

Does sabermetrics have a place in amateur baseball?

Joe Gray ★ 7 March 2009

The term “sabermetrics” is one of the many creations of Bill James, the great baseball theoretician (for details of the term’s derivation and usage see Box 1). Several tight definitions exist for the term, but I feel that rather than presenting one or more of these it is more valuable to offer an alternative, looser definition: sabermetrics is a tree of knowledge with its roots in the philosophy of answering baseball questions in as accurate, objective, and meaningful a fashion as possible. The philosophy is an alternative to accepting traditional thinking without question.

Branches of the sabermetric tree

The metaphor of sabermetrics as a tree extends to describing the various broad concepts and themes of research as branches. Some of the major branches are described below.

Evaluating statistics in relation to wins

One theme of sabermetrics is that a player’s basic role is to help the team score or prevent runs and, by extension, win games. Accordingly, player statistics that are strongly correlated with a team’s run-scoring or run-preventing ability are deemed particularly valuable.

As an example, if you were to use batting average as the primary means of comparing hitters then you would be ignoring several other facets of a player’s game that are related to run scoring. To remedy this it is necessary to take multiple statistics into consideration and to weight them according to how much they contribute to team run scoring (an area known as “linear weights”). For instance, a double would, on average, contribute more than a single (because the batter ends up closer to home and enables greater runner advancement), while a strike-out would have a negative impact on run scoring.

- ★ The term **sabermetrics** combines SABR (the acronym for the Society for American Baseball Research) and metrics (numerical measurements). The extra “e” was presumably added to avoid the difficult-to-pronounce sequence of letters “brm”. An alternative exists without the “e”, but in this the first four letters are capitalized to show that it is a word to which normal rules of pronunciation do not apply.
- ★ It is a singular noun despite the “s” at the end (that is, you would say “sabermetrics is growing in popularity” rather than “sabermetrics are growing in popularity”).
- ★ The adjective **sabermetric** has been back-derived from the term and is exemplified by “a sabermetric tool”, or its plural “sabermetric tools”.
- ★ The adverb **sabermetrically**, built on that back-derived adjective, is illustrated in the phrase “she approached the problem **sabermetrically**”.
- ★ The noun **sabermetrician** can be used to describe any practitioner of sabermetrics, although to some people it might imply some minimum level of skill or contribution.
- ★ Finally, it would probably be best if the alternative adjective “sabermetrical” was reserved for jocular usage, in the way that the Internet is sometimes referred to as the “Interweb”.

Box 1. Etymology and usage of the term “sabermetrics”.

Empirical calculation of appropriate weights for a given league environment is based on the concept of run expectancy, which is the expected number of runs that will score for any combination of outs and base occupancy. To illustrate this, a typical team might score, on average, 0.8 runs with bases loaded and two outs, but it will obviously score 0 runs with bases loaded and three outs. A strike-out is a batting event that would change the former situation into the latter, and so striking out with bases loaded and two down could be seen as making a contribution to run expectancy of -0.8 . Looking at all possible batting events for all possible situations of outs and base occupancy is at the heart of deriving appropriate

weightings. Once the weightings have been derived, they can be used to calculate a single value from a selection of batting statistics.

Separating talent from the environment

In addition to shedding light on talent, a player's statistics also contain "background noise" of the environment in which they were accumulated. For instance, some players may have played more games on baseball fields favouring the hitter than other players. Correcting for this with park-adjusted figures in theory allows for a fairer assessment of a player's ability.

Assessing the extent of luck

A third factor that influences a player's statistics (in addition to talent and the environment) is luck. A good illustration of this point is batting average on balls in play (BABIP). A high BABIP is typically attributable more to luck than talent and so it is not a sustainable statistic: players with a high BABIP can be expected to see their batting averages regress towards the mean.

Regression towards the mean is not some magical force drawing values to the average level for no good reason. Instead it results from the values that experience the most regression being those that the play of chance has pushed furthest from their true value. It can explain many phenomena seen in sports and the wider world, and a good example is the "sophomore slump", in which players who have stand-out rookie years tend to see a drop-off in performance in their second season (see Box 2).

Projection of future talent

Knowing a player's current talent level is one thing, but their value as a long-term investment depends on their projected value over future levels. There are a number of different techniques that have been developed for these projections. Several of them can be found here: <http://www.fangraphs.com/projections.aspx>.

Value over replacement level

It has been argued that since a Major League team can essentially draw on a limitless supply of players who can perform at a basic (or "replacement") level, this should be taken into account when assessing a player's value to the team. That is to say, a player's value lies in what is offered above that replacement level.

- ★ Imagine a pool of rookies who all have the talent to bat exactly .260 over the course of the season. Through the play of chance some will bat more than .260 while others will bat less.
- ★ At the end of the season, there are five rookies who batted over .300 and are thus heralded as star players in the making. Crucially for this argument, they are a selective sample of players who benefited from a good chunk of luck over the season.
- ★ Back on their talent level, imagine that it rises to give each of these batters an expected .265 average over the course of their second seasons.
- ★ Even though their ability has increased, it is still likely that none of these five players will bat over .300 in their second season. For all of them to perform at their rookie-year level would be very unlikely. Thus, we are very likely to see some regression towards the mean.
- ★ Regression towards the mean works in both directions, so if journalists went looking for it (and assuming that under-performing rookies got a chance to prove themselves) then a "sophomore jump" phenomenon would also reveal itself.

Box 2. Explaining the phenomenon of the "sophomore slump" through regression towards the mean.

This idea borrows from the economic concept of scarcity, with the resource that is scarce – and thus worth paying good money for – being ability above the replacement level.

Optimal roster construction

It is teams rather than individual players that win games, and optimal roster construction is about having the set of players that gives a team the best chance of victory. Even without financial constraints, the task of constructing an optimal finite roster of players would require a careful balancing of players' roles. Add in the constraint of a finite budget and the problem becomes substantially more complex.

In-game tactics

The final branch of the sabermetric tree presented here is in-game tactics, or more specifically choosing the option in making a tactical decision that gives the team the best chance of success on average (otherwise known as "playing the percentages"). When it is beneficial to intentionally walk a player, steal a base, or lay down a sacrifice bunt are among the many choices to have been sabermetrically analysed.

Summary of the branches

In summary, then, sabermetrics has been employed:

- 1 to extract the components of players' statistics that are closely related to current talent (as measured in contributing to wins) from environmental noise;
- 2 to consider the extent of luck in players' statistics when evaluating current talent;
- 3 to project future talent;
- 4 to assess value as the fraction of current and future talent that is a rare commodity;
- 5 to consider the roster as a whole when interpreting value rather than evaluating each player in isolation;
- 6 to offer guidance on the tactical decisions made during each game that will give the roster of players as high a chance of winning the game as possible.

Is this relevant to amateur baseball?

For points 3, 4, and 5, the answer to the question of relevance to amateur baseball is an easy one: no. Projections, value over replacement level, and roster construction are obviously not important in a league without trades or roster limits.

- ★ Fielding average is by far the most well-established defensive statistic, but by focusing on errors, it does not do a good job of capturing a fielder's ability to cover ground or makes throws in a fashion that is better than ordinary (a difficult ball that gets past a fielder or a throw that is too weak but mechanically fine will not show up in a player's fielding average).
- ★ As a way of accounting for both of these factors, the proportion of balls hit in the fielder's zone that are converted into outs can be considered (such stats belong to the "zone rating" family). The zone can be defined for each position as all the batted balls that will be handled by a particular fielder for an out at least 50% of the time. It is also possible to consider the proportion of balls hit outside of a fielder's zone that are converted into outs.
- ★ Unfortunately, sophisticated methods for tracking batted balls are required to generate the data, and this is something that is not available in amateur baseball. And even if it was, the seasons in amateur leagues would not be long enough to generate sufficient data for any meaningful conclusions on fielding talent based on stats.
- ★ Stats cannot help us that much, so we need to rely more on subjective views. Fortunately, it is easier to assess the quality of a fielder through isolated observations than it is for a pitcher or batter.

Box 3. A look at why fielding statistics are not particularly useful in amateur baseball.

For point 2, the answer, once more, has to be no. Away from a very useful application to evaluating trades and roster construction (be it in professional baseball or a fantasy setting), to attempt to make a correction for luck in sport would be to bleach its character. Luck is what prevents sport from being mundanely predictable.

So this leaves us with the first and last points on the summary list above. For point 6, the margins of gain from a sabermetric approach to tactics will be small in the relatively short seasons that are characteristic of amateur baseball. In addition, to calculate the percentages as accurately as possible (and thus be able to confidently "play" them) it would be necessary to assess tactical decisions in the context of the league setting, rather than transplanting sabermetric guidance from professional baseball as played in North America. Nevertheless, employing sabermetrically derived strategy on the field will probably be more beneficial than harmful.

Point 1 provides, at last, a branch of sabermetrics that is of real relevance to amateur baseball. Of course, there is no financially motivated need in amateur baseball to assess players' statistics in relation to their contribution to wins, and to separate talent from the environment. However, there may be a desire to do this anyway, and one that is more pure. In amateur baseball, we can still ask questions such as which batter or pitcher has put in the best single-season performance in the history of a league and want an answer that is based on the most appropriate statistics (rather than those engrained through tradition, such as batting average and win-loss records) and that corrects in some way for the environment. (I deliberately avoided giving a fielding example here. The reason is provided in Box 3.)

Conclusion

In a run-through of sabermetrics that was far from being exhaustive, we have encountered a branch that is of genuine relevance to amateur baseball. The answer to the question in the title, therefore, is yes.



To comment on this article, please visit
<http://www.baseballgb.co.uk/?p=2063>