Runs Expectancy in Baseball

From George Lindsey to an Ordinal Extension

Jim Albert
BGSU and Houston Astros

Albert (2015), *Journal of Sports Analytics*, 1(1), 3–18

Overview

Part I: Foundations

- The 24 base-out states
- George Lindsey (1963)
- The RE Matrix

Part II: Applications

- Valuing plays
- Strategy (bunt, steal, IBB)
- wOBA and RE24

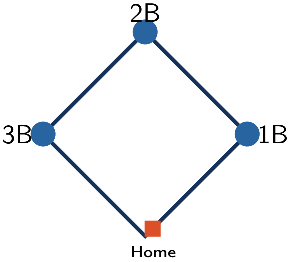
Part III: Beyond RE

- Ordinal regression
- Multilevel modeling
- Covariates & clutch

Part I: Foundations

A **state** = outs + base configuration

- 3 out counts × 8 base configurations
- **24 distinct states**



Examples

```
(1--, 0 outs)   runner on 1st
(-23, 1 out)     2nd & 3rd
(123, 2 outs)   bases loaded
```

George Lindsey (1963)

- Canadian defense researcher and baseball enthusiast
- First to study **run distributions** by game state
- Laid the groundwork for all modern RE analysis

Lindsey's Insight

Strategy decisions can be evaluated by tracking run distributions across states.

Key Papers

Lindsey (1961)
Operations Research

Lindsey (1963)
Operations Research

$$\mathrm{RE}(b, o) = \mathbb{E}\big[\text{runs scored to end of inning} \mid \text{state } (b, o)\big]$$

- b = base configuration, o = number of outs
- Estimated empirically from Retrosheet play-by-play data
- **24 values** form the Runs Expectancy Matrix

The RE Matrix (2019 Season)

Bases	0 Outs	1 Out	2 Outs
--	0.51	0.27	0.10
1--	0.88	0.52	0.22
-2-	1.14	0.71	0.33
12-	1.47	0.93	0.45
--3	1.35	0.97	0.38
1-3	1.77	1.18	0.51
-23	2.05	1.42	0.59
123	2.33	1.61	0.78

Computing RE from Data

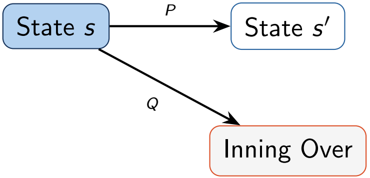
1. Get Retrosheet play-by-play files
2. Track runs scored from each state to end of half-inning
3. Average within each of the 24 states

Data Source

Retrosheet: free play-by-play data back to the 1950s. Typically pool multiple seasons for stable estimates.

- The 24 states + *inning over* form a **Markov chain**
- RE satisfies the recursion:

$$\mathrm{RE}(s) = \sum_{s'} P(s \to s') \cdot \left[r(s, s') + \mathrm{RE}(s')\right]$$



A 60-Year History

Year	Milestone
1963	Lindsey: run distributions by state
1984	Thorn & Palmer: *The Hidden Game of Baseball*
2003	Albert & Bennett: *Curve Ball*
2007	Tango, Lichtman & Dolphin: *The Book*
2015	Albert: *Beyond Runs Expectancy*, JOSA

Part II: Applications

The Core Formula

$$\Delta\mathrm{RE} = \mathrm{RE}(s_{\text{after}}) - \mathrm{RE}(s_{\text{before}}) + \text{runs scored}$$

- $\Delta\mathrm{RE} > 0$: play **helped** the offense
- $\Delta\mathrm{RE} < 0$: play **hurt** the offense

Example: Single, runner 1B $\to$ 3B, 0 outs:

$$\Delta\mathrm{RE} = 1.35 - 0.88 = +0.47$$

Linear Weights

Average $\Delta\mathrm{RE}$ over all instances of each event type:

Event	**Run Value**
Home Run	$+1.40$
Triple	$+1.05$
Double	$+0.78$
Single	$+0.47$
Walk/HBP	$+0.32$
Out	-0.28

Scale run values so that **league-average wOBA $\approx$ league OBP**:

$$\mathrm{wOBA} = \frac{w_{BB} \cdot BB + w_{1B} \cdot 1B + w_{2B} \cdot 2B + \cdots}{\mathrm{PA}}$$

- Weights derived from RE linear weights
- More informative than BA, OBP, or SLG alone

RE24

Sum $\Delta\mathrm{RE}$ across all plate appearances in a season:

- **Context-sensitive**: a hit with RISP is worth more
- RE24 $= +22.4$ means $+22.4$ runs added above average
- Unlike wOBA, RE24 rewards clutch performance

The Sacrifice Bunt

State change: $(1\text{--}, 0) \to (\text{-}2\text{-}, 1)$

$$\Delta\mathrm{RE} = 0.71 - 0.88 = -0.17$$

The bunt **costs** 0.17 runs on average.

Only justified in specific late-game situations.

Runner on 1st, 0 outs
RE $=$ 0.88

$\big\downarrow\; -0.17$

Runner on 2nd, 1 out
RE $=$ 0.71

Stolen Base Break-Even

Set expected $\Delta\mathrm{RE} = 0$ and solve for p:

$$p \cdot [\mathrm{RE}(\texttt{-2-}, o) - \mathrm{RE}(\texttt{1--}, o)] = (1 - p) \cdot [\mathrm{RE}(\texttt{1--}, o) - \mathrm{RE}(\texttt{---}, o{+}1)]$$

Situation	**Break-even success rate**
0 outs	$\approx 70\%$
2 outs	$\approx 75\%$

The Intentional Walk

State change: $(\text{-2-}, 0) \to (\text{12-}, 0)$

$$\Delta\mathrm{RE} = 1.47 - 1.14 = +0.33 \text{ runs given up}$$

Worth it only if the win-probability gain from bypassing a dangerous hitter exceeds the RE cost.

Leverage Index

$$\mathrm{LI} \approx \frac{\sigma^2(\text{WE change in current state})}{\sigma^2(\text{average across all states})}$$

- Measures **how much this PA matters** for winning
- Built on win expectancy, which is built on RE
- Used to deploy relievers and measure clutch context

Inherited Runners & Lineup

Inherited Runners

- Reliever RE24 penalizes stranded runners that score
- Separates starter vs. reliever responsibility

Lineup Construction

- On-base ability most valuable at the top
- Power most valuable with runners on
- Optimal lineups via RE-based simulation

RE values shift with the run environment:

Era	Characteristic
Dead Ball (1910s)	Low RE; few extra-base hits
Offensive Boom (1990s–2000s)	RE peaked league-wide
Three True Outcomes (2015–)	Higher RE empty bases; lower with runners on

Part III: Beyond RE — Albert (2015)

The Fundamental Limitation

$$\mathrm{RE}(\text{-23, 1}) = 1.42$$

This tells us the **mean**. But managers need:

- $P(\text{score} = 0)$ — squeeze play decision
- $P(\text{score} \geq 2)$ — walk-off situation

The mean alone cannot answer these questions.

Why Not a Standard Count Distribution?

Run-scoring has an **irregular shape**:

- Heavy mass at 0 and 1
- Long right tail

Poisson	requires mean = variance	$\times$
Neg. binomial	still fits poorly	$\times$
Zero-inflated	adds complexity, still fails	$\times$

The Ordinal Regression Approach

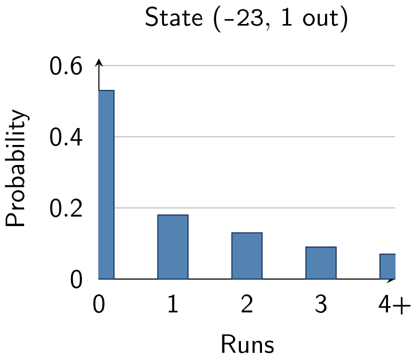
Treat runs as an **ordered categorical** outcome: $0, 1, 2, 3, 4+$

$$\log\left[\frac{P(Y \leq k \mid \mathrm{x})}{P(Y > k \mid \mathrm{x})}\right] = \alpha_k - \boldsymbol{\beta}^\top \mathrm{x}$$

- α_k: thresholds (one per category boundary)
- $\boldsymbol{\beta}$: single coefficient vector (proportional odds)

The ordinal model gives $P(Y = k \mid \text{state})$ for every k:

- $P(\text{score} \geq 1)$: probability of any run
- $P(\text{score} \geq 2)$: probability of a big inning
- Median, quantiles — not just mean



The Runs Advantage Table

The ordinal coefficient $\beta_{(b,o)}$ measures how a state **shifts the run distribution upward**.

	RE Matrix	**Runs Advantage**
Quantity	Mean runs	Ordinal coefficient
Information	Mean only	Full distribution
120 vs 003	$120 > 003$	$003 > 120$
Best use	General valuation	Tail-probability decisions

Runner on 3rd alone has a **higher probability of scoring exactly 1 run** than runners on 1st and 2nd.

- RE says $120 > 003$ (higher *expected* runs)
- Runs Advantage says $003 > 120$ (higher $P(\text{any run})$)

The distributional view matters for squeeze plays, walk-offs, and one-run game strategy.

Estimate all 30 teams **simultaneously**:

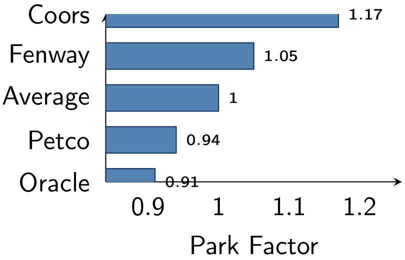
$$\mu_j \mid \mu, \sigma \;\sim\; \mathcal{N}(\mu, \sigma^2), \quad j = 1, \ldots, 30$$

- Treats team abilities as **exchangeable**
- Shrinks extreme estimates toward the league mean
- Separates signal from noise

Covariate 1: Ballpark Effects

$$\eta = \mu_j + \beta_{\text{park}}$$

- Adjusts for park dimensions, altitude, humidity
- Separates true team quality from park inflation



Covariate 2: Pitcher Quality

- Modeled via opponent ERA or FIP
- Elite pitcher $\rightarrow$ run distribution shifts **left**
- Enables **opponent-adjusted** team offense ratings

Covariate 3: Clutch / RISP

Does performance with **runners in scoring position** differ?

- RISP indicator added as a covariate
- Exchangeable priors shrink spurious team effects toward zero

Finding

Most team RISP variation is **random noise**.

Each covariate shifts **all** log-odds by the same amount β:

$$\log \frac{P(Y \leq k)}{P(Y > k)} = \alpha_k - \beta x$$

- One β per covariate regardless of category k
- Parsimonious — far fewer parameters than unconstrained
- Albert finds this **well-supported** by MLB data

Summary: Albert (2015)

1. **Full distribution** of runs scored, not just the mean
2. **Ordinal regression** — parsimonious, well-fitting
3. **30-team multilevel model** with shrinkage
4. **Covariates**: ballpark, pitcher quality, RISP
5. **Runs advantage table** — distributional RE counterpart
6. Much team RISP variation is noise

Open Questions

Modeling

- Relax proportional-odds assumption
- Add batter identity
- Dynamic in-season updating

Applications

- Ordinal-based wOBA weights
- Strategy with distributional RE
- Real-time leverage index

Conclusion

- The RE Matrix: **60 years** at the core of sabermetrics
- Powerful for play valuation, strategy, and metrics
- Its limitation: **the mean hides the distribution**
- Ordinal regression + multilevel models provide the full picture

RE Matrix	gives the *mean*
Beyond RE	gives the *distribution*
Multilevel	accounts for *context*

"The runs expectancy table tells you the average.
The ordinal model tells you the story."