The power hitters swing and miss and strike out. Of the hit they made and walk, at the end of the year their ORP is always going to be higher than most of the other guys on the team because they brag the bases.” - Harold Reynolds

We aim to answer 4 questions about the relationship between baserunning and batting skill:

1. Which baserunning method increases/decreases run scoring the most? (The double play)
2. Is there a case that this baserunning method differs based on a player’s batting abilities?
3. Are there differences in the change in run scoring for the same baserunning method across players with different batting abilities?
4. Are there certain players who stand out given their unique set of baserunning and batting skills?

**Model**

We define baserunning and batting abilities for each player $j$ by separating the player’s parameter space into three parts using a triple

$$y_j = (q_j, p_j, r_j),$$

where $y_j$ is a 3-dimensional simplex of batting ability, $q_j$ and $c_j$ are 10-dimensional vectors over $[0, 1]$ representing baserunning attempt and success ability, respectively. We explore the notion of a gradient for a continuous function $f$, where $f$ is a mapping of $y_j$ to run scored, through the following steps:

1. Fit the models for initial parameter estimates.
2. For $x_j$, we use nested Dirichlet distribution as a prior in batting abilities, make updates assuming that the prior batting outcomes are multinomial and estimate via maximum likelihood estimation as in (Nut, 2010).
3. For $q_j$ and $q_j$, similar nested Dirichlet priors pertaining to the attempt and success abilities for any one of the 10 baserunning events are set. The underlying data likelihoods are binomial for attempts and conditional binomial for successes (when conditioned on outcome, similar to model 1.3 as explained above).
4. Determine the current run value of a player’s total offensive abilities. We fit each player’s batting and baserunning profiles at their initial state $(y_j)^0$. We estimate $f$ at that state by performing simulations from the baserunning outcome simulator (SimuOutCome, detailed in (Baumer 2009)).
5. Calculate the batting-baserunning gradients. We examine the partial derivative of $f$ in a neighborhood of $(y_j)^0$, looking for how $f(y_j)$ changes as we make changes exclusively in $q_j$ and $c_j$. This is done by estimating $f$ on a grid of values in the neighborhood of $(y_j)^0$, via the same SimuOutCome used for the initial state.

Schematic

Key

- 9 batting events (i.e., outcomes of a plate appearance):
  - (1B, 2B, 3B, HR, BB, SO, HBP, QO, AO)
- 10 baserunning events

References