



# Baseball VI



Produced by Dr. Mario | UNC STOR 390



# Player Win Averages

- Let  $p$  = Probability My Team Wins (%)
- Let  $q = 100 - p$  = Probability Opponent's Team Wins
- How Does an Individual Player Impact  $p$ ?
- Winning Probability Difference (WINDIFF)

$$WINDIFF_t = p_t - q_t \text{ where } t = \text{Time}$$

- WINDIFF Before Game Begins
  - Assume Each Team Equally Likely to Win

$$WINDIFF_0 = 50 - 50 = 0$$





# Player Win Averages

- Jeff Sagarin
  - Developed Ranking Methods in Variety of Sports
  - Publishes Rankings in USA Today
  - Known for MLB Player Win Average Analysis (1957-2006)
  - To Avoid Decimals,

$$SAGWINDIFF_t = 10(p_t - q_t) \text{ where } t = \text{Time}$$

- Scenario
  - Home Team Losing by 2 Runs in Bottom of 9th
  - Bases Loaded and Mariano Rivera Comes in to Pitch
  - Current WINDIFF

$$SAGWINDIFF_t = 10(52.3 - 47.7) = 46$$

- Batter Hits into Double Play and 1 Run Scores
- Next WINDIFF

$$SAGWINDIFF_{t+1} = 10(17.2 - 82.8) = -656$$





# Player Win Averages

- Scenario
  - Observe the Change in WINDIFF

$$SAGWINDIFF_{t+1} - SAGWINDIFF_t = -656 - 46 = -702$$

- Clearly, Swing Was in Mariano Rivera's Favor
  - Batter Loses 702 Points
  - Mariano Gains 702 Points
- How Does This Methodology Improve ERA?

- Key Conversion: 2000 SAGWINDIFF = 1 WIN Above 500
  - Changes in SAGWINDIFF For Team = 1000 Points
  - Suppose Team Record is 82W and 80L
  - Across Season

$$82 \times 1000 - 80 \times 1000 = 2000 \text{ Points}$$

- Team Ends Season 1 Game Over .500





# Player Win Averages

- Win Average Leaders from 2004 to 2006

Sagarin Win Average Leaders, 2004–6

Year	Player	Position	Total Points	Situations
2006	Albert Pujols	batter (outfield)	+18,950	653
2006	Francisco Rodriguez	relief pitcher	+10,562	312
2005	David Ortiz	outfielder/ designated hitter	+18,145	718
2005	Roger Clemens	starting pitcher	+12,590	852
2004	Barry Bonds	outfielder	+25,398	637
2004	Brad Lidge	relief pitcher	+11,906	382





# Player Win Averages



- Applied to Fielding Ability
  - Use of Rating System by John Dewan's *Fielding Bible*
  - Derek Jeter's Rating was -34 (Caused 34 Hits More Than Average)
  - Derek Jeter Cost the Team

$$\begin{aligned}-34 \text{ Hits} &= -0.8 \times 34 \text{ Runs} = -27.2 \text{ Runs} \\ &= -\frac{27.2}{10} \text{ Wins} = 2000 \times 2.72 \text{ Win Points} \\ &= -5,440 \text{ Win Points}\end{aligned}$$

- Adjust Derek Jeter's Win Points by Subtraction
- Historically, Fielding Has Been Overrated
- Based on *Fielding Bible*, only 7 players have positive fielding ratings equivalent to 2 more wins above average.





# Player Win Averages

- Applied to Baserunning Ability
  - Good Base Runner Described
    - Rarely Caught Stealing
    - Rarely Caught in Double Plays
    - Able to Take Extra Bases
  - Player Win Averages Reward Stolen Bases and Preventing Double Plays
  - Player Win Averages Do Not Reward
    - Fast Base Running is not Rewarded by Player Win Averages
  - Analyzed By Dan Fox (Director Of Baseball Informatics for Pirates)
  - Compare Runner's Number of Runs to Average Runner
  - Best Base Running Has Little Effect on SAGWIN





# Player Win Averages

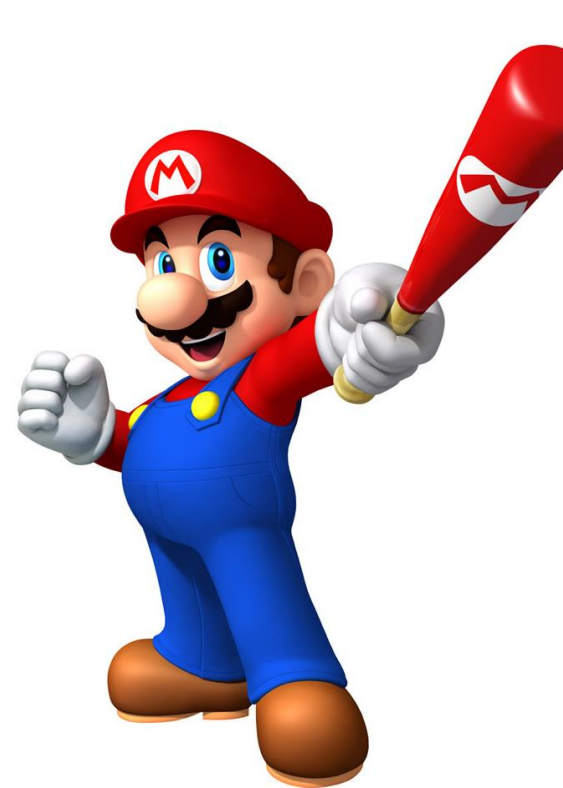
- Interesting Application: Hitting Versus Pitching
  - 1969 World Champion New York Mets
  - Table of Player Win Averages
    - Hitters = Bad
    - Pitchers = Good
  - Season 100 Wins
  - Expect 38,000 Win Points
- Not Equal Because Trades

$$(100 - 81) \times 2000 = 38,000$$

	Hitters total		Pitchers total	Team total
	-9319		41057	3173
Batter	Winpoints	Pitcher	Winpoints	
Grote	-1960	Seaver	13471	
Kranepool	-765	Koosman	13218	
Boswell	-42	Cardwell	761	
Garrett	-3819	McAndrew	1332	
Harrelson	-131	Ryan	23	
Agee	5410	McGraw	10902	
Jones	6334	Koonce	185	
Swoboda	6278	DiLauro	890	
Shamsky	4998	Taylor	1625	
Weis	-2054	Frisella	-605	
Gaspar	-1772	Jackson	-745	
Pfeil	-3629			
Clendenon	240			
Martin	-2196			
Charles	-1021			
Otis	-1789			
Dyer	-290			







# Player Win Averages

---

- Winning Probabilities for All Game Scenarios
  - Recall:  $p$  = Probability My Team Wins (%)
  - How is this Probability Calculated?
  - We Can Use a Massive Set of Play-by-Play Data (1977-2006)
  - Scenarios Based on Inning, Score, Runner Locations, and Outs
  - Problem Some Scenarios Are Rare
  - Advised to Use Markov Chain Monte Carlo (MCMC)





# Value of Replacement Players

- Winning Probabilities for All Game Scenarios
  - Recall:  $p$  = Probability My Team Wins (%)
  - How is this Probability Calculated?
  - We Can Use a Massive Set of Play-by-Play Data (1977-2006)
  - Scenarios Based on Inning, Score, Runner Locations, and Outs
  - Problem Some Scenarios Are Rare
  - Advised to Use Markov Chain Monte Carlo (MCMC)





# Value of Replacement Players

- **Keith Woolner**
  - Works for the Cleveland Indians
  - Authored for *Baseball Prospectus*
  - Respected for Player Analysis and Market Evaluation
  - Created Value of Replacement Player (VORP)
- **Replacement Player**
  - Players Create Value by Keeping Bad Players Off the Field
  - Acquiring Replacement Players (Minor League)
  - Assume List of Replacement Players is Infinite
  - Woolner Defined Replacement Player as a Player in the Bottom 20% of List Ordered by Plate Appearances
  - Team of Replacement Players Would Get Approximately 44 Wins and 118 Losses





# Value of Replacement Players

PA = Plate  
Appearance  
BFP = Batters Faced

- Value of a Replacement Player Points (VORPP)

- Recall: Each Loss Below .500 = -2,000 SAGWIN
- For All Replacement Players = -74,000 SAGWIN
- Batters and Pitchers Get Equal Blame = -37,000 SAGWIN Points Each
- Approximately 6,200 Plate Appearances in a Season

$$\frac{-37,000}{6,200} = -5.97 \text{ Points Per Plate Appearance}$$

- Formula for Hitters

$$VORPP = SAGWIN + 5.97 \times (PA)$$

- Formula for Pitchers

$$VORPP = SAGWIN + 5.97 \times (BFP)$$





# Value of Replacement Players

- VORPP Used to Evaluate Trades

$$\text{Hoffman VORPP} = 7,963 + 5.97(255) = 9,485.$$

$$\text{Young VORPP} = 6,117 + 5.97(781) = 10,780.$$

- VORPP Used to Determine Salary

$$\text{Pujols 2006 value} = 22,848 \times 1,040 = \$23.8 \text{ million.}$$

$$\text{Hoffman 2006 value} = 9,485 \times 1,040 = \$9.9 \text{ million.}$$

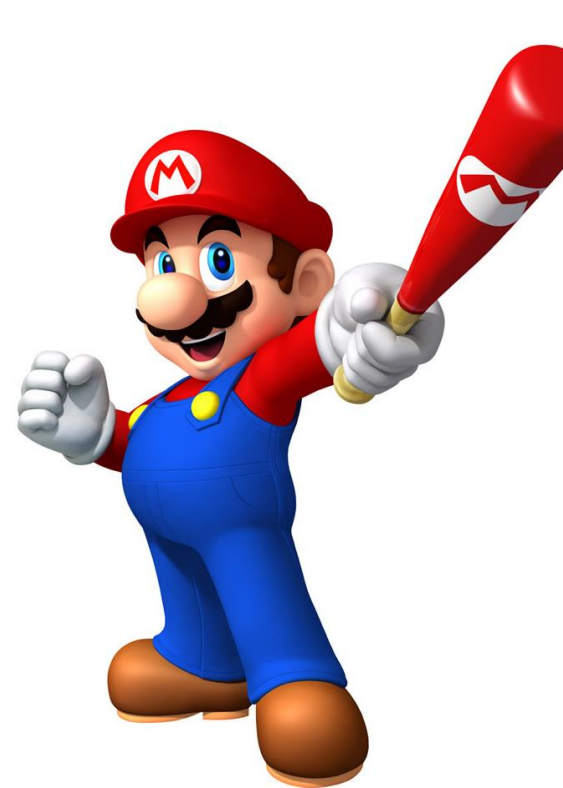
$$\text{Young 2006 value} = 10,780 \times 1,040 = \$11.1 \text{ million.}$$

$$\$77M = 74,000 \text{ VORPP}$$

$$\$1,040 = 1 \text{ VORPP}$$







# Value of Replacement Players

- Was Alex Rodriguez Overpaid by Yankees? Yes
  - Deal: \$275M for 10 Years

Year	Plate Appearances	SAGWIN Points	VORP Points	Fair Salary in 2006 Dollars (millions)
2003	741	10593	15016.77	\$ 15.62
2004	734	6575	10956.98	\$ 11.40
2005	752	12521	17010.44	\$ 17.69
2006	695	2205	6354.15	\$ 6.61
2007	736	14193	18586.92	\$ 19.33

- Fair Deal Based Off 2007 Statistics  
 $\$19.33M \times 1.05 \times 10 \approx \$202M$





# Value of Replacement Players

- Extra Plate Appearances Create Value

$$\text{Inge VORPP} = 15 + 5.97(617) = 3,699.$$

$$\text{Therriot VORPP} = 2,480 + 5.97(174) = 3,519.$$

- VORPP Used by Theo Epstein in 2004 Red Sox Season





# Final Inspiration

Yesterday's home runs don't  
win today's games.

-Babe Ruth