -1, between yellow and red -2, and beyond red -3. This means that the higher you score and still lose, the unluckier you are.

The second equation is the case when you win with a score of x, and this time if your score was higher than the leftmost green line you receive a score of 0. If your score is between the green and yellow it is 1, between yellow and red it is 2, and beyond the red 3. These scores mean that if you scored less and less points yet still won, you were very lucky.

2.3 Close Game Luck Score

The final piece of calculating luck score comes from the Close Game Luck Score. This is the simplest of all three pieces as it has the easiest equation. We have all had weeks where we lost by a tiny margin, and felt like we got a little unlucky. Regardless of how many points you scored, losing by a tiny amount is an unlucky part of that week. So we quantify this luck score by adding a point of luck if you lose by less than three points, or subtract a luck point if you win by less than three. Let A be the set of scores and B be the set of corresponding opponents scores. In other words, a point $(a, b) \in A \times B$ represents a matchup where you scored a points and your opponent scored b points. The equation for this would be $\Phi : A \times B \to \{-1, 0, 1\}$ defined by:

$$\Phi(a, b) = \begin{cases} -1 & \text{if } 0 < b - a < 3 \\ 1 & \text{if } 0 < a - b < 3 \\ 0 & \text{otherwise.} \end{cases}$$

Now that we have all three pieces of the main Luck Score, we can officially define how we compute the overall luck score for a player.

For any given player and a week, we calculate the luck score by summing $\Omega(x) + \Gamma(x, r) + \Phi(a, b)$, where Ω, Γ, Φ are all defined as above.

3 Arbitrary Decisions and Plans for the Future

Throughout this project there have been some decisions made which may appear arbitrary. The first decision made was when we introduced the standard deviations in the methodology. We actually are not taking the full standard deviation, as that typically is about 24. This number tends to be too large, so the intervals we consider are not as interesting when computing the luck scores. So we decided to take 50% of the actual standard deviation as our σ. This allows us to have more interesting values when it comes to the luck scores.

Another decision made was on quantifying the luck by assigning integers from -3 to 3, and by also only considering three deviations away from the mean. There is room to attempt using more or fewer than three deviations. As well as using differing values for the luck scores rather than increasing linearly.

An approach considering a continuous function was considered, and may be applied in the future. The considered approach was using the function:

$$f(x) = \pm \frac{1}{625}(x - \mu)^2,$$

where μ is the mean previously defined and x is a score. This function over accounted for unlucky scores as we have experienced outliers in the high range of scores, while not having any in the lower range. Other attempts were given with higher order polynomials, but they grew too quickly at the extrema while a linear function treated values too high within the original deviation window. These alternative luck functions will continue to be studied.

An alternative method of calculating luck scores could come from comparing historically how likely you were to win with a certain score.

4 Revisions to Formulas

After the initial definitions of Ω, Γ, and Φ, it was decided that these functions should be defined by closed formulas rather than long piecewise rules. Beginning with the Opponent Luck Score, Ω, it was possible to redefine the function by only incorporating μ, x, σ, and the floor function:

$$\zeta(x) = \frac{\mu - x}{|\mu - x|} \left\lfloor \frac{|\mu - x|}{\sigma} \right\rfloor.$$

This closed formula makes it much easier to plug in any value of x (which is the score your opponent had in a given matchup) and output the Opponent Luck Score generated.

To redefine the Your Luck Score function Γ, a little more work was necessary. First the following function

$$\xi(x) = \left\lfloor \frac{\mu + x}{2\mu} \right\rfloor$$

was necessary to produce the value of 0 when our score was less than average, and 1 when our score was more than average. It is used to redefine Γ as such:

$$\psi(x, y) = \left\lfloor \frac{|\mu - x|}{\sigma} \right\rfloor \Big(y\xi(2\mu - x) - (1 - y)\xi(x) \Big).$$

While this is a considerable improvement upon the extended piecewise function, the closed formula for Γ can be tricky to read. Evaluating $\xi(2\mu - x)$ and replacing it in the equation would be one way to reduce some of the clutter in this new version. The new definitions for Ω and Γ allow for, in theory, infinitely many values rather than being limited to the set $\{-3, \ldots, 3\}$. It would be possible to remove the floor functions in each, which in turn would convert the functions from being discrete-valued. There is argument to be made that luck scores should be more of a continuous output. Currently research and development is dedicated to this.

5 Statistical Tests

With the first season of statistics tracked and all luck scores calculated, there may arise questions about what certain luck scores mean. Without proper contextualizations, a luck score of 10 over the course of a fantasy season does not have much meaning more than someone being considered very lucky during a single season. However with context provided we may see that a luck score of 10 is a very common occurrence over the course of a season. Or we could see that luck scores of 10 are relatively rare, in which the significance of that season could be more understood.

Due to the fact that the first season has completed, with only 168 scores to use as data, another approach will be necessary to begin contextualizing luck scores. The approach will be creating simulations of fantasy season and computing luck scores from these created situations. With a simulation model, it will be possible to enumerate over a large sample size to see what trends arise with the luck scores. During these simulations, we may also track other trends such as the number of wins, average points scored, etc. The pure methodology of constructing these simulations will be described in the following section. After the methodology, we will begin to compute different statistics using the samples found in the simulations. All the assumptions made will be stated in the methodology and results sections. To finish this chapter, other hypotheses will be made in regards to expanding to new equations of luck scores.

5.1 Methodology

As described in the introduction of the chapter, simulations of fantasy seasons were created to collect samples of luck scores and other statistical features. This process began with noticing from the data gathered in the previous season that scores followed a normal distribution as in Figure 2. The notation μ and σ will continue to be used to model the mean and standard deviation of all observed scores. Note that now σ will represent the full standard deviation, where earlier we defined σ to be half of the standard deviation.

Another observation may be made from Figure 2 in that there are considerably more outliers greater than μ than there are ones less than μ. Concretely, there were no scores below 70 while there were over 10 greater than 160. While it is possible that there was simply not enough data to fill in the distribution correctly, it can be noted from other fantasy seasons that scores below 70 are much rarer than scores above 160. Therefore we will make the assumption that there is a slight right skew to the data. To account for this, we introduce a skew function that corrects values of the sample below 70 to values near 70. This will be described more concretely below.

This right skewed normal sample will makeup all the scores that took place during a simulated season. Another assumption being made here is that no individual team is being given preference in their 14 scores. Instead all 168 scores are selected randomly in the sample and then randomly applied to each team's 14 scores. In practice, this has still allowed for some teams to be considerably better than others, which achieves both goals of having the overall distribution of scores being normal while also having some teams doing better than others.

The next step in performing the simulation is creating a schedule that matches the ESPN Fantasy Football system. Each season begins with a swiss format, where every team plays all other teams (11 total matchups with 12 teams) prior to repeating any matchups. Constructing the initial part was done using a method referenced as the "circle method" in [AL]

5.2 Results

References

[AL] Scott Anson and Sam Lester. Sports scheduling: Algorithms and applications.