Time Through the Order Penalty: A Bayesian Analysis of Starting Pitcher Progressive Performance

CODY MCCAMMON

THE UNIVERSITY OF MISSISSIPPI

December 5, 2023

Contents

1 Introduction

At its core, baseball is a battle between pitcher and hitter. The pitcher's goal is to get the batter out, while the batter's goal is to get on base by whatever means possible.

There is a dogma that exists in clubhouses across the world: once a starter is seeing a hitter for the third time, his performance drops off considerably. It is not unreasonable to think that professional hitters improve the more times they see a pitcher; they get used to the speed, the arsenal, and the mental games the pitcher is playing. However, this *Time Through the Order Penalty* (TTOP) has been used in some of the most important moments in baseball, and it has not always been the right decision.

The most famous of this is in Game 6 of the 2020 World Series. Starter Blake Snell was pitching a remarkable game, his metrics looked good, and his count was low. Snell was only on pitch number 73 (most starters reach around 100 before they are pulled). On top of that, he had only given up two hits and one earned run. In those 73 pitches, he had struck out 9 batters. He was on a roll.

Despite this beautiful statline, as he reached the third time through the order, his manager came and pulled him from the game. The next pitcher gave up two runs fairly quickly. They lost the lead, and ultimately, lost the World Series. The entire season hinged upon this one decision, and it was the wrong one.

Why is this idea so prevalent though? This dogma goes back decades, but it was first formalized by Tango et al. in 2007. Tango quantified the improvement from hitter-to-hitter as a steady increase in *weighted on-base average* (wOBA). They found a remarkable 9 point increase in wOBA between the first time through the order and the second time through the order.

In this battle between hitter and pitcher, the roles are reversed: the pitcher seeks to minimize wOBA while the hitter seeks to maximize. A 9 point increase is a significant advantage in favor of the hitter. The goal of this paper is to model the probability of each outcome using variables that change throughout the game.

2 Method

Before the model is introduced, it is important to state why the multinomial logistic regression model is chosen. It is also necessary to define the variables which will be used within the model.

2.1 Modeling highly variable outcomes

The game of baseball is a game of rapidly shifting probabilities. Each pitch seen drastically affects the outcome of the next. Modeling these rapid shifts is inherently complex, and in order to handle this, a complex model is needed.

The multinomial logistic regression model used allows flexibility and the ability to incorporate time-dependent effects. The game of baseball has seven categorical outcomes, and the probabilities of these categorical outcomes change on a per-batter basis. It is necessary to have a model that can change along with the game. The model can also effectively represent the different skillsets of athletes (does an individual athlete get more walks, singles, homeruns, or strikeouts?)

Before the model can be introduced, it is necessary to specify the eight variables that will be used.

There are a total of $K = 7$ possible outcomes in a plate-appearance: out, unintentional walk (uBB), hit by pitch (HBP), single (1B), double (2B), triple (3B), and home run (HR). We let y_i be the categorical variable which indicates the outcome of the i^{TH} plate appearance. This can be written notationally as

$$y_i \in \{1, 2, ..., 7\} = \{\texttt{Out, uBB, HBP, 1B, 2B, 3B, HR}\}$$

We also need to introduce the *batter sequence number*, t. This number is simply a consecutive count of the batters faced by the pitcher. In this case, the number ends at 27 (the end of the 3rd time through the order). There are also two binary variables corresponding to the 2nd and 3rd TTO. They are $\mathbb{I}(t_i \in 2\text{TTO})$ and $\mathbb{I}(t_i \in 3\text{TTO})$ respectively.

It is also important to introduce the variables $x_i^{(b)}$ which represents batter quality and $x_i^{(p)}$ which represents pitcher quality. A more extensive definition can be found in §2.5 of the paper paper [?].

In addition to these, there are two more binary variables. A batter who has the opposite handedness of the pitcher has an advantage, based on the pitches a player can throw as well as the slightly different angle the ball comes towards the plate. This advantage is represented by the binary variable $\texttt{hand}_i$. On top of this, a hitter also has an advantage when batting at home compared to on the road. This advantage comes from familiarity with the background the pitcher is throwing against, the ability to follow their normal routine, as well as other intangibles that will not be elaborated here. This variable is represented by $\texttt{home}_i$.

2.2 The Model

With the preliminary information out of the way, it is now possible to introduce a new multinominal logistic regression model.

$$log \left(\frac{\mathbb{P}(y_i = k)}{\mathbb{P}(y_i = 1)} \right) = \alpha_{0k} + \alpha_{1k} t_i + \beta_{2k} \mathbb{I}(t_i \in 2\text{TTO}) + \beta_{3k} \mathbb{I}(t_i \in 3\text{TTO}) + \mathbf{x}_i^\top \eta_k$$

In addition to the eight variables, four more parameters are introduced. These are α_{0k}, α_{1k}, β_{2k}, and β_{3k}. The α terms control the evolution of the probability of each plate appearance outcome as the game progresses. This can be thought of as the *continuous* effect of a change in pitcher performance.

The β terms represent the discontinuity (once out, a batter does not hit again until the order turns over, 9 hitters later). β_{2k} and $\beta_{3k} - \beta_{3k}$ thus represent the *discontinuous* change of batter performance during the second and third times through the order.

2.3 Priors

For simplicity, the model assumes all pitchers and batters deteriorate at the same rate throughout the game. While this is not true, Brill et al. (2023) found that this does not effect the qualitative results of the study.

Greenhouse (2011) also found continuous decreases in pitch velocity as pitchers deteriorated. The longer the pitcher is in, the more likely he is to allow an outcome that is not an out. To

model this, the slopes of α_{1k} were constrained to positive values and the prior was truncated using a half-t distribution:

$$\alpha_{1k} \sim \text{half-t}_\tau$$

The other priors are simply standard Gaussian priors,

$$\alpha_{0k}, \beta_{2k}, \beta_{3k}, \eta_k \sim \mathcal{N}(0, 1)$$

2.4 Measuring performance via wOBA

A common statistic in baseball is the *On-base percent* (OBP). This statistic is simply

$$\text{OBP} = \frac{\text{Hit + Walks + HBP}}{\text{Total Plate Appearances}}$$

However, this does not capture the full impact of the outcomes. Instead, it is important to judge hitters (and pitchers) according to the quality of hits they are responsible for.

Tango et al. introduced the idea of wOBA in 2007. In it, they weight the method of getting on base and the hit quality and generate a number (out of 1000). This wOBA is scaled so that average wOBA equals leage average OBP. The formula for wOBA is

$$\texttt{wOBA} = \frac{.69 \times uBB + .719 \times HBP + .87 \times 1B + 1.217 \times 2B + 1.529 \times 3B + 1.94 \times HR}{AB + BB - IBB + SF + HBP}$$

AB represent at-bats and IBB represent intentional walks (when the pitcher deliberately walks the hitter to get to the next one). SF represent sacrifice flies (when the hitter gets out but still manages to drive in a run).

3 Results

Instead of focusing solely on the results obtained from the TTO penalty, this section will focus on a different metric: hit quality through the order.

The original paper showed a progressive increase in wOBA as the game progressed. We instead take a look at how the lineup is constructed and the resulting wOBA throughout the order.

The first hitter of a lineup is typically a person whose skillset involves getting on base and being fast. There is a strong preference for singles and walks at this position. The second batter is typically someone who is also able to get on base but who can also drive in the runner if the first batter successfully manages to get on base.

The third and fourth hitters are typically "power hitters;" this is where the best hitters on the team go. There is an uptick in homeruns from this position as these hitters often hit the ball much harder and much farther more consistently than the rest of the team. However, there is also in increase in walks as pitchers (often wisely) are afraid to give the batter something
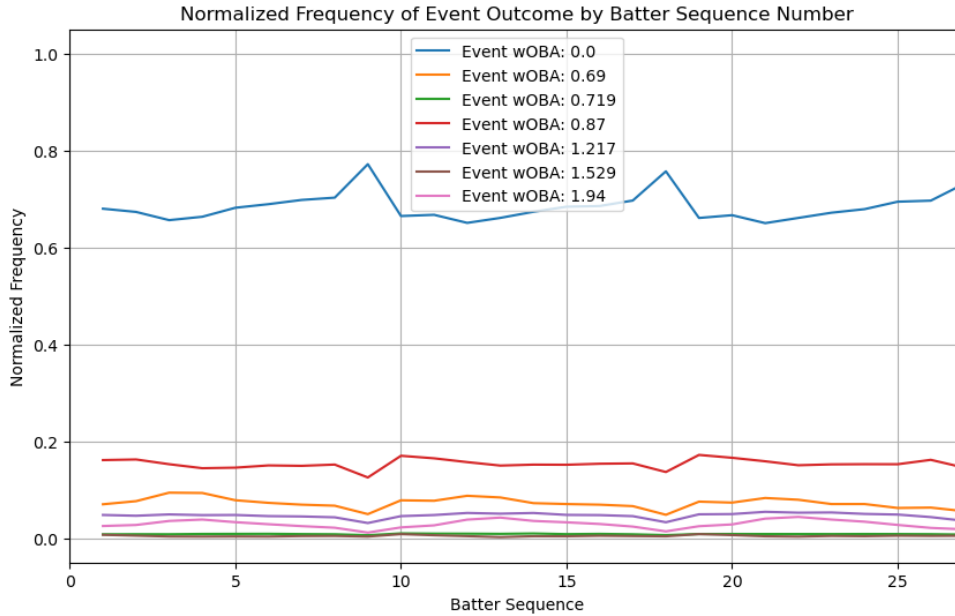
Figure 1: This figure shows the frequency of each outcome vs batter sequence number. The wOBA values correspond to out, walk, HBP, 1B, 2B, 3B, HR respectively. Notice the sharp uptick in the outs from the 9th position of the sequence, but the subsequent reduction in hits. The 9th position is typically where the worst hitter is placed (often pitchers) and so it is no surprise that this position in the lineup sees the most outs.

he can hit. These power hitters are often not as fast on their feet as the "scrappy" one and two hitters so the frequency of doubles is slightly reduced.

After that, it's more of the same. The lineup then proceeds in the order of best to worst hitters. Especially significant is the huge discrepency for the 9th position of the order. The 9th position is where the worst hitters go. Although the rules have changed, at the time of this dataset, the pitcher was required to hit. While they are great athletes, they are not often as practiced as their other teammates when it comes to hitting. It is interesting to see this clear dropoff in performance at this position, as well as the clear difference between each position at the front of the lineup.

Finally worth noting is the slight decrease in outs and increase in the frequency of singles (and other hits) as the batter sequence increases, a finding which further supports the idea of pitcher deterioration as a game progresses.

4   Conclusion

Baseball is a game that is defined by numbers. Each game of baseball results in 6-7TB of data, ranging from field positioning to launch angle of hits to bat velocity to pitch movement to pitch spin and so on. Careers are started and ended based entirely on the decimal points they produce.

In this paper, we take a brief overlook at just one statistic: the *weighted on-base average* and compare it over time.

Statistics and baseball go hand-in-hand. For over 100 years, people have been tracking the trends and evolving what it means to understand the game. Through the use of MCMC

models and other Bayesian methods, we can better characterize baseball for what it is: the greatest game in the world.

References