Executive Summary - Major League Baseball (MLB) Umpire Bias

Vaughn Hajra and Andrew Mayer

Background

In this analysis, we were looking to observe what environmental factors of a baseball game setting have significant impacts on umpires' performance. Very much inspired by the findings of Eric Fesselmeyer at Monmouth University who found that temperature had a significant impact on umpire accuracy, we investigate weather conditions, game duration, attendance, and more.

Methods

- *Multiple Linear Regression* - To begin our analysis, we ran multiple linear regression on our data with total run impact as the response. We found that all of the variables of interest, except for attendance, are statistically significant. In addition to the statistical significance, we find that all of the coefficients of the variables follow our intuitive expectations. This regression model has an adjusted R-squared of 0.165 and mean squared error of 0.52

- *Lasso (L1) Regression* - We are interested in Lasso regression to further investigate the importance of specific variables and also feature importance within our model. In our case, lasso regression yielded very similar results to multiple linear regression, likely because of the robust dataset. Again, we had a mean squared error of 0.52

- *Multivariate Adaptive Regression Splines* (MARS) - Multivariate Adaptive Regression Splines were used as many variables intuitively lent themselves to non-linear relationships with Total Run Impact. MARS is perfect for examining these kinds of relationships as it allows for a piecewise decomposition of our input variables. MARS also allows for easy interpretation of coefficients and variable importance. Again MARS results were similar to the other models but provided some insight into the cutoff points of our input variables.

Results & Suggestions

As previously mentioned, our models are all consistent and intuitive. We are accurately predicting total run impact with a mean squared error of 0.5 to 0.7 which is surprisingly low given the little information provided. We find that elapsed time is positively correlated with total runs impact (a longer shift will lead to lower performance) and score differential is positively correlated with total runs impact (lower intensity likely is related to lower levels of attentiveness). Colder temperatures are also associated with a higher runs impact and later start times are associated with lower runs impacted (potentially big games are broadcasted at prime time, for example Sunday Night Baseball which could have more potential scrutiny). MARS adds to all of these findings giving discrete cutoff points for some variables.

Although we are cautious to extrapolate our results regarding MLB Umpires' accuracy to other industries, the same methods could be used to assess workplace performance and improve overall efficiency. For example, how long should a lifeguard sit in the sun to maximize attentiveness? Is there a benefit in providing temperature control (extra heating or cooling) to construction workers? What intensity in an office setting yields the best results? Multiple regression, lasso regression, and MARS could be used to answer all of these questions and more.