A Bayesian Hierarchical Model To Estimate the Framing Ability of Major League Baseball Catchers

Sameer K. Deshpande and Abraham J. Wyner
Statistics Department, The Wharton School
University of Pennsylvania

26 September 2015

Introduction

- What exactly is framing?
  - A catcher's effect on the likelihood a taken pitch is called a strike.
- Should we even care about framing?
  - Lots of attention in popular presst
  - Several articles on Baseball Prospectus and Hardball Times
  - "Good framer" can save $\sim$ 15 - 25 runs per season.
  - Teams seem to care: Hank Conger, Russell Martin

Fundamental Question and Our Contribution

For any given *taken* pitch, what is the catcher's effect on likelihood of pitch being called strike over and above factors like:

- Pitch Location
- Pitch Context: Count, base runners, score differential, etc.
- Pitch Participants: batter, pitcher, umpire

Our contributions:

- Hierarchical Bayesian logistic regression model of called strike probability
- Value of a called strike as function of count
- Uncertainty estimates of framing impact (runs saved on average)

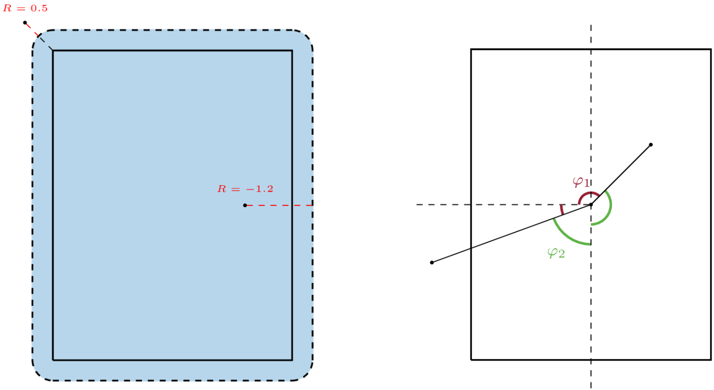
PITCHf/x Data

PITCHf/x data scraped from MLB Advanced Media:

- Horizontal and vertical coordinate of pitch as it crosses plate
- Approximate vertical boundaries for the strike zone
- Umpire, pitcher, catcher, batter identities
- Count

Focus only on the 320,308 taken pitches within 1 ft. of the strike zone.

Parameterizing Pitch Location



(a) Distance from strike zone, R (b) Angles from horizontal and vertical

Figure : R, φ_1, φ_2 for RHB

Bayesian Logistic Regression Model

For umpire u, log-odds of calling a strike is a linear function of:

$$R, \varphi_1, \varphi_2, \text{and indicators for batter, catcher, pitcher, and count}$$

Let $\Theta^{(u)}$ be vector of covariate partial effects.
Place common prior on $\Theta^{(u)}$'s ("borrow strength" between umpires)

$$\Theta^{(1)}, \ldots, \Theta^{(93)} \overset{\text{i.i.d}}{\sim} \bigotimes_{j} \text{Cauchy}(0, \lambda_j).$$

- $\lambda_j = 2.5$ as in [Gelman et. al (2008)].
- Gibbs sampling facilitated by Polya-Gamma data augmentation [Polson, Scott, and Windle (2013)]
- Identifiability: designate one batter, catcher, and pitcher as baseline

Differences between umpires' probability of calling strikes

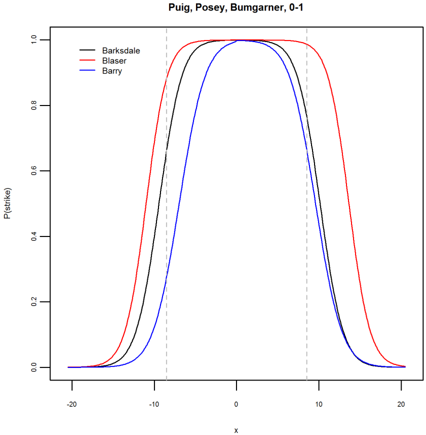


Figure : Inside $\longrightarrow$ Outside

Impact of Framing

For each catcher, look at all of the called pitches he received:

- $\hat{p}$: fitted probability of strike
- $\hat{p}^0$: fitted probability of strike with catcher replaced by baseline catcher
- $\hat{p} - \hat{p}^0$: catcher's "framing effect"
- Value of called strike, based on count, ρ
- Sum $\rho \times \left(\hat{p} - \hat{p}^0\right)$ over all called pitches received

Value of a called strike depends on the count!

Value of a called strike on an $0-1$ pitch

Between 2011 and 2014:

- 182,405 $0-1$ pitches taken: 140,667 balls, 41,738 called strikes
- Avg. # runs allowed in rest of inning after called ball: 0.322
- Avg. # runs allowed in rest of inning after called strike: 0.265

Conditional on an $0-1$ pitch being taken:

called strike saves $\rho = 0.057$ runs, on average

Expected Runs Saved Over All Pitches

Rank	Catcher	Runs Saved	95% Interval	Num. Pitches
1.	Hank Conger	13.95	[6.12, 22.38]	5515
2.	Miguel Montero	11.87	[3.51, 21.48]	9272
3.	Brian McCann	9.89	[1.90, 17.54]	7350
4.	Jose Molina	9.69	[1.90, 17.54]	5301
5.	Jonathan Lucroy	8.99	[0.95, 19.08]	9571
6.	Mike Zunino	8.97	[0.16, 18.34]	8822
7.	Rene Rivera	8.92	[1.34, 13.76]	5925
8.	Christian Vazquez	7.5	[1.44,13.76]	3770
9.	Russell Martin	7.37	[-0.24, 15.61]	7228
10.	Buster Posey	6.37	[-0.75, 14.69]	7441

Table : $\rho \times \left(\hat{p} - \hat{p}^0\right)$ summed over all of catcher's called pitches

Ongoing and Future Work:

- Other ways to incorporate pitch location:
  - Alternative parameterizations of strike zone
  - Non-parametric approach: generalized additive models
- Out-of-sample performance
- Improved Runs saved calculation
  - Non-uniform distribution of framing opportunities
  - Integrate $\rho \times \left(\hat{p} - \hat{p}^{0}\right)$ over batter, pitcher, umpire, count, and location.
  - Framing analog of SAFE [Jensen, Shirley, and Wyner (2008)]

References

Gelman, A., Jakulin, A., Pittau, M., and Su, Y. (2008) "A weakly informative default prior for logistic and other regression models." *Annals of Applied Statistics*, 2:4, 1360–1383.

Polson, N.G., Scott, J.G., and Windle, J. (2013) "Bayesian inference for logistic models using Polya-Gamma latent variables." *Annals of Statistics*, 108:504, 1339-1349

Jensen, S.T., Shirley, K.E., Wyner, A.J. (2008) "Bayesball: A Bayesian hierarchical model for evaluating fielding in Major League Baseball." *Annals of Applied Statistics*, 3:2, 491–520.

Acknowledgements

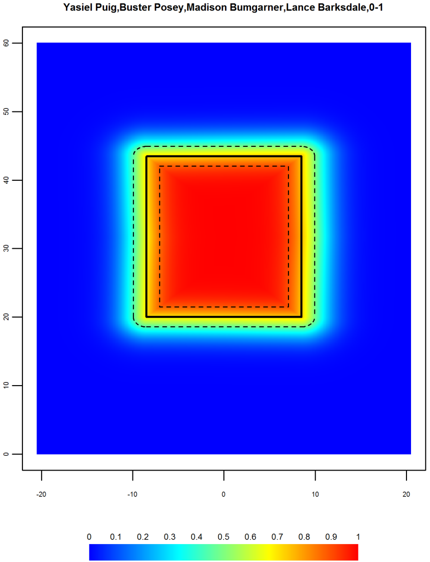
Special thanks to

- Hugh MacMullan and Wharton High-Performance Computing team.
- NESSIS Organizers
- All of y'all

Please send any ideas or questions!

Email: dsameer@wharton.upenn.edu
Paper and code will be available soon:
`www-stat.wharton.upenn.edu/~dsameer/pitchFraming/pitchFraming.html`

Estimated Strike Probabilities



Average # Runs Given Up and Value of Strike

Count	Ball	Strike	Value of strike, ρ
0-0	0.367 (0.002)	0.305 (0.002)	0.062 (0.002)
0-1	0.322 (0.002)	0.265 (0.004)	0.057 (0.004)
0-2	0.276 (0.003)	0.178 (0.007)	0.098 (0.008)
1-0	0.427 (0.003)	0.324 (0.003)	0.103 (0.005)
1-1	0.364 (0.003)	0.280 (0.004)	0.084 (0.005)
1-2	0.302 (0.003)	0.162 (0.006)	0.140 (0.006)
2-0	0.571 (0.007)	0.370 (0.006)	0.201 (0.009)
2-1	0.468 (0.005)	0.309 (0.006)	0.159 (0.008)
2-2	0.383 (0.004)	0.165 (0.006)	0.218 (0.007)
3-0	0.786 (0.013)	0.481 (0.008)	0.305 (0.015)
3-1	0.730 (0.010)	0.403 (0.009)	0.327 (0.014)
3-2	0.706 (0.008)	0.166 (0.008)	0.540 (0.011)

Table : Standard errors in parentheses