Is the Most Valuable Player Award Actually Rewarding the Most Valuable Player?

Johnny Asel
7 November, 2018
HL Mathematics -- Mr. Snyder

Problem/Background:

Baseball has always been a game of numbers. Since its inception, statistics were created to measure how good players were. Hits, errors, total bases, batting average; all just attempts to measure how good a player is. While no single statistic could sum up how good a player was, one could look at many and get an idea of who the best players were. This was until the creation of Wins Above Replacement (WAR).

WAR is a single number that estimates the value of a player in units of wins above a replacement level player, or a player any team could quickly acquire in free agency. The basic idea of how WAR is calculated is that you take the difference in value between the player and the estimated value of a replacement level player multiplied by the amount of playing time (see below). Of course there is a lot of complication within all of that, but that is the general concept.

$$\text{Playing time} \left(\begin{array}{c} \text{Value of Player being measured per X playing time} \end{array} - \begin{array}{c} \text{Value of replacement level player per X playing time} \end{array} \right)$$

What is interesting is that the WAR formula works in estimating the value of players back to before the beginning of modern baseball. This provides a way of judging how well the voters chose the MVPs across all of baseball history.

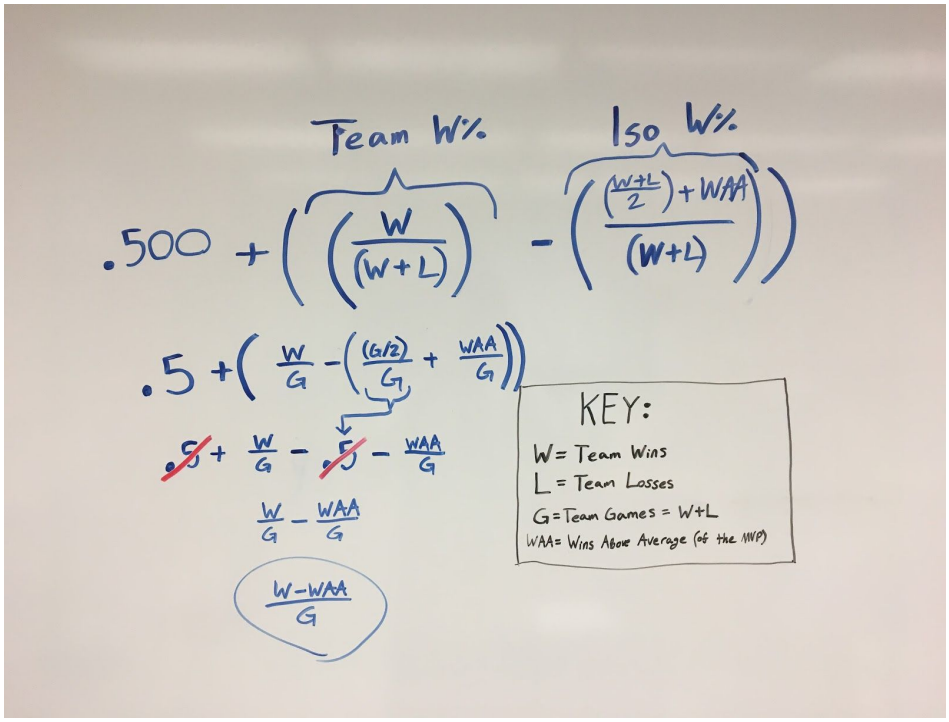
One key concept to understand is that there have been multiple forms of the MVP award. The Chalmers Award was the first form of a Most Valuable Player (MVP) Award being presented to the best players from both the American and National Leagues from 1911 to 1914.

The next form of the award came in the League Awards. The problem with this award was that player-coaches and repeat winners were ineligible on the American League side of the award. This led to outrage that Babe Ruth could not win the MVP award in recognition of his historic 60 home run season. These controversies led the award to only last from 1922 to 1929. In the 1931 season, the Baseball Writers' Association of America's Most Valuable Player Award was introduced. This stands as the modern version of the award to this day. For my following tests, I used data from all three awards counting them all as the same since all three awards had the intention of recognizing the most valuable player in the league.

In my time as a baseball fan, I have noticed two trends that I think have caused voters to not vote for the MOST valuable player. These two trends are voting for batters and players on winning teams. While there are exceptions, it seemed to me like too many MVPs followed these trends. **The goal of this project is to statistically prove (with 95% confidence) that there is unfair bias against pitchers and players on sub-par teams in MVP voting.**

Sub-par Team Bias

To test for the bias against players on sub-par teams, I had to find what every MVP's team winning percentage would've been had the MVP not been on the team. To find this winning percentage, I needed to find the difference between each MVP's team real winning percentage (TeamW%) and subtract it from what an average team's winning percentage would be with the addition of the MVP (162WL%). The resulting number indicates how far above/below .500 the MVP's team would have been without him. Therefore, adding .5 to this number would give the winning percentage we are looking for. The formula below shows the formula for this winning percentage and how it can be derived into something more simple.



If no bias existed, the average MVP-less winning percentages should average out to be about .500 and the amount over .500 should be roughly equal to the number under .500. Unfortunately, this was not the case. Not even close. The average projected winning percentage of the MVP's team was .558 with 169 over .500 and just 26 under .500 with one was exactly at .500. To put that into perspective, that is like flipping a coin 195 times and only getting 26 heads (assuming the voters are unbiased). But maybe this just happened by chance. What are the chances of this happening if the chances really were 50/50?

$$\left(\frac{195!}{(195-26)!26!} 0.5^{26}(0.5)^{195-26} + \frac{195!}{(195-25)!25!} 0.5^{25}(0.5)^{195-25} \ldots \frac{195!}{(195-0)!0!} 0.5^{0}(0.5)^{195-0} \right)$$

Using binomcdf (see above), we can calculate the that probability by plugging in .5 as the probability, 195 as the number of trials, and 26 as the number of successes. In stats class, the standard significance/confidence level to disprove something is 95%, so you want the chance of your null hypothesis (the thing you're trying to disprove) being correct to be less than 5%. In this case, the chance of the null hypothesis being true is 5.36·10^-28. So done then right? The MLB MVP award clearly has some bias that is leading the voters to choose players on winning teams. Not quite yet.

This is because most MVP-level talents have the tendency to gravitate toward teams that can win. Since the first season of free agency in 1974, players have had much better control over where they play. If a good player is tired of playing for a bad team, they can just wait for their current contract to expire, then sign with a better team. In addition, teams can trade future assets such as young players and picks for good players. This leads to MVP-caliber talent generally being on good teams.

To test the extent of good player's gravitation toward good teams, I wanted to take a sample of MVP-caliber (not necessarily winners) from seasons of which the MVP award was in existence. To do so, three conditions must be met; random, independent, and large/normal. To meet the random condition, I randomly selected the years of which I would be taking the sample by having my computer randomly choosing one by one. From each year selected, I would take the top ten players (by WAR) and put them under the same procedure that the MVPs went through to determine what their team's winning percentage would be without them. I also satisfied the independent condition by selecting nine years (containing ninety players). Since this

in less than ten percent of the desired population, it meets the independent condition. Finally, I also met the large counts condition because I selected over thirty players.

The mean winning percentage was .508 with 36 out of the 90 players being below .500. I put the results through the same binomcdf test as I did for the MVPs with 90 as the number of trials, 36 as the number of successes, and .5 and the probability. So if the average MVP-caliber player is on a .500 team (hence the .5 probability), the probability of this little amount of sub .500 teams in a 90 player data set is .036. While not as convincing as the MVP players' results, the probability is still less than .05 allowing me to reject the null hypothesis that the average MVP-caliber player is on a .500 team.

So finally, I re-ran the test for the MVPs, kind of. Instead of looking at the number of MVPs above/below .500, I found the number above/below .508. Instead of the previous 26/195 being below .500, the number increased to 30/195 being over .508. Run that through binomcdf again (trials=195, successes = 30, probability = .5) and the probability comes out to 4.37•10^-24. Since this probability is still much less than .05, I can confidently assume that there is some voting bias against MVP candidates on sub-par teams.

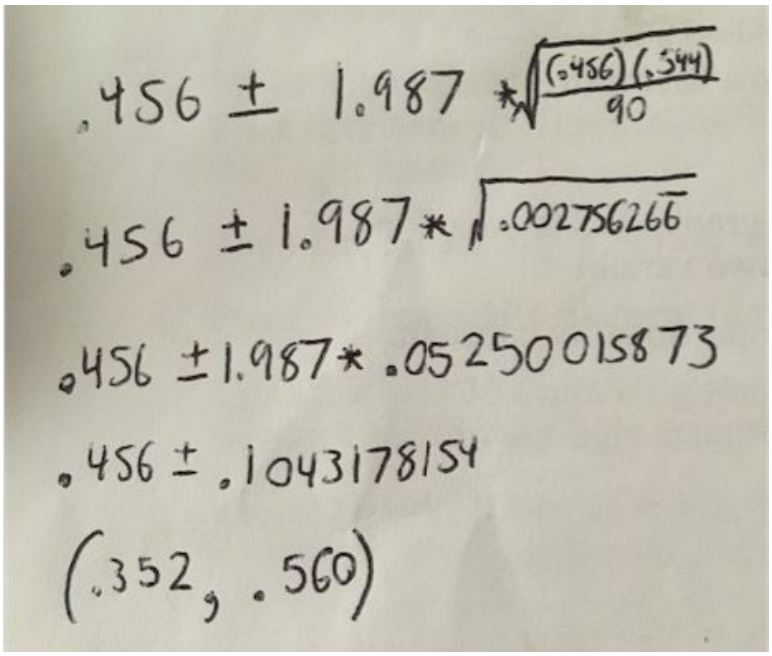
Anti-pitcher Bias

The other bias I wanted to test was the bias against pitchers. The common argument against pitchers winning the MVP award is that they only play in one fifth of their teams games. While this is true, the top pitchers tend to have similar WARs to the top hitters therefore indicating they can hold the same amount of value over the course of a season.

To test the bias against pitchers as opposed to position players, I first needed to get an idea of what proportion of MVP-caliber players were pitchers. To find this proportion, I again

randomly selected nine years of which MVP awards were awarded. I would take the top ten players from each year selected and find the proportion of pitchers within the entire sample and create a 95% confidence interval for the true proportion of pitchers amongst MVP-caliber players.

After selecting and testing, I found the sample proportion of pitchers to be .456. After I created the t-based 95% confidence interval (see below), I can be 95% confident that the true proportion of pitchers amongst MVP-caliber players to be between .352 and .560. This is interesting considering that the proportion of pitchers amongst MVP winners is .128 (25/196).



Even if I use the bottom end of the interval (.352) for the true probability of a pitcher winning the MVP award for a binomcdf test (trials=196, successes=25, probability=.352), the probability of this low number of pitchers winning MVPs is 9.93·10^-13. Since this is much less than 0.05, I can confidently reject the null hypothesis that there is no difference between the

probability of an MVP-caliber player being a pitcher and the probability of an MVP being a pitcher.

The statistical evidence is clear, across the history of the MVP voting, there has been a statistically significant bias against pitchers and players on sub-par teams. These awards help define players' legacies, hopefully as WAR becomes more widely accepted and fine tuned, the rightful winners begin to be rewarded more often, regardless of position or team.

Reference List

"MLB Stats, Scores, History, & Records." *Baseball-Reference.com*, www.baseball-reference.com/.

- All statistics taken from this website

"WAR Comparison Chart." *Baseball-Reference.com*, www.baseball-reference.com/about/war_explained_comparison.shtml.

- Used for the history of the WAR (Wins Above Replacement) statistic

Honor Pledge

On my honor, I've neither given nor received any unauthorized aid on this assignment.

-Johnny Asel