Analysis of Convention Betting Strategies in Sports Betting

Douglas Silverman, Calvin Lok, Jason Le *University of Amherst*

Over the course of the past 10 years fantasy sports have seen an explosion in popularity and interest. People participate in various forms of fantasy sports gambling ranging from a single $10 league with friends, to multiple money leagues, all the way to multiple wagers on draftkings and other one-day-Sunday fantasy teams. The recent popularity of websites like draftkings and fanduel are a way for Americans to feel the thrill of gambling. Sports gambling in the United States was banned in 1992 by the PASPA (Professional and Amateur Sports Protection Act of 1992). This made it difficult for states, particularly New Jersey, to have some sort of sports betting. However, in May of 2018, the Supreme Court overturned PASPA allowing states to pass their own legislation for sports betting. Since the decision, 17 states have fully legalized sports betting and 48 states have some sort of legislation. As more states trend to legalize sports betting, the analysis of different models or betting strategies is interesting to look at.

Traditional Betting Models

For our project, we started by looking at conventional betting models. These models are strategies many casual gamblers use to minimize losses. The concept behind this is known as the gambler's fallacy. An example of this is tossing a fair coin. Each toss is statistically independent and the result of the toss does not depend on the previous tosses. Let's say we toss a coin 10 times. If the first 9 tosses are heads, many will believe the next toss must land tails. This is known as the gambler's fallacy. This belief is wrong however. Let the event H_i represent the ith toss lands hands. The probability of 10 heads in a row is what most people look at:

$$P(10\ heads) = P(\bigcap_{i=1}^{10} H_i) = \prod_{i=1}^{10} P(H_i) = \frac{1}{2^{10}}$$

However, the probability we are looking at is "the 10th toss being heads given the first 9 are heads." This probability can be described as:

$$P(H_{10} \mid H_1, H_2, ..., H_9) = P(H_{10} = \prod_{i=1}^{9} P(H_i)) = H_{10} = \frac{1}{2}$$

The Probability is $\frac{1}{2}$ because every toss is an independent event. Another way of looking at this is looking at the probability of 9 heads in a row and then another heads:

$$P(H_1, H_2, ...H_9, H_{10}) = \prod_{i=1}^{9} P(H_i) \cdot P(H_{10}) = \frac{1}{2^9} \cdot \frac{1}{2} = \frac{1}{2^{10}}$$

This Probability is also equal to having 9 heads in a row and then a tails:

$$P(H_1, H_2, ...H_9, H_{10}) = \prod_{i=1}^{9} P(H_i) \cdot P(T_{10}) = \frac{1}{2^9} \cdot \frac{1}{2} = \frac{1}{2^{10}}$$

This shows how the gambler's fallacy is incorrect. The most infamous case of the gambler's fallacy occurred in August of 1913 during a roulette game at the Monte Carlo Casino. The roulette ball landed on black 26 times in a row. The probability of this happening was : $(\frac{18}{37})^{26-1}$ or around 1 in 66 million

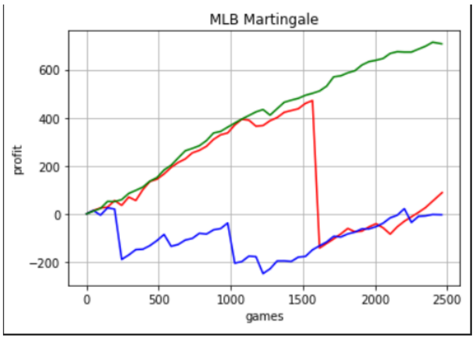
Martingale Betting Algorithm

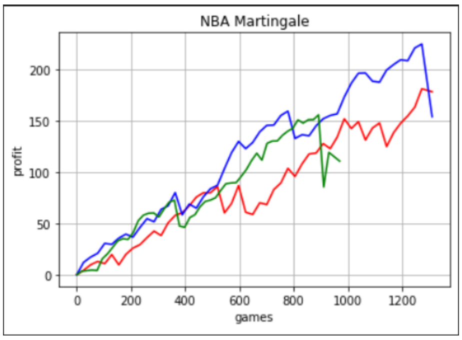
The first model is based on a martingale of the initial capital. A martingale is a sequence of random variables such that the next value in the sequence has an expected value equal to the present value. The formal definition of discrete-time martingale (i.e. a discrete-time stochastic process, meaning a sequence of random variable $X_1, X_2, X_3, ...$) that satisfies:

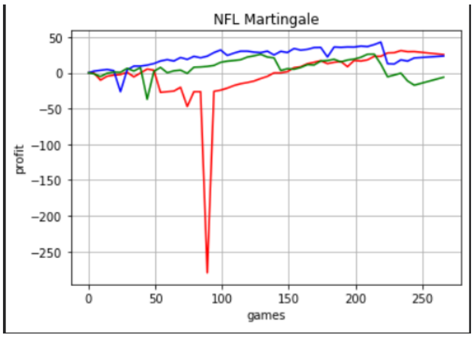
$$\forall n : E[|X_n|] < 0 \land E[X_{n+1} \mid X_1, X_2, ..., X_n] = X_n$$

Our model uses a martingale based on the initial loss value, which starts at 0. We start with an initial bet size so that losing 10 bets in a row would bankrupt us (this value can be adjusted). After a win, the bet size stays the same and the loss remains the same. After a loss, the next bet size will be equal to the amount needed to wager so that the entire loss will be made up. If loss is 0 then the bet size will be the initial bet size. As consecutive losses occur, the bet size grows exponentially. This shows how gamblers can amass debt very quickly.

Martingale Performance







Expected Value for a single round of gambling:

A single round can be defined as a stochastic process (a sequence of random variables) such that all discrete events are losses. After a set of losses, the first win happens the round is over because the gambler has "reset". The next round of gambling occurs and the process continues. A continuous sequence of martingale bets can be split into independent rounds. let

q = the probability of losing a given bet (For American roulette, $q = \frac{20}{38}$ for a bet on red or black)

b = the initial bet size

n = number of bets the gambler can afford to lose before being bankrupt. n must be a positive integer ($n \in +Z$).

The probability that all n bets will be lost is q^n

$$P(n\ losses\ in\ a\ row) = \prod_{i=1}^{n} P(loss) = \prod_{i=1}^{n} q = q^n$$

The total amount lost in this case is:

$$\sum_{i=1}^{n} b \cdot 2^{i-1} = b \cdot (2^n - 1)$$

The probability of a not losing all bets is:

$$P(not\ losing\ all\ bets) = 1 - q^n$$

Therefore the expected profit is:

$$b \cdot (1 - q^n) - q^n \cdot b(2^n - 1) = b(1 - (2q)^n)$$

Notice that whenever $q > \frac{1}{2}$ the expected value of the round will be negative, for all $n > 0$. This is the reason why the house always wins for any casino game. This means that for any wager where the gambler is more likely to lose than win, the gambler is expected to lose money using a martingale strategy. Increasing the wager per round only increases the average loss rather than making up the difference.

Oscar's Grind

Oscar's Grind is a betting strategy first documented in 1965 in The Casino Gambler's Guide, and is designed to minimize risks for steady profits. It applies to all even money bets, and the only goal is to win one unit of profit. Each bet is considered a series that is continued until one unit is won, then reset to begin a new cycle. Initially, one unit is bet. If this is a win, the series ends and a new one begins. If it is a loss, the bet stays at the same size and you continue betting. When the bet is a win and the profit threshold has not been met, the successive bet is increased by one unit. These two rules are repeated until the series ends in one unit of profit, then reset to a bet of one unit for the next series. Given infinite time and infinite wagers, Oscar's Grind always return a profit.

Kelly Criterion

The Kelly Criterion is a simple formula to assist gamblers in deciding how much money each bet should receive. Each bet is a fraction of the current amount of assets, scaling after each bet. The equation is :

$$K\% = \frac{ap - q}{a}$$

Where

K = the fraction of the bankroll to bet

a = the decimal odds of a bet – 1

p = the probability of winning

q = the probability of losing

The result is a percentage of your capital that is recommended to bet on the gamble. If this is negative then it is a sign to avoid the bet and maybe consider betting on the other option.

Datasets for the Analysis

In order to analyze these different models we have obtained 9 datasets from (website name). We obtained data for MLB, NFL, and NBA professional games. The benefit of using multiple sports for analysis is that there is variation. Each sport varies in the length of the game, the length of the season, and the likelihood for an upset (i.e. a team with a better record losing to another team). For each sport, we looked at every game over the past 3 seasons. Each game included: the date of the game, the 2 teams playing, the scores after each quarter of play (or inning for MLB), the final score, and the american odds [1] for each team winning. For this analysis, we required a bet being placed on every game.

(Here is the github) https://github.com/Douglas-Silverman/Math-456-Final-Project

Poisson Based Prediction

The Poisson Distribution is a discrete probability distribution that can be used to express the probability of certain events from happening when known how often the event has occurred. In terms of sports, it can be used to determine the probability of the number of points a team scores in a game. With this method, we can calculate the probability that a favored team wins against a certain team. Below is the poisson distribution formula:

$$P(X) = \frac{\lambda^x e^{-\lambda}}{X!}$$

[1] American odds refer to the payout for winning the bet. A positive number like 250 means that on a bet of $100 you will win $250 if you win. A negative number refers to the amount needed to place on a bet in order to win $100. For example, if the odds are -115, then a bet of $115 would yield $100 if won.

For X is the desired number of points a team will score and λ is the expected number of points a team scores. Before we can use the poisson distribution to calculate winning odds, we need to find the value of λ. We can find this value by calculating the attack and defence strengths of a home team and an away team. Attack and defence strengths are based on the number of points a team has scored and the number of points they have conceded in a given season respectively. We must also know the average number of points scored per game for both at home and away teams. To find the averages we will use these equations:

$$\text{Average points scored at home} = \frac{\text{total points scored by home in the season}}{\text{total number of home games}}$$

$$\text{Average points scored away} = \frac{\text{total points scored by away in the season}}{\text{total number of away games}}$$

We will also need to find the average number of points conceded by both the home and away team. However, these are just the opposite to the average points scored per game. The average number of points conceded by the home team would be equal to the average points scored away and the average number of points conceded by the away team would be equal to the average points scored at home.

Next we will calculate the home team's attack strength which is needed to calculate the expected value of points for the home team. We will use these two equations below:

$$\text{Home team's average score per home game} = \frac{\text{Points scored at home by the home team}}{\text{number of home games played}}$$

$$\text{Home team's attack strength} = \frac{\text{Home team's average score per home game}}{\text{average points scored at home}}$$

With this, we will find the home team's attack strength for the given season. The same can almost be done when calculating the away team's defensive strength:

$$\text{Away team's average points conceded per away game} = \frac{\text{Number of points conceded away last season by the away team}}{\text{number of away games played}}$$

$$\text{Away team's defense strength} = \frac{\text{Away team's average points conceded per away game}}{\text{average points conceded by away team}}$$

With this, we now have everything needed to calculate the expected points a team will score. To find the expected value of the home team, multiply the home team's attack strength with the away team's defense strength and the average number of home games. Now we plug this value into λ and we can find the probability that the home team will score X points against the away team. We can also find the away team's expected points by simply flipping all the home and away variables.

Based on the sport, X should be a reasonable value that an average team could achieve. For example, in soccer, there is no reason to find the probability that a team will score above 5 goals because it is highly unlikely to score that many points in the sport. Once X has been established, We should get a list of probabilities for the desired range of scores for both teams. If we want to find the probability of a specific final score of a game for example, an outcome of 2-0 home vs away, simply multiply the respective probabilities together to obtain the probability of the outcome. To find the probability that the home team wins however, we will have to find the probability of every possible outcome where the home team wins and add them all together.