Money Lines and Moneyball: An Intersection of Betting and Analytics

Miles Kee
University of Connecticut, Storrs, Connecticut, USA

Abbreviated abstract: The number one thing a sports bettor tries to do is to give themselves an “edge” over the sportsbook by having information that signals that a line is not accurate. For example, if you believe a team has a 50% chance of winning a game and they are listed by the sportsbook as having a 45% chance to win, your edge is 5%. The rise of both the use and availability of advanced statistics has opened new avenues to creating an edge, as these stats often have more predictive value. I used some of these advanced statistics to create a betting model for Major League Baseball to see if there was any betting value in their use.

The Model: Data

Projecting Hitting

- The advanced stat that gives the most insight to future run creation is xwOBA. wOBA is highly correlated with runs created, and xwOBA is much more predictive than wOBA since it uses Statcast data of what was expected to happen
- I used the average xwOBA of each team's daily lineup to project the number of runs that lineup would create over a full season

Projecting Pitching

- There are many advanced pitching stats that have some predictive value, my favorite is SIERA since I think it does the best job of isolating pitcher skill
- I standardized each starting pitcher and team bullpen's SIERA relative to the league average SIERA to create a "run factor" (.85 run factor means they would allow 15% fewer runs than an average pitcher)
- Using a weighted average calculation for starter innings vs bullpen innings, I created a game long run factor for every team

Combining Hitting and Pitching

- To incorporate pitcher strength into the run total, I multiplied each team's projected run total by the opposing team's pitcher factor to obtain a season long run scored and allowed total for each team
- I then used Pythagorean win expectancy with exponent 1.83 to find each team's percentage of winning the game

Using the Model

Placing A Bet

- I converted each team's winning percentage into an implied money line for each team
  - For example, a 50% winning percentage converts to a money line of 100
- This became my "play to" value- the money line at which I would no longer place a bet on a team
  - If my "play to" value was -150 and a team was listed at -145, I would place a bet on them
  - If my "play to" value was -150 and a team was listed at -165, I would not place a bet on them
- **Money line: A bet on a team to win the game straight up**
- **-150: A $150 bet will win $100**
- **100: A $100 bet will win $100**
- **+150: A $100 bet will win $150**

		xWOBA	Pitcher Factor	Offense Runs	Factored Runs	Pythag Win %	Win % w Home ADV	Play To
Away	DET	0.314	1.10	709	664	36%	32%	216
Home	MIN	0.332	0.94	836	917	64%	68%	-216

Results and Improvements

Results

- Through 399 bets, the model is up 1.46 units
  - A unit is your standard bet- meaning that if you bet $100 a game, you would be up $146
- This is better than your expected profit with no model, which is 0 assuming no juice (the % of bets the sportsbook takes)
  - In baseball betting, there is 1.2% juice, so a better with no model is expected to lose .12 units per bet
  - Over the course of 399 games, a standard bettor would expect to lose 47.88 units
    - This simple calculation shows how difficult sports betting is and illustrates how “the house always wins”

Conclusion and Improvements

- While it is not up much money at all, the model has vastly outperformed a standard bettor
  - There were many ups and downs throughout the season, as is expected due to random variation
- Some improvements I would like to include for next season:
  - Baseline preseason projections
  - Better rookie forecasting (no previous MLB data for them)
  - Weighted averages for bullpen SIERA so relievers used more have a bigger impact on rating