236B Zachry Engineering Center
3131 TAMU
College Station, Texas 77843
979.845.4347
ebickel@tamu.edu

Applying OR across the Baseball Decision Hierarchy

Prepared by:
J. Eric Bickel
Assistant Professor
Industrial and Systems Engineering
Texas A&M University

Presented to:
Symposium on Statistics and
Operations Research in Baseball
California State Univ., East Bay
Chair: Mitch Watnik
Hayward, California

11 July 2007

Abstract

Operations research can be applied at all levels of baseball decision making from the field to the front office. In this talk I will present two examples spanning this range of application. At the field level, I will consider the decision of whether or not to take a particular pitch. For example, should a batter take on 3-0? Why? What about 2-0? From the perspective of the front office and player valuation, I will discuss a new method to determine how much a particular starting pitcher increases his team's chance of winning. For example, by how much does Roger Clemens increase the Yankees' chance of winning in games he pitches? These examples illustrate the breadth and power of operations research and highlight its ability to improve tactical and strategic decision making in baseball and beyond.

Decision analysis can bring clarity of action to all levels of the baseball decision hierarchy.

Front Office

- Which players should a team draft/trade
- Player valuation
- Creating a portfolio of players
  - Power vs. speed
  - Pitching vs. hitting

Managerial (ball is not in play)

- Design of optimal batting order
- Intentional walk
- Hit, Steal, Sacrifice Bunt, Hit & Run
- Pitch Count Strategy

Players (while ball is in play)

- Attempt a double play allowing a run to score
- Catch a foul ball allowing a runner to score

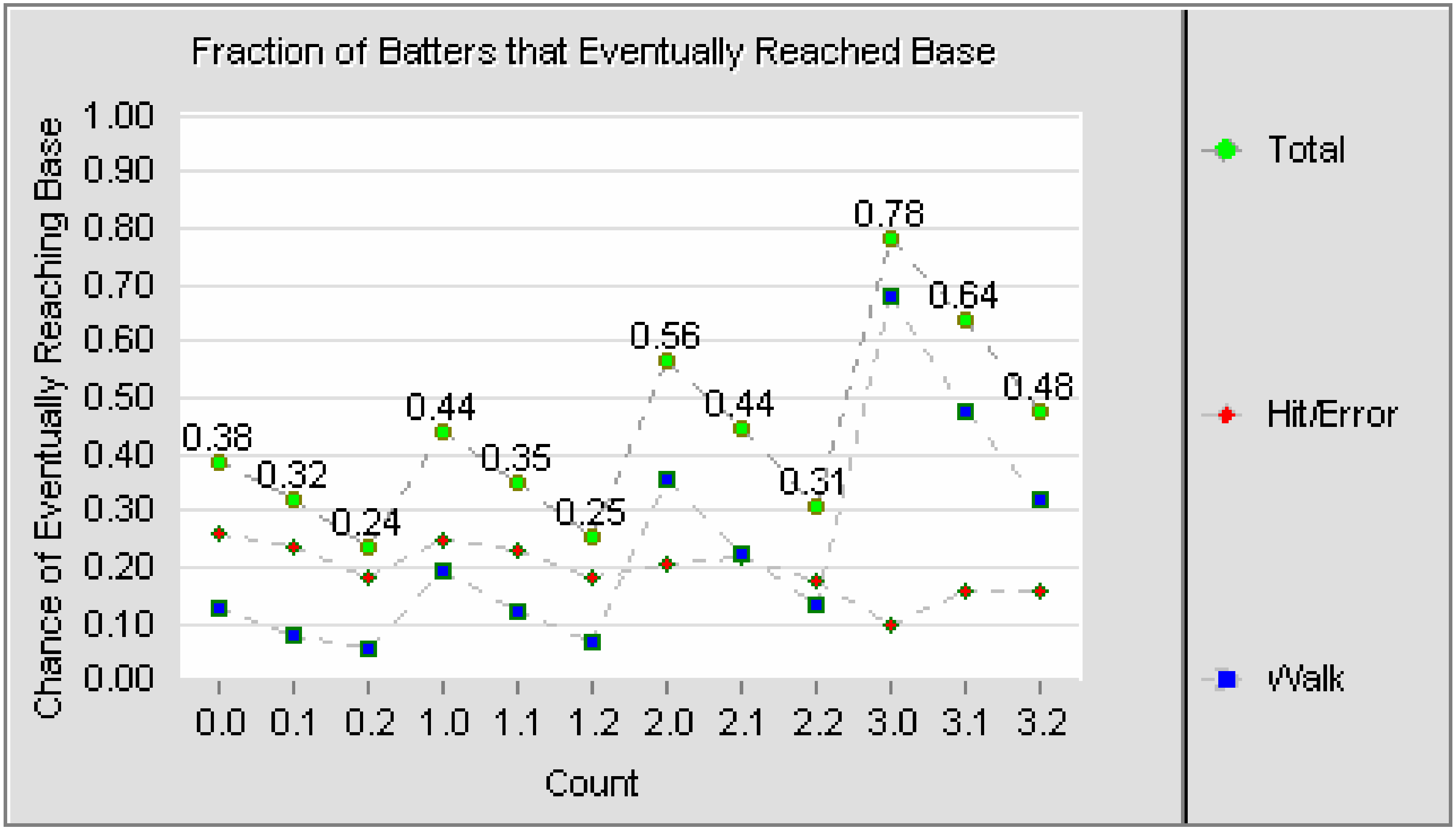
This analysis grew out of work I did with the Stanford baseball team.

Background

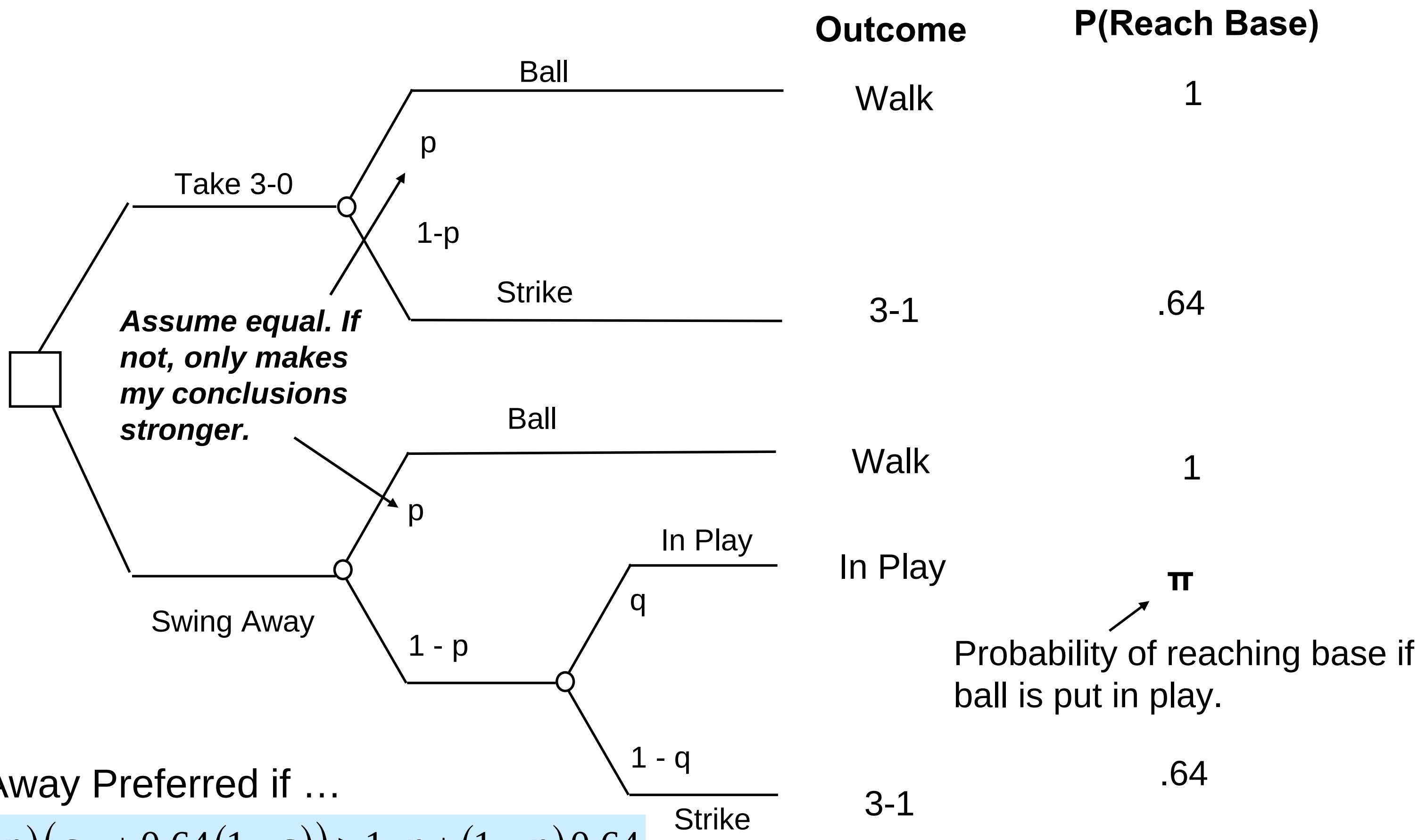
- Developed this model while in graduate school.
  - Wanted to help Stanford beat USC, which they did—again and again!
- Built a Markov model of pitch count dynamics.
  - Assessed key inputs from Dean Stotz (Associate Head Coach, Stanford Baseball).
  - Compared model results to actual game statistics.
- This led to the development of a pitch/hit charting decision support system called ChartMine® (www.chartmine.net).
  - Dean and I founded Competitive Edge Decision Systems to market ChartMine.
  - Used by over 300 teams, including 30% of NCAA Division 1
  - Named the best baseball product by Collegiate Baseball Newspaper for three straight years.

For the rest of this talk, I am going to use actual game data (about 150,000 pitches), instead of a Markov model.

ChartMine® Screen Shot



Let's look at the decision of whether or not to take the 3-0 pitch.



Swing Away Preferred if …

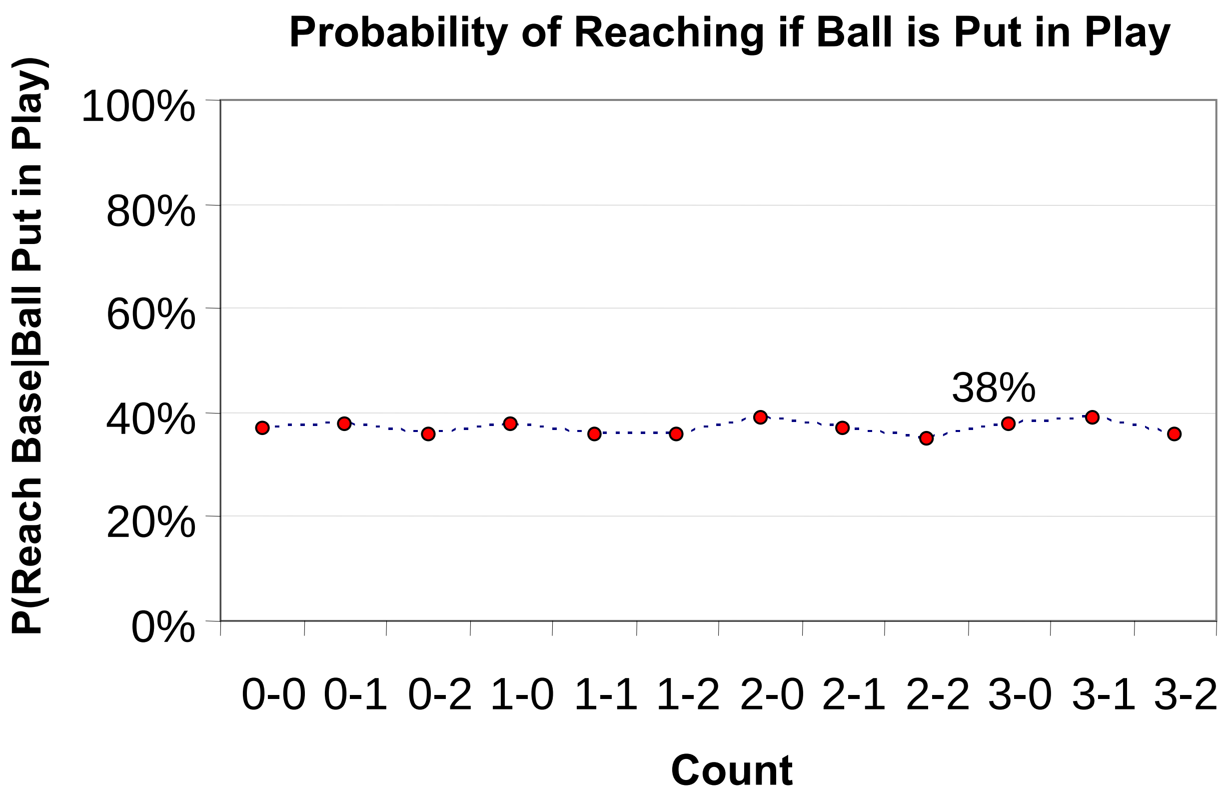
$$1 \cdot p + (1-p)\big(q\pi + 0.64(1-q)\big) > 1 \cdot p + (1-p)0.64$$

$$q\pi + 0.64(1-q) > 0.64$$

$$q(\pi - 0.64) > 0$$

$$\pi > 0.64$$

Batters have about a 38% chance of reaching base if they put the ball in play on 3-0.

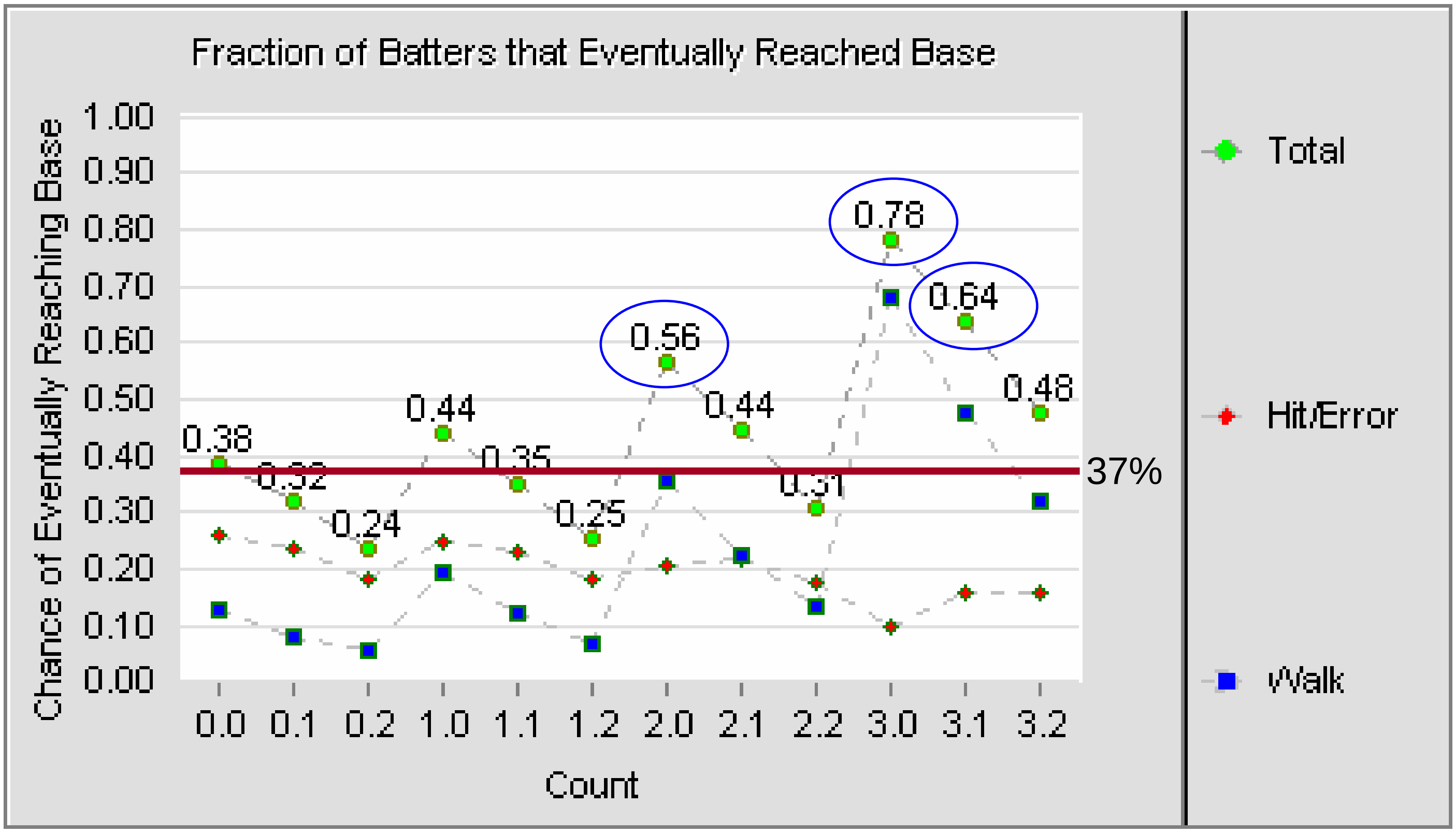


A good rule of thumb is that about 37% of all batters that put the ball in play reach base (33% get a hit and 4% get on with an error).

Therefore, the batter should take 3-0, even if he/she knows the pitch is going to be a strike. As soon as the ball is put in play the chance the batter will reach base has been reduced from at worst 64% (if a strike is taken) to 38%!

If batters are trying to maximize the chance of reaching base, they should take 2-0, 3-0, and 3-1.

When to Take Pitches



We can summarize this analysis in a decision table.

Decision Table for Taking a Strike Based on Reaching Base

Current Count	Probability of Eventually Reaching Base if a Strike is Taken	Probability of Reaching if Ball is Put in Play	Difference	Take?
0-0	32% (new count: 0-1)	37%	-5%	No
0-1	24% (0-2)	38%	-14%	No
0-2	0% (K)	36%	-36%	No
1-0	35% (1-1)	38%	-3%	No
1-1	25% (1-2)	36%	-11%	No
1-2	0% (K)	36%	-36%	No
2-0	**44% (2-1)**	**39%**	**+5%**	**Yes**
2-1	31% (2-2)	37%	-6%	No
2-2	0% (K)	35%	-35%	No
3-0	**64% (3-1)**	**38%**	**+26%**	**Yes**
3-1	**48% (3-2)**	**39%**	**+9%**	**Yes**
3-2	0% (K)	36%	-36%	No

Might our conclusions change if the objective was to maximize the average number of bases obtained?

No! Batters should take 2-0, 3-0, and 3-1 if they are trying to maximize expected bases.

Decision Table for Taking a Strike Based on Expected Bases

Current Count	Average Bases Eventually Obtained if a is Strike Taken	Average Bases Obtained if Ball is Put in Play	Difference	Take?
0-0	0.43 (new count: 0-1)	0.47	-0.04	No
0-1	0.32 (0-2)	0.47	-0.15	No
0-2	0.00 (K)	0.44	-0.44	No
1-0	0.46 (1-1)	0.50	-0.04	No
1-1	0.33 (1-2)	0.46	-0.13	No
1-2	0.00 (K)	0.43	-0.43	No
2-0	**0.56 (2-1)**	**0.52**	**+0.04**	**Yes**
2-1	0.40 (2-2)	0.47	-0.07	No
2-2	0.00 (K)	0.44	-0.44	No
3-0	**0.73 (3-1)**	**0.53**	**+0.20**	**Yes**
3-1	**0.58 (3-2)**	**0.50**	**+0.08**	**Yes**
3-2	0.00 (K)	0.47	-0.47	No

Does that mean batters should always take 2-0, 3-0 and 3-1?

No! Taking is never optimal if you are trying to maximize the probability of a hit.

Decision Table for Taking a Strike Based on Getting a Hit

Current Count	Probability of Eventually Getting a Hit if Strike is Taken	Probability of Getting a Hit if Ball is Put in Play	Difference	Take?
0-0	0.24 (new count: 0-1)	0.34	-0.10	No
0-1	0.18 (0-2)	0.34	-0.16	No
0-2	0.00 (K)	0.33	-0.33	No
1-0	0.23 (1-1)	0.36	-0.13	No
1-1	0.18 (1-2)	0.33	-0.15	No
1-2	0.00 (K)	0.31	-0.31	No
2-0	**0.22 (2-1)**	**0.37**	**-0.15**	**No**
2-1	0.18 (2-2)	0.33	-0.15	No
2-2	0.00 (K)	0.32	-0.32	No
3-0	**0.16 (3-1)**	**0.33**	**-0.17**	**No**
3-1	**0.16 (3-2)**	**0.36**	**-0.20**	**No**
3-2	0.00 (K)	0.34	-0.34	No

What can we conclude about taking pitches?

Conclusions

- Batters should not “take a strike”, i.e, taking 0-0 or 1-0.
- When runners are needed (max prob reach base) it makes sense to take 2-0, 3-0, and 3-1.
- Early in the game where bases are proportional to runs, batters should take 2-0, 3-0, and 3-1.
- Batters should not take a pitch when a hit is needed (to drive in a runner on second or third, for example).

Decision analysis can bring clarity of action to all levels of the baseball decision hierarchy.

Front Office

- Which players should a team draft/trade
- Player valuation
- Creating a portfolio of players
  - Power vs. speed
  - Pitching vs. hitting

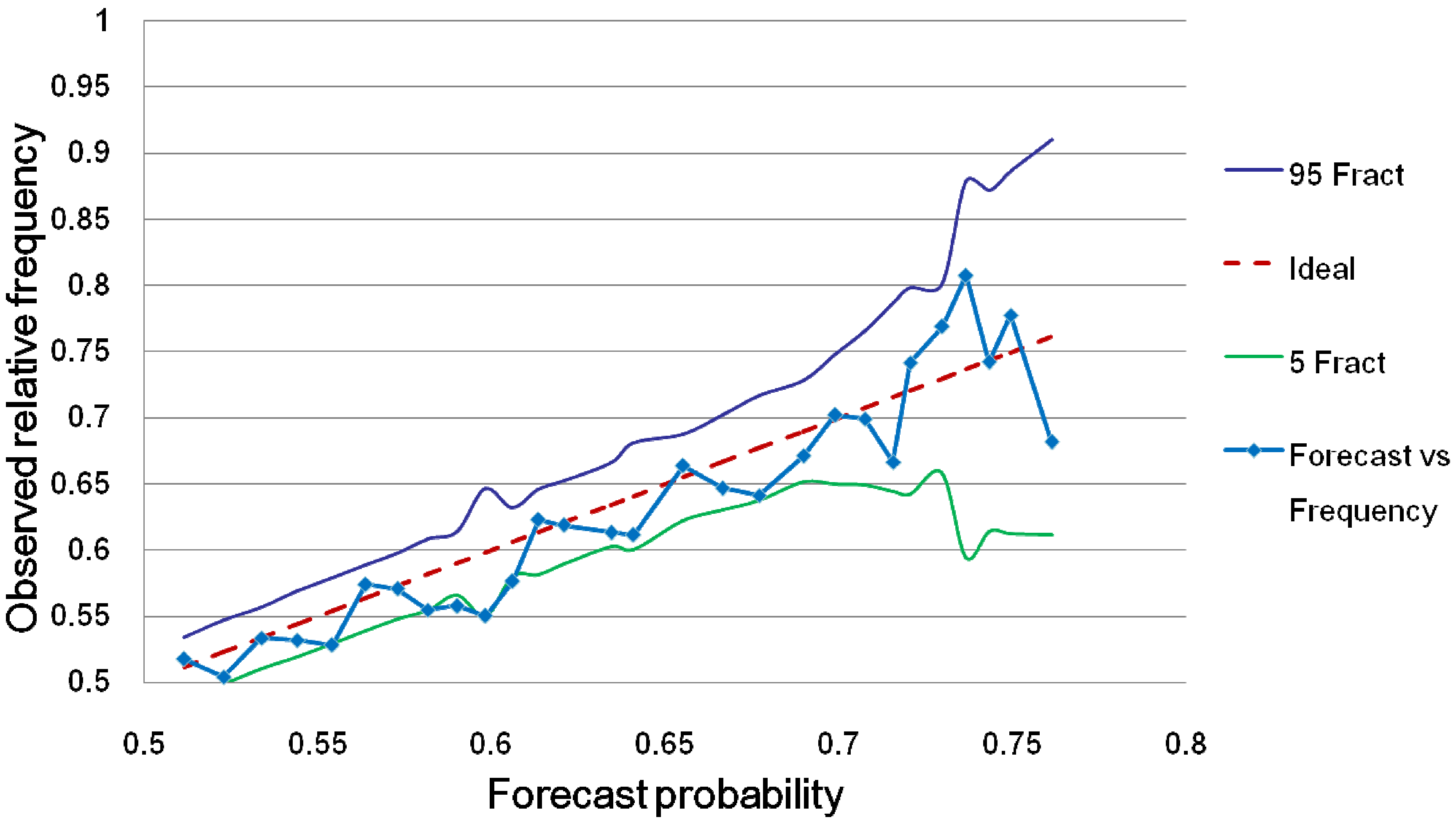
Managerial (ball is not in play)

- Design of optimal batting order
- Intentional walk
- Hit, Steal, Sacrifice Bunt, Hit & Run
- Pitch Count Strategy

Players (while ball is in play)

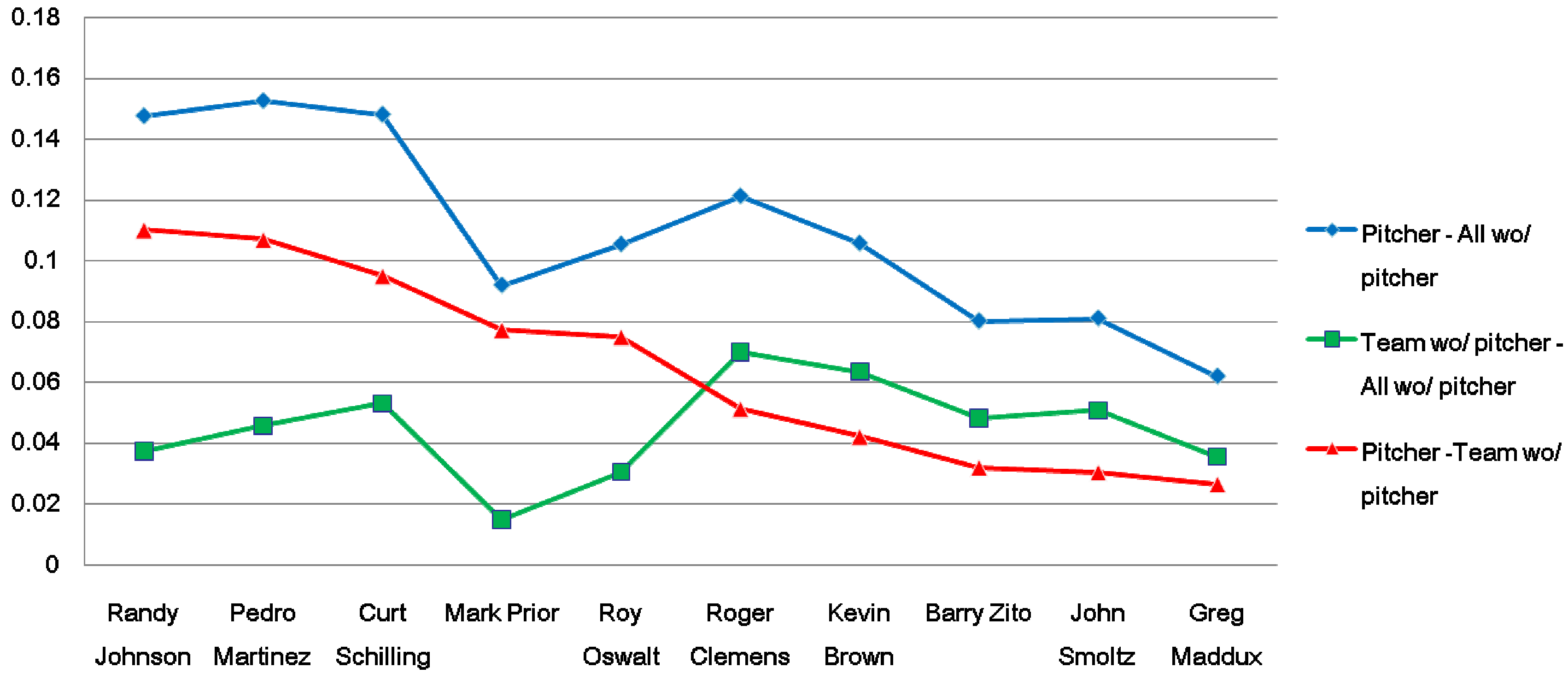
- Attempt a double play allowing a run to score
- Catch a foul ball allowing a runner to score

Calibration (Hilton 2000-2006 with final odds)



* 16,185 games are analyzed

TOP 10 Pitchers' ability



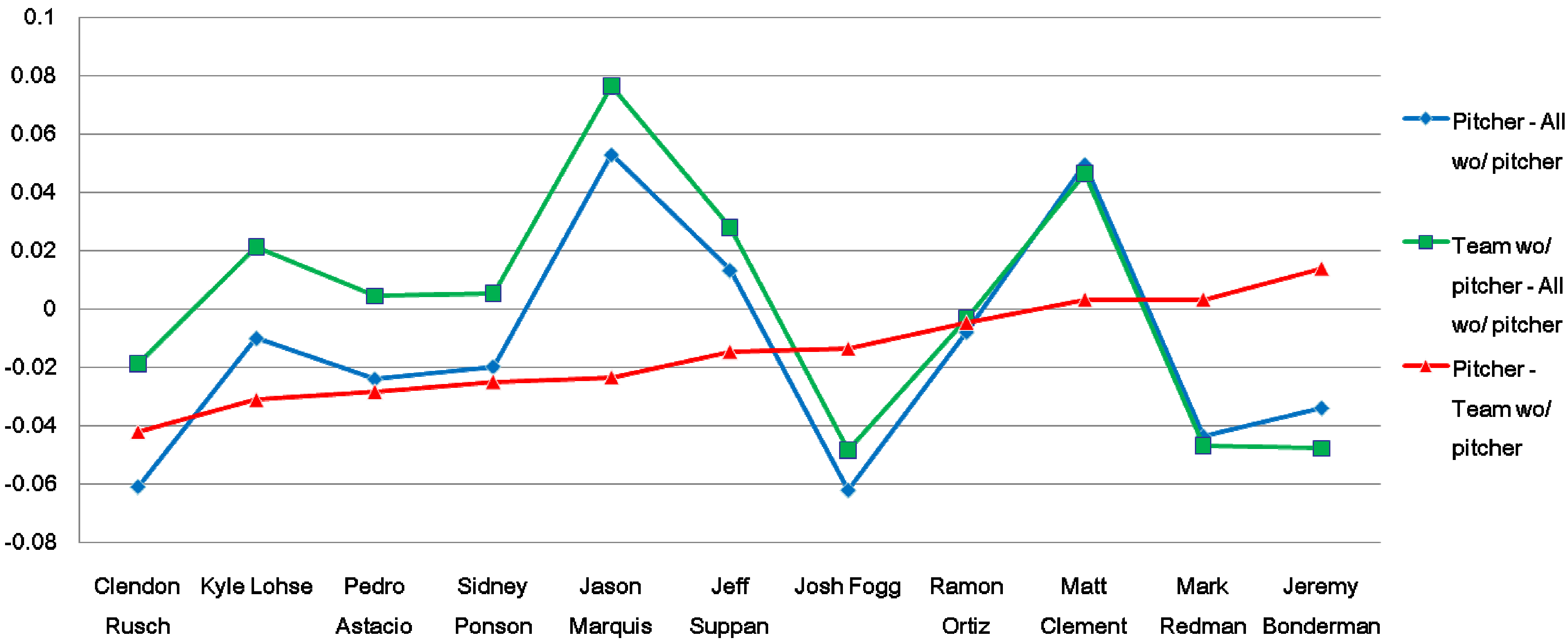
Pitcher	169	149	160	147	168	172	85	189	69	187
All wo/ pitcher	12963	12976	12967	11987	12959	12963	13005	12951	13010	12952
Team wo/ pitcher	716	742	735	700	715	743	626	687	816	853

$$Pitcher\ : P\left(win \,|\, the\ \ pitcher\ \ is\ started\right)$$

$$All\ \ wo\ /\ \ pitcher\ \ : P\left(win \,|\, the\ \ pitcher\ \ is\ not\ started\right)$$

$$Team\ \ wo\ /\ \ pitcher\ \ : P\left(win \,|\, the\ \ pitcher\ 's\ team\ \ plays\ , the\ \ pitcher\ \ is\ not\ started\right)$$

Worst 10 Pitchers' ability



Pitcher	102	156	83	136	130	176	147	152	147	151	124
All wo/ pitcher	12995	12971	13002	12974	12980	12952	12972	12967	12976	12968	13000
Team wo/ pitcher	758	880	1201	1603	783	887	709	720	729	789	531

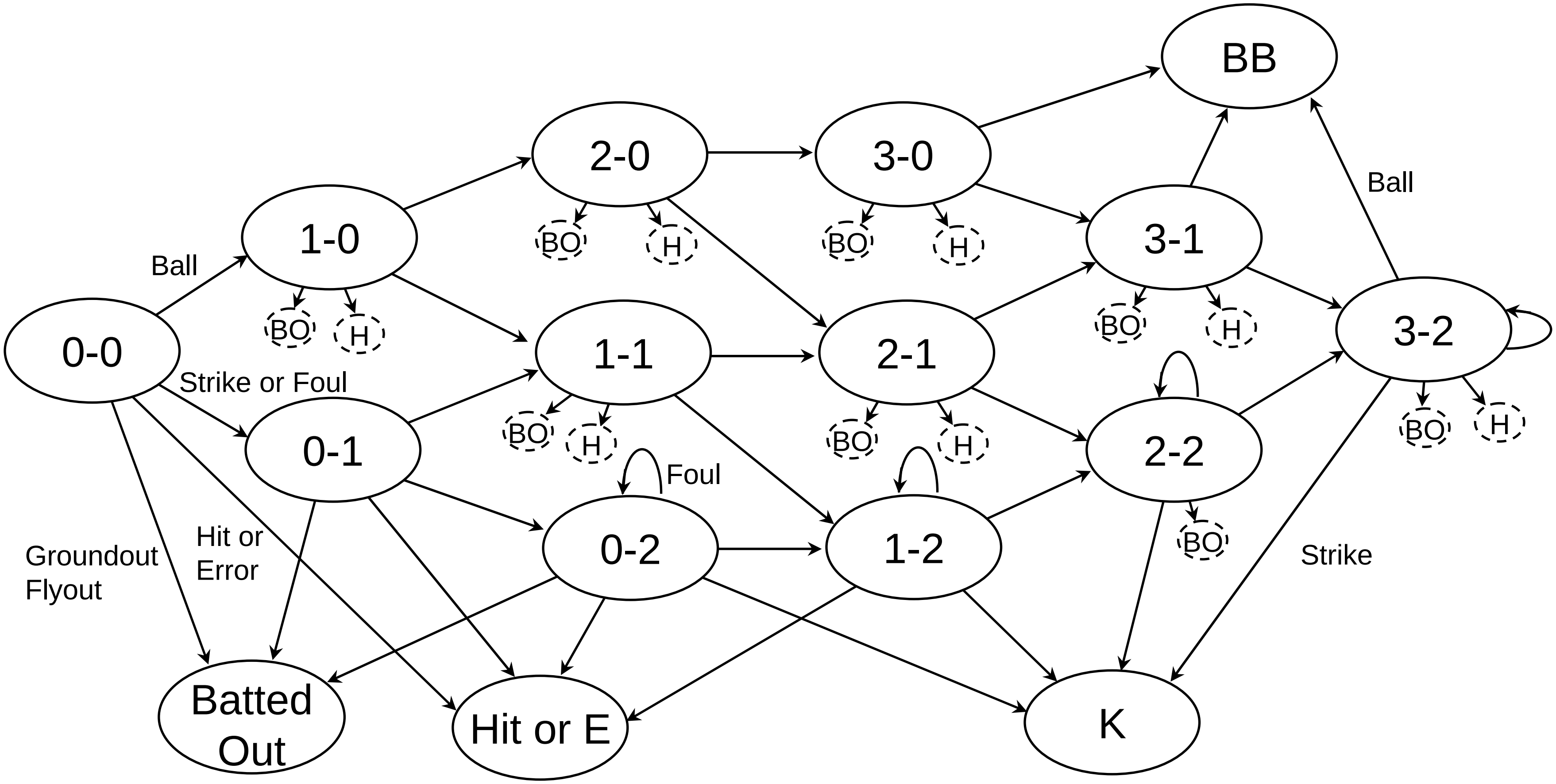
$$Pitcher\ :P\left(win\,|the\ pitcher\ is\ started\right)$$

$$All\ wo\,/\,pitcher\ :P\left(win\,|the\ pitcher\ is\ not\ started\right)$$

$$Team\ wo\,/\,pitcher\ :P\left(win\,|the\ pitcher\ 's\ team\ plays\ ,the\ pitcher\ is\ not\ started\right)$$

Appendix

We can model the chance of moving to different pitch counts, given the current count, as a Markov process.



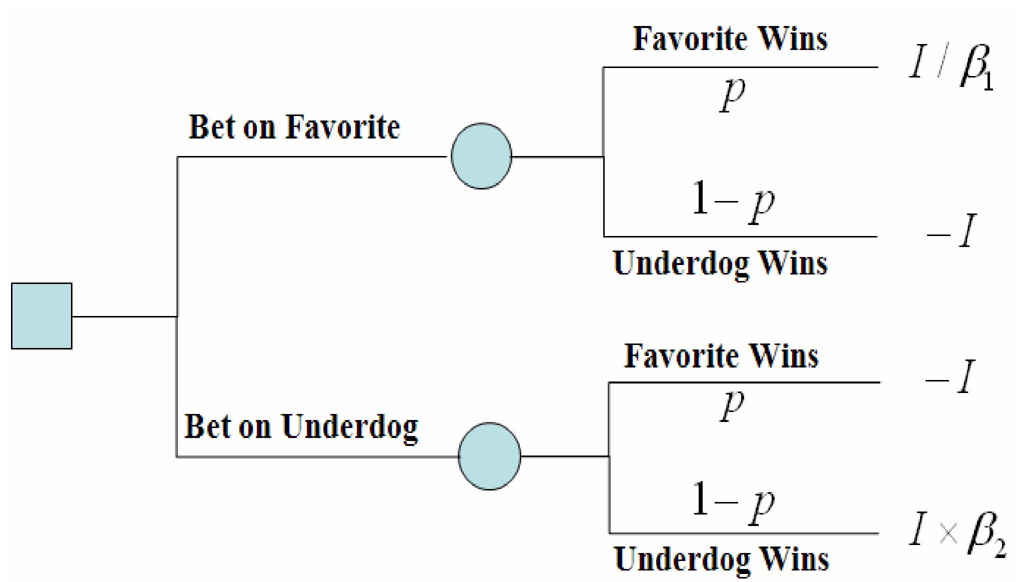
This model ignores hit batters and the fielder's choice.

How to go from money lines to Prob.

- Notations
  - β_1 : Favorite odd (-favorite money line/100)
  - β_2 : Underdog odd (underdog money line/100)
  - p : Probability that a favorite wins
  - I : Investment (The amount of money you wager)

  e.g. for (-140,+130), β_1 =1.4, β_2 =1.3

- To balance wagers, the bookmaker adjust money lines so that the expected val ues of both decisions are the same (his commission)



$$p(I/\beta_1) - I(1-p) = -Ip + (1-p)\beta_2$$

$$p = \left(\frac{\beta_2 + 1}{1/\beta_1 + \beta_2 + 2} \right)$$