LS88: Sports Analytics

Classical Baseball Statistics
& Measuring Performance

Road Map

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

1. Classic statistics in baseball; then two advanced metrics
2. Break down the metrics and study what they are trying to do
3. Then: run expectancy (RE) and the Run Expectancy Matrix
4. Use RE to build better metrics for run creation

What are we trying to do?

Non-pitchers possess two skills: hitting and fielding

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: use outcome data 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" situations
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 talent is impossible to determine.

We view performance as tied to ability and that's what we try to quantify

Towards Run Scoring

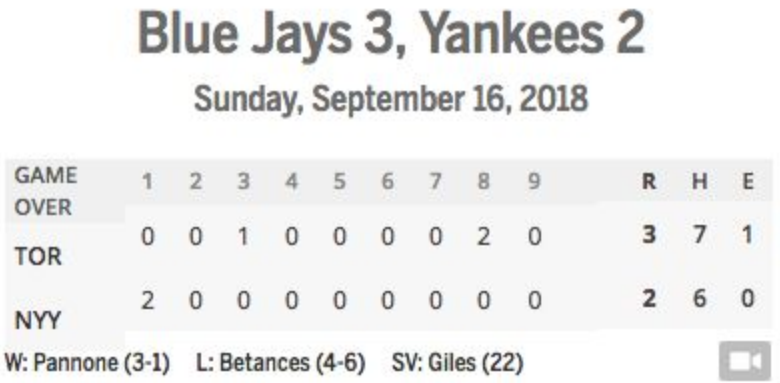
➔ Pythagorean expectation showed us how run scoring relates to winning

$$\text{Pythagorean Expectation Win Pct} = \frac{\text{Runs Scored}^2}{\text{Runs Scored}^2 + \text{Runs Allowed}^2} = \frac{(\text{Runs Scored} \ / \ \text{Runs Allowed})^2}{1 + (\text{Runs Scored} \ / \ \text{Runs Allowed})^2}$$

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

Which of the classic stats, if any, correspond to creating runs? How do we even determine if a stat really can measure run creation?

Box Scores



Blue Jays 3, Yankees 2

Sunday, September 16, 2018

GAME OVER	1	2	3	4	5	6	7	8	9	R	H	E
TOR	0	0	1	0	0	0	0	2	0	**3**	**7**	**1**
NYY	2	0	0	0	0	0	0	0	0	**2**	**6**	**0**

W: Pannone (3-1) L: Betances (4-6) SV: Giles (22)

Box Scores

TORONTO BATTERS	AB	R	H	RBI	BB	SO	LOB	AVG	OPS
1 McKinney *LF-1B-LF*	4	0	0	0	0	2	3	.289	.861
2 Smoak *1B*	4	0	1	0	0	2	1	.248	.820
a-Pillar *PR-CF*	0	1	0	0	0	0	0	.246	.689
3 Morales *DH*	1	0	0	0	2	0	0	.255	.791
1-Davis *PR*	0	0	0	0	0	0	0	.167	.334
Tellez *PH-DH-1B*	1	0	1	1	0	0	0	.407	1.207
4 Grichuk *CF-RF*	4	0	1	1	0	2	2	.247	.801
5 Solarte *3B*	4	0	0	0	0	1	4	.232	.676
3-Díaz, A *PR-3B*	0	0	0	0	0	0	0	.257	.739
6 Hernández, T *RF-LF*	4	0	0	0	0	1	3	.243	.770
Giles *P*	0	0	0	0	0	0	0	.000	.000
7 Ureña, R *SS*	4	1	1	0	0	1	0	.262	.673
8 Travis *2B*	4	0	1	0	0	0	1	.236	.669
9 McGuire, R *C*	4	1	2	1	0	0	1	.308	.819
TOTALS	34	3	7	3	2	9	15		

a-Singled for Smoak in the 8th. 1-Ran for Morales in the 6th. 2-Ran for Smoak in the 8th. 3-Ran for Solarte in the 8th.

BATTING

2B: Grichuk (28, Betances); McGuire, R (2, Lynn).

TB: Grichuk 2; McGuire, R 3; Smoak; Tellez; Travis; Ureña, R.

RBI: Grichuk (54); McGuire, R (1); Tellez (5).

Runners left in scoring position, 2 out: Smoak; McKinney; Hernández, T.

Team RISP: 3-for-8.

Team LOB: 6.

NY YANKEES BATTERS	AB	R	H	RBI	BB	SO	LOB	AVG	OPS
1 McCutchen *RF*	4	1	1	1	0	2	0	.252	.785
2 Stanton *DH*	3	1	2	0	1	0	0	.265	.842
3 Hicks, A *CF*	4	0	1	0	0	2	2	.242	.811
4 Andújar *3B*	3	0	0	0	0	1	2	.298	.851
Hechavarria *3B*	0	0	0	0	0	0	0	.254	.633
a-Walker *PH*	1	0	0	0	0	0	0	.218	.640
5 Gregorius *SS*	2	0	1	1	1	0	0	.270	.838
6 Torres *2B*	4	0	1	0	0	1	2	.279	.843
7 Sánchez, G *C*	4	0	0	0	0	2	3	.188	.692
8 Voit *1B*	3	0	0	0	0	0	2	.280	.908
9 Gardner, B *LF*	3	0	0	0	0	1	0	.237	.689
TOTALS	31	2	6	2	2	9	11		

a-Flied out for Hechavarria in the 9th.

BATTING

HR: McCutchen (19, 1st inning off Pannone, 0 on, 0 out).

TB: Gregorius; Hicks, A; McCutchen 4; Stanton 2; Torres.

RBI: Gregorius (84); McCutchen (62).

Runners left in scoring position, 2 out: Hicks, A; Voit.

SF: Gregorius.

Team RISP: 0-for-4.

Team LOB: 5.

Single


MLB.com


YANKEESBEISBOL.COM
NEW ERA
ON-FIELD
NEW ERA
StubHub
MLB.com

Home run

Strikeout

Measuring Performance

We know...

- ➔ Different events occur in a game that build towards runs
- ➔ Visually and in the scoresheet/box score we can distinguish these
- ➔ You need to round the bases to score runs

We can start measuring performance with classical stats: BA, OBP, SLG.

Also RBI and Runs: but we'll see why those are no good

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. RBOE stands for Reached Base On 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

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.

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

Definitions

Sacrifice:

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

Plate Appearance (PA):

When a hitter steps up to the plate (we ignore edge cases like switching hitter in during 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

Classical Baseball Statistics

Classical Baseball Statistics

The Slash Line

- ➔ Batting average
- ➔ On-Base Percentage
- ➔ Slugging Percentage

Also:

- ➔ Runs Batted In (RBI)
- ➔ Runs (R)

Classical Baseball Statistics

Quick overview

- 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

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

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

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!

Runs and Runs Batted In

Runs (R): tallied to a runner who crosses home plate

Runs Batted In (RBI): all runs scored in a PA tallied to the batter

Ex: A runner is driven in from second base on a double. Batter gets an RBI, runner a run

Demo

Let’s get to know these stats with data

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 (R and RBI obviously do)? Do they tell us about what we really care about (run creation)?
- ➔ Is it helpful to weight events differently?
- ➔ If so, how should we weight events differently?
- ➔ What about other potential events?

Correlation and Errors

We can answer the first two questions using two methods:

1. Correlation: the value between -1 and 1 that measures the strength of the linear relationship. .5 is pretty good. .8 is great. .9+ is amazing.
2. Standard deviation of errors (aka Root-Mean-Square-Error aka RMSE):

$$y_{\text{Model}} = \alpha + \beta \times x$$

$$\text{Error} = y_{\text{observed}} - y_{\text{Model}}$$

We compute the standard deviation of all those errors

Run Creation and Classical Baseball Stats

Back to the demo.

Run Creation and Classical Baseball Stats

Okay, so yes, the stats are correlated with run scoring!

Correlation: the good and the bad

The Good:

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

The Bad:

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

This is why we also inspect the magnitude of errors

Digging Deeper on the Classical Baseball Stats

Aside from the data, we can inspect what the classical stats are assuming or how they do their job

Revisiting Batting Average

Batting average

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

Observations on BA:

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

Revisiting On-Base Percentage

On-Base Percentage

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

Observations on OBP:

1. 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
2. OBP like BA weights events equally
3. OBP like BA is *situation independent*

Revisiting Slugging Percentage

Slugging Percentage

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

Observations on SLG:

1. Measures power level: extra base hits should be worth more
    a. The weight is assigned as proportional to number of bases
2. Events are no longer weighted equally
3. *Situation independent*

Revisiting Runs and Runs Batted In

Recall: 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
- ➔ Correlates highly with run scoring!
- ➔ Rate statistics (more PAs doesn't increase value): BA, OBP, SLG
- ➔ Count statistics (more opportunities boosts value): R, RBI
- ➔ Sometimes equally weighted, sometimes not
- ➔ Different events may count (BB in OBP)
- ➔ Attribution confusion (R and RBI)
- ➔ Situational independence: BA, OBP, SLG
- ➔ Situational dependence: R, RBI
- ➔ *Bases Fallacy*

Summary of Classical Baseball Stats

Absent anything else, we could/should use the slash line

- ➔ BA is not as reliable an indicator of run scoring as OBP or SLG
- ➔ BA misses two import concepts: walks matter and power matters
- ➔ SLG uses a simple heuristic to weight different hits

The classical stats are *proxies* for run scoring

Do not directly measure run creation, just something that is related

This will motivate later work: actual estimation of run creation

Advanced Metrics

There are two advanced statistics that can be used in place of the slash line

➔ OPS (pronounced *oh-pee-ess* or *ops*)

➔ wOBA (pronounced *whoa-buh*)

OPS: On-Base Plus Slugging

OPS is a fairly famous advanced stat developed 30 or so years ago

Its name sums up its construction

An alternate form multiplies OBP by 1.8 first and then adds the two

$$OPS = OBP + SLG$$

OPS clearly tries to combines the positives of OBP and SLG

wOBA: Weighted On-Base Average

- ➔ Developed by Tom Tango and his coauthors for *The Book*
- ➔ Designed to feel just like OBP (ie. .330 is about average, .400+ is amazing)
- ➔ The original formula

$$wOBA = (0.72 \cdot BB + 0.75 \cdot HBP + 0.90 \cdot 1B + 0.92 \cdot RBOE + 1.24 \cdot 2B + 1.56 \cdot 3B + 1.95 \cdot HR) / PA$$

On Fangraphs, the weights are updated each year

We'll dive deeper into this but here is where you can find the weights
https://www.fangraphs.com/guts.aspx?type=cn

wOBA: Weighted On-Base Average

$$wOBA = (0.72 \cdot BB + 0.75 \cdot HBP + 0.90 \cdot 1B + 0.92 \cdot RBOE + 1.24 \cdot 2B + 1.56 \cdot 3B + 1.95 \cdot HR) / PA$$

wOBA recognizes you need to weight the events differently but not according to bases (designed to ignore the *bases fallacy*)

➔ See the values for 1B, 2B, 3B, HR: not in 1:2:3:4 proportion

➔ BB and HBP not as valuable as 1B

Advanced Metrics and Run Creation

How do they perform??

Back to the demo...

Final Thoughts: Five Good Stats

- ➔ The slash line works
- ➔ Even without empirical reasons, R and RBI are conceptually bad
- ➔ OPS and wOBA are two advanced metrics that perform exceptionally
- ➔ Use the slash line, OPS, or wOBA depending on who you're talking with
- ➔ Know the levels of the stats that tell you what's bad, average, good, amazing. Like .300 BA is good and .350 is amazing.

Final Thoughts: Player Measurement

Team measurement vs player measurement

Our framework so far:

1. Determine how a metric performs at the team level, ie. team run scoring
2. Take it on assumption that results should pass over to the player level

We'll revisit this as we go but: it's an okay assumption but one we need to know we're making

Final Thoughts: Moneyball

A quote from Moneyball:

By analyzing baseball statistics you could see through a lot of baseball nonsense. For instance, when baseball managers talked about scoring runs, they tended to focus on team batting average, but if you ran the analysis you could see that the number of runs a team scored bore little relation to that team's batting average. It correlated much more exactly with a team's on-base and slugging percentages.

What do you think about that quote?

Final Thoughts: Henry Chadwick

A very brief history lesson

- ➔ Baseball traces its roots to Cricket and Rounders
- ➔ Henry Chadwick is the Father of Baseball/Baseball Statistics
  As a Brit, he initially played Cricket and Rounders

And now I'll let Colin Wyers take it away...

Final Thoughts: Henry Chadwick

And so baseball won [against Cricket]— but there are still some strong influences from cricket left over that persist into today. Very few of them influence how the game is played, mind you—they all almost show up in how the game is observed and recorded. For, you see, Henry Chadwick was a cricket player. He was also the most influential sports journalist... probably of all time. [...] And one of the things he did [...] is create the box score.

https://www.baseballprospectus.com/news/article/11054/manufactured-runs-everything-you-wanted-to-know-about-run-prevention-but-were-afraid-to-ask-part-2/

Final Thoughts: Henry Chadwick

Chadwick seemed to misunderstand the role of walks in baseball. The closest analogue in cricket is the "wide," a ball so high or so far off that the "striker" [...] has no chance of reaching it. In cricket these count as "extras," which are akin to baseball's battery errors, wild pitches, and passed balls. [...] Notably, extras were not counted for individual strikers in cricket. By accounting for walks in much the same fashion, you wholly ignore both the skill it takes for a batter to "work a walk" by not swinging at pitches outside the zone and the effect a walk has on a team's ability to score runs. Chadwick's favored offensive metric, batting average, excluded walks entirely, as though the batter never stepped to the plate.

Final Thoughts: Henry Chadwick

This is not the only case of Henry Chadwick's work persisting til this day

Old habits die hard.
Science and empirics may show batting average isn't what we want but...