ASSESSING OFFENSIVE CONTRIBUTION

The Linear Weights Method

by Chuck Korb

The purpose of this article is to illustrate the linear weights method of assessing the offensive contribution of major league baseball players. I originally wrote this for a class I taught at MIT during their Splash program called "Markov and Baseball, Assessing Offense through Linear Weights." The premise of this assessment tool is based on the brilliant work of Pete Palmer, and the metrics used in this article borrow liberally from the further research of people like Mitchell Lichtman, Tom Tango, and many more.

Baseball is Markovian in nature: What happens in an at bat is dependent on past events, just as future events are influenced by present events. What I mean is that a single with two outs and no runners on base has a different value than a single with the bases loaded and no outs, with the events leading up to those respective singles being the difference, and the subsequent events also influencing the total run result. So how does one determine with any degree of accuracy the "value" of a single (or the result of any at bat)? The following chart will help answer that question. The chart gives the "net expected run value" for each of the 24 different situations a batter can see in an at bat, from none on, none out, to bases loaded, two outs. The run values are calculated by looking at each at bat over the course of a season and determining how many runs score subsequent to each of the situations. These totals are then averaged to give the "net expected run values" for each situation. This may be easier to understand when looking at the first batter of an inning--with no outs and none on, an average team will score 0.517 runs. Multiplying this by 9 innings in a game yields just over 4.6 runs, which is the average number of runs scored per game by Major League Baseball teams last season.

Looking at the chart, one can see that with a man on first and one out, an average team (that scored the major league average number of runs) would be expected to score 0.533 runs. If the batter coming up in that situation doubles, and the run scores, he has influenced the game by +1.00 for the run that scored and +0.155 for the difference between the resulting runner on second with one out (0.688) and the former runner on first with one out (0.533), for a total contribution of +1.155 runs. Conversely, if the batter popped out he would have contributed -0.310 runs, the difference between the situation he inherited and the one he left. (If you're not familiar with this chart, take a few minutes and play with the situations--it's really pretty cool and shows clearly that most "productive outs" are not really all that productive.)

Expected Runs by Situation

OUTS	RUNNERS							
	None	1st	2nd	3rd	1st, 2nd	1st, 3rd	2nd, 3rd	Loaded
0	0.517	0.884	1.140	1.315	1.484	1.769	2.014	2.279
1	0.279	0.533	0.688	0.966	0.921	1.202	1.416	1.558
2	0.106	0.223	0.322	0.370	0.456	0.520	0.565	0.750

One way of assessing a player's offensive contribution would be to go through each of their at bats individually over the course of a season and calculate the results of their contributions based on the situations they inherited and how their at bats affected them. If you were to do this for every non-pitcher in baseball and total the results, you would logically have a total sum close to zero runs (you would miss non-batter or runner events like balks, wild throws to first on pick-offs, etc.), as all these players together make up the MLB average. Besides being incredibly work intensive, this method of assessing offense can be somewhat misleading in that it is situationally based, and a player coming to bat after two batters with .400 OBP could have an advantage over a player batting after weak hitters. So what else can we do?

I like using the following method, which averages all the batting results over the course of a season—i.e., the average contribution of all singles, doubles, fly balls, stolen bases, etc.—and finds a value for each of those specific results. In other words, if you took the total contribution, as determined using the chart, of every single over the course of an entire season and divided it by the total number of singles, you would have the average value of a single independent of the situation. The tricky thing here is how detailed you want to get in your breakdown. Do you want to know the value of all ground balls, or do you want to break them down by ground balls to each position? I will leave the level of breakdown to you for your own future forays into this type of assessment, but will say that a simple breakdown yields very similar results to a more detailed one, and is much less labor intensive. What I have done for the 2009 season is use the average value of 11 offensive results to determine the offensive contribution in runs above or below average. The values I determined were:

- singles +0.48
- doubles +0.77
- triples +1.07
- home runs +1.40
- walks/hpb +0.32
- intentional walks +0.185
- strikeouts -0.280
- ground outs -0.301
- fly outs -0.266
- stolen bases +0.193
- caught stealing -0.437

Applying these values to the total offensive statistics for all players in baseball in 2009 yields a net result of (almost) zero, meaning these average numbers "work."

The above is done to illustrate how the linear weights method of assessing runs is calculated, and is somewhat simplified. For more accurate and meaningful results, I would calculate the leagues separately, and remove pitcher at bats from my National League numbers. **MSP**