
ISyE 2028 – Basic Statistical Methods - Fall 2015

Bonus Project: “Big” Data Analytics

Proposal (or Final Report)

Jonathan Pang

Sports betting, although controversial, is an extremely popular form of gambling in the United States. Although many types of bets can be placed on popular websites or in casinos, the most popular type of bet that is placed on the results of major American sports is the spread bet.

A spread bet is a wager that the difference between the scores of two teams will be less than or greater than a certain value, which is specified by the bookmaker. For example, if the spread betting line for the Detroit Lions at Kansas City Chiefs game in Week 8 is +3.5 for the Detroit Lions, a spread bet placed on the Detroit Lion would be won if the score for the Detroit Lions (D) + the spread betting line (+3.5) is greater than the score for the Kansas City Chiefs (K).

$$D + 3.5 > K$$

If the score for the Detroit Lions (D) + the spread betting line (+3.5) is less than the score for the Kansas City Chiefs, the wager will be lost.

$$D + 3.5 < K$$

In other words, the bettors are wagering on whether or not the Kansas City Chiefs will outscore the Detroit Lions by +3.5 points. It is also noted that the scores sports are only represented by integers. Thus, oftentimes, the bookmakers will set the betting line at half a point intervals, making it impossible for neither side to win the bet. Otherwise, if the Kansas City Chiefs outscored the Detroit Lions by exactly the spread betting line, a push would occur and neither side would win (or lose).

In order to maximize the odds of winning a spread bet, historical data can be utilized to determine patterns in the final spread and the actual results in games. In this study, trends will be analyzed to determine whether or not any teams in the NFL are consistently able to “beat the spread”. This will be answered by hypothesis testing, comparing the spread betting lines and score differentials for each game a team plays. If the difference is significant, it is thus a safer bet to wager money on the team. Further analyses of different conditions (home or away, opponent) may also be analyzed to determine whether a bet is considered safer. However, it is expected that there is no significant difference between spread betting lines and score differentials for each game a team plays

An aggregate of historical and current sports betting information can be found at the Odds Sharks website, from which the final point spreads and over under betting lines are recorded for professional sports such as the NFL (National Football League), MLB (Major League Baseball), NBA (National Basketball Association), and NHL (National Hockey League) as well as college sports such as NCAAF (National Collegiate Athletic Association Football) and NCAAB (National

Collegiate Athletic Association Basketball). A web scraper will be utilized in order to extract past sports betting lines as well as actual scores of games.