LS88: Sports Analytics

Linear Weights

Linear Weights

Recall BA, OBP, and SLG

➔ Weight events by a value, and then add up (linear)

Linear Weights

Batter_ID	Lineup_Order	Inning	Outs	Event_Type	Start_Bases	End_Bases	Event_Runs	Run_Expectancy	Run_Expectancy_Next	RE24
seguj002	1	1	0	BB	None on	1st	0	0.529067	0.913562	0.384494
hanim001	2	1	0	K	1st	1st	0	0.913562	0.526186	-0.387376
canor001	3	1	1	Generic out	1st	1st	0	0.526186	0.231741	-0.294445
cruzn002	4	1	2	K	2nd	2nd	0	0.330982	0	-0.330982
escoy001	1	1	0	1B	None on	1st	0	0.529067	0.913562	0.384494
calhk001	2	1	0	2B	1st	2nd and 3rd	0	0.913562	2.0538	1.14024
troum001	3	1	0	Generic out	2nd and 3rd	2nd	1	2.0538	0.718396	-0.335402
pujoa001	4	1	1	Generic out	2nd	2nd	0	0.718396	0.330982	-0.387414
cronc002	5	1	2	Generic out	2nd	2nd	0	0.330982	0	-0.330982
seagk001	5	2	0	Generic out	None on	None on	0	0.529067	0.279063	-0.250004

From Retrosheet, we have RE24 for every plate appearance

Linear Weights

For each type of event, average the observed RE24 values

This is *Linear Weights*: average run production values by event

➔ We can use them to directly weight events

➔ Weight values give contribution in *runs*, not some other unit

Later, we'll use the weights to build wOBA

Linear Weights Interpretation

- ➔ Randomly pick a single that occurred during the season
- ➔ What do we expect this single to have produced in runs?

- ➔ Let's use the single as an example

  Single with no one on: just adds a runner at first. That's an easy value boost to compute.
  Single with runner on first: sometimes leads to 1st and 2nd, sometimes 1st and 3rd
  Single with the bases loaded: sometimes 1 run, sometimes 2, maybe 3?? Maybe runner out at home?
  - ◆ 1st and 3rd is better than 1st and 2nd
  - ◆ Singles with bases loaded are very high value

Linear Weights Interpretation

$$\text{Run Value for Event } E = \sum_{\text{All Outcomes}} \{\text{Prob. of Outcome } O \text{ from Event } E\} \times \{\text{RE24 of Outcome } O\}$$

The average balances frequency of situation/baserunning outcomes against observed run production values

- 1st and 3rd is better but maybe not as likely as 1st and 2nd
- Bases loaded is a rarer situation, so the high value is tempered by rarity (a runner thrown out at home is even rarer)

Interpreting the Values

Event_Type	Generic out	K	FC	IBB	BB	HBP	Interference	1B	RBOE	2B	3B	HR
RE24	-0.284	-0.279	-0.233	0.182	0.314	0.335	0.399	0.449	0.491	0.779	1.056	1.38

All values are net runs produced above average

- ➔ A home run is worth +1.38. Why?
  Accounts for frequency of runners on base.
- ➔ Intentional walk worth less than a regular walk
  Typically done when first base is open. Can lead to more force out opportunities. Also, who's getting IBBed?
- ➔ HBP is worth slightly more than a walk

wRAA

wRAA: Weighted Runs Above Average

	Double	Generic out	Hit by pitch	Home run	Intentional walk	PA	Single	Triple	Walk
Barry Bonds (2001)	32	320	9	73	35	664	49	2	177

$$\begin{aligned}122.8 \text{ Runs Above Average} &= 0.778 \cdot 32 - 0.283 \cdot 320 + 0.336 \cdot 9 \\ &+ 1.379 \cdot 73 + 0.17 \cdot 35 + 0.461 \cdot 49 + 1.081 \cdot 2 + 0.306 \cdot 177\end{aligned}$$

LWTS and wRAA are also known as Batting Runs by Pete Palmer

- Pete Palmer fixes the positive weights and lets the out value fluctuate to account for variable run scoring environments

wRAA

If you multiply the weights by frequencies of the events and sum:

Result is (near) 0

Why?

- ➔ When building RE24, we subtract an expectation: we're computing *changes* in expectation
  Also known as the marginal impact
- ➔ The linear weights are telling you about *net runs produced above/below average.*
  An event with weight 0 would maintain average scoring rate
- ➔ Weight tells you the incremental expected run value of an event

wRAA tells a similar story to RE24 but we remove the situational dependence

wOBA

We want to rate a player on events *relative to outs*

BA, OBP, and SLG set the value of an out at 0.

We do the same for wOBA. How?

➔ Compute the difference of the value of an event with the value of an out

$$\text{Relative Value} = \text{LWTS Event Value} - \text{LWTS Out Value}$$

wOBA

One last thing to do for wOBA: the wOBA scale

wOBA scale is used to alter the presentation to make it similar scale to OBP

- ➔ For most people, quoting runs relative to outs per PA is tough to understand
- ➔ We do however have a feel for good, average, and bad OBP values

$$\text{wOBA Value} = (\text{LWTS Event Value} - \text{LWTS Out Value}) \times \text{wOBA Scale}$$

Season	wOBA	wOBAScale	wBB	wHBP	w1B	w2B	w3B	wHR
2017	.321	1.185	.693	.723	.877	1.232	1.552	1.980
2016	.318	1.212	.691	.721	.878	1.242	1.569	2.015
2015	.313	1.251	.687	.718	.881	1.256	1.594	2.065

wOBA

We have the wOBA weights, we can build the wOBA value

This is similar to how we build BA, OBP, or SLG

Event_Type	Walk	Hit by pitch	Single	Double	Triple	Home run
RE24	0.696	0.732	0.879	1.254	1.613	1.965

wOBA weights for 2001 (wOBA scale 1.182)

$$wOBA = (0.696 \cdot BB + 0.732 \cdot HBP + 0.879 \cdot 1B + 1.254 \cdot 2B + 1.613 \cdot 3B + 1.965 \cdot HR)/PA$$

Linear Weights and wOBA

Linear Weights (LWTS)

Incremental change in expected runs by an event. Events like hits are a plus. Outs are naturally minus.

wOBA Weights

Incremental value of an event *relative to an out*.

Is a home run 3x more valuable than a single or 2.2x more valuable?

LWTS says 3, wOBA says 2.2. Which is right?

LWTS. Why?

Linear Weights and wOBA

wOBA weights quantify the tradeoff between the events and an out.

Ex: Convert 2.2 PA from an out to a single to have the same impact as 1 PA to a home run

Better and more usual to think in terms of LWTS values.

Another way to think about them/how to use them:

Adding a walk to a batter's line (and therefore also adding a PA)

Use the LWTS value: ~0.32 runs

Exchanging a walk for an out in a batter's line (keeping total PA the same)

Use the unscaled wOBA value: ~0.62 runs

Classical Stats, LWTS, and wOBA

Let’s revisit the classic stats one last time…

Let’s also recall OPS:

$$OPS = OBP + SLG$$

Classical Stats, LWTS, and wOBA

A little bit of algebra applied to OPS...

$$OPS = OBP + SLG$$

$$\Rightarrow$$

$$OPS = \frac{1}{PA}\left(2 \cdot 1B + 3 \cdot 2B + 4 \cdot 3B + 5 \cdot HR + 1 \cdot BB\right) + \text{Other Stuff}$$

$$\Rightarrow$$

$$OPS = \frac{4}{PA}\left(0.50 \cdot 1B + 0.75 \cdot 2B + 1.00 \cdot 3B + 1.25 \cdot HR + 0.25 \cdot BB\right) + \text{Other Stuff}$$

OPS has a linear component with weights that look kind of familiar!

Classical Stats, LWTS, and wOBA

Weights comparison

	$1B$	$2B$	$3B$	HR	BB	HBP	O
BA	1	1	1	1			
OBP	1	1	1	1	1	1	
SLG	1	2	3	4			
OPS	2	3	4	6	1	1	
$LWTS$	0.48	0.78	1.07	1.40	0.32	0.32	-0.3
$wOBA$	0.879	1.254	1.613	1.965	0.696	0.732	

Classical Stats, LWTS, and wOBA

Weights comparison with a bit of rescaling

	$1B$	$2B$	$3B$	HR	BB	HBP	O
BA	1	1	1	1			
OBP	1	1	1	1	1	1	
SLG	0.333	0.667	1.000	1.333			
OPS	0.500	0.750	1.000	1.250	0.250	0.250	
$LWTS$	0.461	0.778	1.081	1.379	0.306	0.336	−0.283
$wOBA$	0.879	1.254	1.613	1.965	0.696	0.732	

- ➔ SLG, OPS, and LWTS are all close, but only LWTS has a value for an out
- ➔ Add out value and scale by 1.182 to SLG/OPS weights and you get close to wOBA