Data 8 Connector: Sports Analytics

...

Measuring Offensive Performance in Baseball

Road Map

We’re going to get into measuring offensive performance in baseball

1. Classic statistics in baseball
2. Break down the classics and study what they are trying to do
3. Introduce run expectancy (RE) and the Run Expectancy Matrix
4. Use RE to construct linear weight estimates of run creation

What are we trying to do?

Non-pitchers possess two skills: hitting and fielding

These are separate issues and fielding is more complicated so we start with offense

Hitting allows us to isolate a player at the plate and try to measure their performance

Data: box score performance or play-by-play

Our goal: collect outcomes of plate appearances and try to summarize hitter ability

Value, Ability, and Talent

Value

Player's actually worth, factoring in clutch plays and important or high value situations.
Backward looking in how it is specific to events that have happened

Ability

Remove situation/context of a play: imagine the player on an "average" team in "average" circumstances
Forward looking by removing context of history to allow projection of future performance

Talent

True skills that are not obviously quantifiable. Little happens in a vacuum so impossible to determine.

We take the view of performance is tied to ability and that's what we try to quantify

Measuring Offensive Performance in Baseball

Classic statistics for measuring offense

- ➔ Batting Average (BA)
- ➔ On-Base Percentage (OBP)
- ➔ Slugging Percentage (SLG)
- ➔ Runs Batted In (RBI)
- ➔ Runs

Measuring Offensive Performance in Baseball

The Classics

Hits are good: Batting Average

Walks are also good: On-Base Percentage

More bases is good: Slugging Percentage

Runs are good: Runs scored or RBI

Measuring Offensive Performance in Baseball

Batting Average treats all hits the same. On-Base Percentage includes walks.

Slugging Percentage weights extra base hits more. How much more?

- Weight according to number of bases (single = 1, double = 2, etc)

High average can be good, but it can also be misleading

- Think of boxing: hitting singles is like hitting weak jabs. Numbers may look good but no power so is it actually effective?

OBP helpful for quantifying skill of getting on base

SLG helpful for quantifying power skill

Run Creation

Recall Pythagorean Expectation: we need to create runs when batting

So what are we really trying to do? Measure run creation!

Which of these classic stats, if any, correspond to creating runs? How do we determine if a stat measures run creation?

Run Creation

Take batting average:

- ➔ It tells you proportion of hits vs outs. That's it. Very primitive and simple.
- ➔ More hits tends to lead to more runs but this can't be the best we can do, can it?

BA is also a "proxy"

Does not directly measure run creation, just something that should be related

The Classical statistics are all proxies

This is now our motivation: work towards better estimation of run creation

How do "The Classics" perform? How do they relate to run scoring?

Can we directly estimate run creation?

Or can we build better proxies?

Let’s Look Closer at the Classic Stats

Definitions

We need to define some events first

Hit (H):

> Plate appearance ending with a batter hitting the ball and getting on base safely without an error or fielder's choice

Error:

> Misplay by a fielder leading to a batter getting on base safely or runners advancing. Scorer determines the credited event the hitter or runners should be given were there no error.

Fielder's Choice (FC):

> A fielder opts for a play that allows the batter to reach safely even though he was likely to be out. No offensive credit is given to the batter.

Definitions

Single/Double/Triple/Homerun (1B/2B/3B/HR):

Batter reaching safely at first, second, third, or home, respectively.

Walk (BB):

Plate appearance ending with 4 balls. Intentional walks denoted IBB. Does not count as an AB

Hit-by-pitch (HBP):

Similar to walk but with the batter being hit by a pitch. Does not count as an AB

Sacrifice:

The batter intentionally taking an out via bunt or outfield flyball to advance (and/or score) a runner

Definitions

Plate Appearance (PA):

When a hitter steps up to the plate (ignore edge cases like switching hitter in the middle of a PA)

At-bat (AB):

Plate appearances that do not end in a walk, hit-by-pitch, sacrifice, or catcher interference. Errors and fielder's choices count as ABs.

Batting Average

Batting average

$$BA = \frac{H}{AB}$$

Very simple: count up hits and at-bats and divide

Measures proportion of at-bats (so not all PAs!) that yield a hit

Three key observations:

1. Ignores walks (and also does not use PA as the denominator)
2. Weights all hits equally: a homerun is equal to a single
3. Independent of situation/context: a hit with bases loaded equals a hit with bases empty

On-Base Percentage

On-Base Percentage

$$OBP = \frac{H + BB + HBP}{AB + BB + HBP + SF} \approx \frac{H + BB + HBP}{PA}$$

Similar to BA: include BB and HBP and divide by PA (approximately) instead

Walks (and HBP) help score runs and shouldn't be discarded

- This is a skill: not making an out and at least getting on base to prolong an inning

Still weights events equally, still situation independent

Slugging Percentage

Slugging Percentage

$$SLG = \frac{(1 \times 1B) + (2 \times 2B) + (3 \times 3B) + (4 \times HR)}{AB}$$

Very different now: no BB or HBP but weighted by bases!

Measures power level: extra base hits should be worth more

- The weight is assigned as proportional to number of bases

Events no longer weighted equally but still situation independent

Runs and Runs Batted In

Runs: tallied to a runner who crosses home plate

Runs Batted In: all runs scored on a hitting event tallied to the batter

- ➔ Ex: A runner is driven in from second base on a double. Batter gets an RBI, runner a run
- ➔ That's confusing: two runs credited for one run scored?
  Runner got to second but almost anyone can score from second on a double
  Batter gets an RBI but did nothing to get the runner to second
- ➔ Attribution problem: who's responsible, ie. who should get credit?
  Should getting a hit with a runner on be more valuable than the same hit in another situation?

What a mess!

Summary of Classical Baseball Stats

- ➔ Measuring some kind of performance: hitting, on-base, power
- ➔ Rate statistics (more PA doesn't increase value): BA, OBP, SLG
- ➔ Sometimes equally weighted, sometimes not
- ➔ Different events may count (walks)
- ➔ Attribution confusion (runs and RBI)
- ➔ Situational independence: BA, OBP, SLG
- ➔ Situational dependence: runs, RBI
- ➔ Bases Fallacy

Absent anything else, we could use these. But we have some data, let's try to do better!

Run Creation and Classical Baseball Stats

Now we know a bit about classical baseball stats...

Here are some questions

- ➔ Do these stats actually correlate with runs scored (runs and RBI obviously do)? Do they tell us about what we really care about (run creation)?
- ➔ We know we should weight events differently, but how?
- ➔ What about other potential events?

Correlation and Errors

Start with the first question:

- Do these stats actually correlate with runs scored?
- How well do they predict run scoring?

Let’s do a quick demo to address this...

Correlation and Errors

Okay, so yes, the stats are correlated!

Correlation: the good and the bad

Good:

- Two quantities have a relation: measuring one can potentially help you infer the other
- Stronger correlations tends to make for better inference

Bad:

- $\rho = \text{Corr}(x, y) \Rightarrow \rho = \text{Corr}(x + 42, y)$ that is x does not necessarily mean anything
- $y = a \cdot x + \text{error}$ => Magnitude of errors in relation to a drive correlation

This is why we also inspect the magnitude of errors

Correlation and Errors

BA is not as reliable an indicator of run scoring as OBP or SLG

Looks like our insight is confirmed:

- Walks matter, power matters

SLG used a simple heuristic to weight different hits

Our goal: a data-driven approach to weighting events

Run Expectancy

Determining Event Value

How should we weight the batting events to improve measurement?

- ➔ Baseball is filled with events. Singles, strikeouts, stolen bases, etc.
- ➔ If a single occurs with no one on so no run is scored, is it still worthless? How do we assign a value?
- ➔ How about a double? Should we assign it double the value of a single?
- ➔ What's the value of a stolen base (ie. advancing an extra base)?
  Or a sacrifice bunt/fly (trading an out for advancing/scoring a runner)?

Let's use the *expected value* an event adds to run creation

Wait, expected value...? We need to introduce the concept of Run Expectancy first

Run Expectancy

Intuition:

A batter walks up to the plate with 1 out and runners on 1st and 2nd

From this PA to the end of the inning, how many runs do we expect the team to score?

1 run? 2 runs? 3 runs?

If we have this number, that would be the Run Expectancy value

Run Expectancy

A bit more rigorous...

Given a state of:

- Runners on base (runner on 1st; runners on 2nd and 3rd; etc)
- Number of outs (0, 1, 2)

Run Expectancy:

> The number of runs we expect a team to score from a plate appearance in the given state to the end of the inning.

The values for the 24 different combinations of runners and outs are called the *Run Expectancy Matrix*

Run Expectancy

Here's a bit of explanation/insight

No runners on, no outs:

> This is usually the start of the inning. The expected number of runs scored from this point and after should be approximately Runs per Game / 9

Bases loaded, no outs:

> This is an amazing position to be in. Our team should expect to score a lot of runs from this and after. This should have the highest Run Expectancy

No runners on, two outs:

> We have nothing. We're likely to be out in this next plate appearance. Run Expectancy should be lowest.

Run Expectancy

Three questions for Run Expectancy:

How do we compute the Run Expectancy Matrix?

If we can compute it, how can we use it?

What would it tell us?

Computing the Run Expectancy Matrix

How do we compute the Run Expectancy Matrix?

Surprisingly simple!

Using the Retrosheet play-by-play data:

1. For observed each plate appearance, log the following:
   - ➔ State, i.e. the number of outs and the runners on base
   - ➔ Total number of runs scored in the remainder of the inning including those during the PA
2. For each of the 24 states, compute the average runs scored in remainder of the inning

Done. That's the Run Expectancy Matrix.

Computing the Run Expectancy Matrix

Here is an example of a Run Expectancy Matrix

Notice:

- In terms of RE, an out varies in value
- A walk varies in value
  - (None on to 1st vs. 2nd and 3rd to loaded)
- Sacrifice fly: 2nd vs 3rd

Outs	0	1	2
Bases			
None on	0.534	0.288	0.116
1st	0.925	0.551	0.247
2nd	1.171	0.709	0.351
3rd	1.517	0.979	0.372
1st and 2nd	1.527	0.917	0.437
1st and 3rd	1.848	1.260	0.527
2nd and 3rd	2.037	1.427	0.607
Bases Loaded	2.347	1.598	0.804

RE from 2001

Using the Run Expectancy Matrix

What does it tell us?

Expected number of runs scored under average circumstances

- ➔ Expected: Repeat situation a lot and runs scored will concentrate around the value.
- ➔ Average Circumstances: average hitters, pitchers, defenders, inning, etc.
  We averaged over all teams/players/circumstances.

Using the Run Expectancy Matrix

Computing it was easy. So how do we use it?

Scoring potential: high run expectancy = high potential

Event valuation

- ➔ No runners on, no outs. Batter gets a single. Now there's a runner on first, still no outs. What was the single worth in terms of Run Expectancy?
- ➔ 1 out, runner on first. In Run Expectancy, what can you gain by stealing second? What can you lose?

Creates an expectation for run scoring so we can better value an event

Run Expectancy Demo

Let’s do a demo.

Strategies: Bunting and Stealing

A bunt trades an out for advancing a runner

Assuming 100% bunt success, we can compare the Run Expectancies to determine the effectiveness of bunting

Runner on 1st:	0.925 - 0 out	0.551 - 1 out
Runner on 2nd:	0.709 - 1 out	0.351 - 2 out

So even if 100% successful, RE goes down if we bunt. Similar for other runner combos.

Remember: this is under average circumstances, ie. average hitters.

A pitcher is almost guaranteed out, might as well bunt.

Strategies: Bunting and Stealing

A steal is much more likely to be unsuccessful so we compute a probability equivalent

$$RE_{\text{Steal Attempt}} = RE_{SB} \cdot p_{\text{SB}} + RE_{CS} \cdot (1 - p_{\text{SB}})$$

$$RE_{\text{Current}} = RE_{\text{Steal Attempt}} \Rightarrow p_{\text{SB}} = \frac{RE_{\text{Current}} - RE_{CS}}{RE_{\text{SB}} - RE_{CS}}$$

This probability tells us how good a base stealer needs to be to make it a break-even strategy on average

Strategies: Bunting and Stealing

Runner on 1st, 0 out

$$p_{\text{SB}} = \frac{0.925 - 0.288}{1.171 - 0.288} = 0.722$$

Need an okay base stealer to break even (MLB average is right around this value)

Runner on 2nd, 2 out

$$p_{\text{SB}} = \frac{0.351 - 0.0}{0.372 - 0.0} = 0.943$$

Remember the old rule: “Never make the 1st or 3rd out at third base”

Strategies: Bunting and Stealing

Demo Section 2.5: on your own evaluating strategies (part of it will be on HW2)