True wOBA: Estimation of true talent level for batters

Scott Powers and Eli Shayer

January 11, 2016

1 Introduction

Batting statistics at the play-by-play level are a solved problem. Thanks to weighted on-base average (wOBA), we know the marginal run value of each possible outcome of a plate appearance (PA) given a specific run environment (Tango et al., 2007). But player evaluation is a two-step process: (1) estimate the value of each outcome and (2) estimate the frequency with which the player produces each outcome. Relative to the first step, sabermetric literature has paid little attention to the second step. In this paper we develop a model for estimating batters' true talent level for wOBA while simultaneously adjusting for sample size, park effects and opponent quality.

2 Review of related work

Several recent papers (Brown, 2008; Null, 2009; Neal et al., 2010; Albert, 2015) have proposed methods for estimating or predicting batting statistics. In contrast with projection systems, these methods implicitly assume that players' skills remain constant between training sample and test sample (as opposed to modeling the aging of skills). True wOBA differs from the above by adjusting for park effects and quality of opposition. Deserved Run Average (Judge and BP Stats Team, 2015) uses a very similar model to True wOBA to estimate pitchers' component skills (e.g. strikeout rate). True wOBA could accurately be described as a version of DRA for batters.

3 Research methodology

3.1 Data

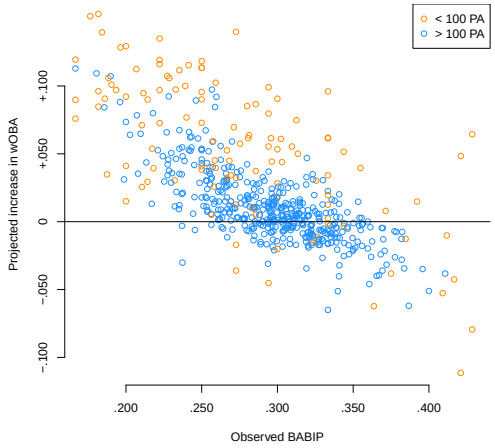
Our dataset, from Retrosheet, comprises all plate appearances (PAs) from the 2015 MLB regular season. For each plate appearance (PA) we have the identity and handedness of the batter and pitcher, the ballpark at which the PA took place, and the result of the PA.

3.2 Methods

True wOBA is based on a ridge regressed multinomial Rasch model which models the probability of each outcome of an at-bat as a function of the difference between the batter's propensity for producing the outcome and the pitcher's propensity for preventing the outcome, with adjustments for ballpark and handedness.

4 Anticipated results

We split the data into a training sample and a test sample, using the training sample to estimate batters' skills and the test sample to evaluate these estimates. Our results compare the performance of True wOBA with both naive estimators and the state of the art. Additionally, we present the impact of sample size and BABIP (e.g. see figure below) on the difference between observed wOBA and True wOBA. Intermediate results compare an opponent-agnostic special case of True wOBA with standard regression to the mean, thus illustrating the relationship between the two.



5 Expected contribution

Because we cannot observe an infinite sequence of plate appearances for a batter, we will never know the true probability with which he produces each outcome. Major League clubs can benefit from making the most out of a finite sample of data when learning the batters' skills, which has immediate applications to lineup and roster decisions. Baseball writers can also benefit from understanding how to quantitatively adjust for small sample sizes and low/high BABIPs, instead of qualitatively citing those as reasons that a batter's wOBA may not be an accurate reflection of his abilities.

References

Albert, J. (2015). Improved component predictions of batting measures. arXiv preprint arXiv:1505.05557.

Brown, L. D. (2008). In-season prediction of batting averages: A field test of empirical Bayes and Bayes methodologies. The Annals of Applied Statistics, 2(1):113–152.

Judge, J. and BP Stats Team (2015). DRA: An in-depth discussion. Baseball Prospectus.

Neal, D., Tan, J., Hao, F., and Wu, S. S. (2010). Simply better: Using regression models to estimate major league batting averages. Journal of Quantitative Analysis in Sports, 6(3).

Null, B. (2009). Modeling baseball player ability with a nested dirichlet distribution. Journal of Quantitative Analysis in Sports, 5(2).

Tango, T. M., Lichtman, M. G., and Dolphin, A. E. (2007). The Book: Playing the percentages in baseball. Potomac Books.