Jointly predicting exit velocity and launch angle for batter-pitcher matchups

Scott Powers

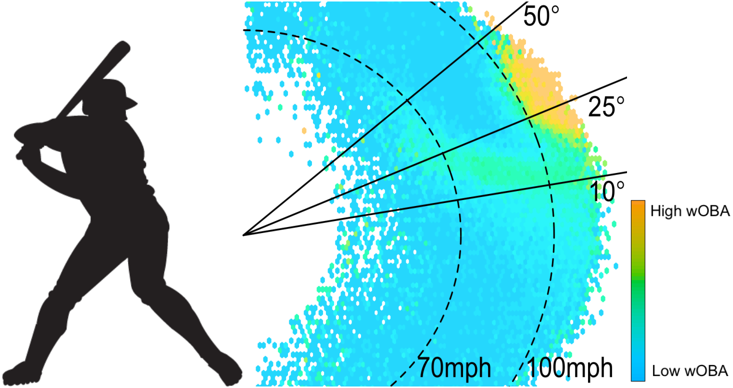
Stanford University

Saberseminar 2016

Statcast data

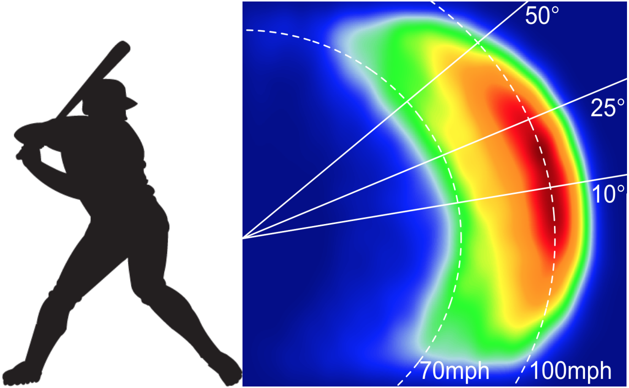
- Pitch-by-pitch data downloaded from BaseballSavant.com
    - Today's talk uses 2016 regular season data through July

wOBA by trajectory



- Averages insufficient; necessary to model **distribution**
- Necessary to model velo, angle **jointly**—not separately

Heatmap of trajectory



- Averages insufficient; necessary to model **distribution**
- Necessary to model velo, angle **jointly**—not separately

Outline of research

Goal

Given batter, pitcher, predict **joint distribution** of launch angle, exit velocity.

Value

- Batter/pitcher evaluation
- Fielder positioning

Strategy

1. Model distribution of launch angle
2. Model distribution of exit velo **conditional on launch angle**

Strategy explained

Launch angle $\sim$ **Contact**

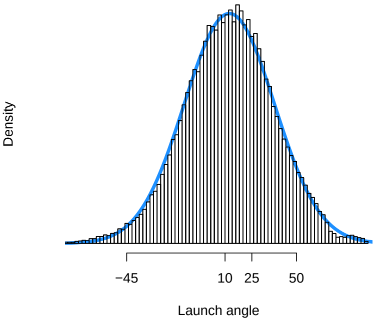
Exit velocity $\sim$ **Power** $+$ **Contact**

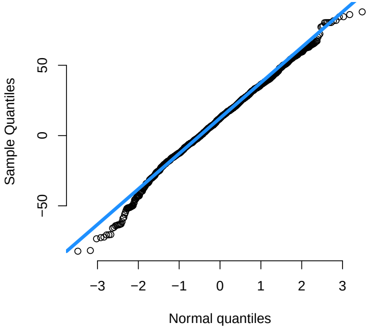
Launch angle $\sim$ **Contact**

Exit velocity $|$ **Launch angle** $\sim$ **Power**

Part I:
Modelling launch angle

Distribution of launch angle





- Normal model a good fit for launch angle distribution

Feasible generalized least squares

1. Use ordinary least squares to estimate residuals:

$$r_i = a_i - \hat{\mathbb{E}}_{\mathrm{OLS}}[a_i]$$

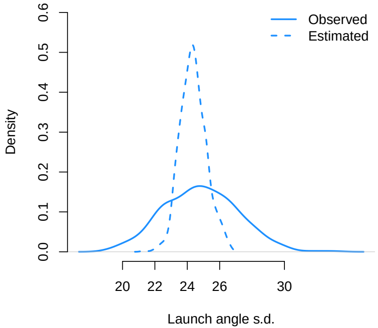
2. Fit linear model (ridge regression) to predict squared residuals:

$$r_i^2 = \alpha + \beta_{B_i} + \gamma_{P_i} + \delta_{S_i} + \zeta H_i + \theta O_i + \epsilon_i$$

$$\epsilon_i \overset{\text{i.i.d.}}{\sim} \mathrm{Normal}(0, \sigma_r^2)$$

3. Fit (generalized) linear model to predict launch angle
    - Ridge penalty!

Launch angle standard deviation results



Player	$\hat{sd}(a)$
Todd Frazier	26.7
Maikel Franco	26.6
Kevin Plawecki	26.6
Steven Wright	26.4
Kevin Kiermaier	26.4
...	
D.J. LeMahieu	22.3
Starlin Castro	22.3
Nick Castellanos	22.3
Jon Jay	22.1
Joey Votto	21.9

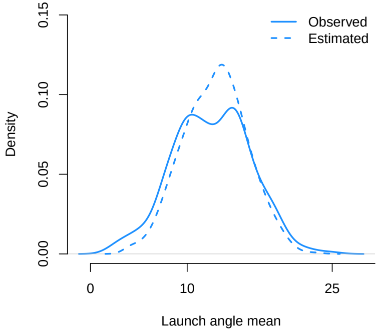
Launch angle model

$$a_i = \alpha + \beta_{B_i} + \gamma_{P_i} + \delta_{S_i} + \zeta H_i + \theta O_i + \epsilon_i$$

$$\epsilon_i \overset{\text{ind.}}{\sim} \text{Normal}(0, \sigma_a^2 \cdot \hat{r}_i^2)$$

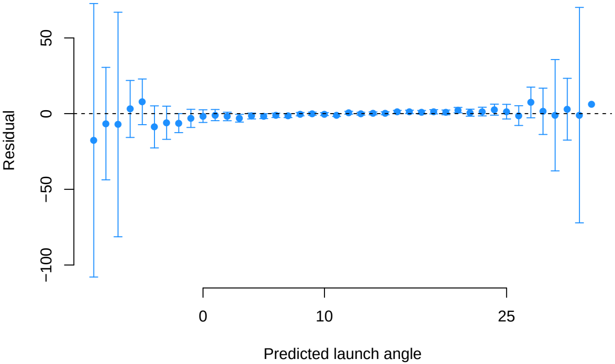
- Fit via ridge regression

Mean launch angle results



Player	$\hat{\mathbb{E}}[a]$
Ryan Buchter	23.4
Zach McAllister	22.1
Koji Uehara	21.0
Bryan Holaday	21.0
Nolan Arenado	20.7
...	
Marcus Stroman	4.3
Jeurys Familia	4.2
Jeremy Jeffress	4.2
Cameron Maybin	4.0
Christian Yelich	3.9

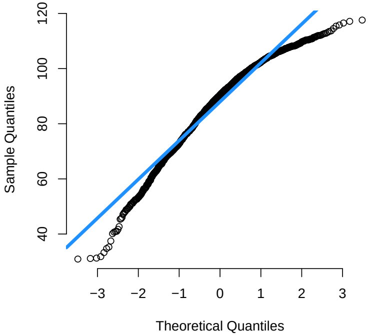
Launch angle model diagnostics



- Conclusion: no significant evidence against additive model

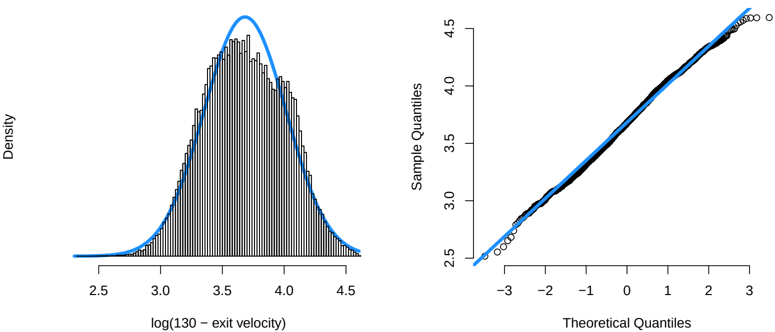
Part II:
Modelling exit velocity conditional on launch angle

Distribution of exit velocity



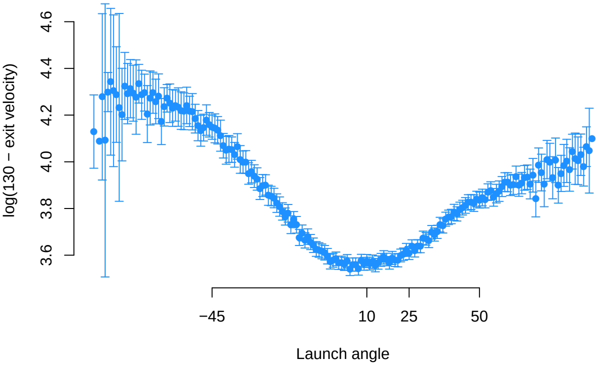
Let's try the transformation log(130 – exit velocity)

Distribution of transformed exit velocity



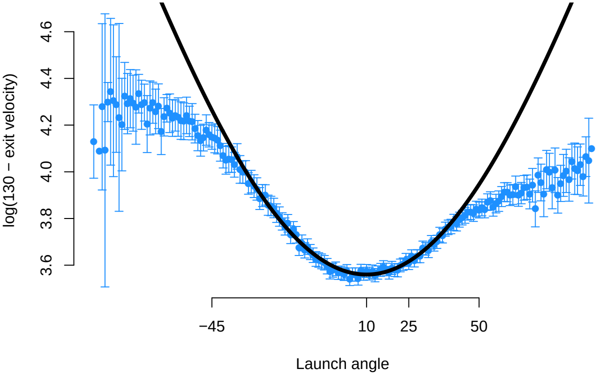
$$t = \log(130 - s)$$

Mean (transformed) exit velocity v. launch angle



-

Mean (transformed) exit velocity v. launch angle



- Between -35° and 45°, $\log(130 - s) \approx \alpha + \beta \cos(a - 10)$

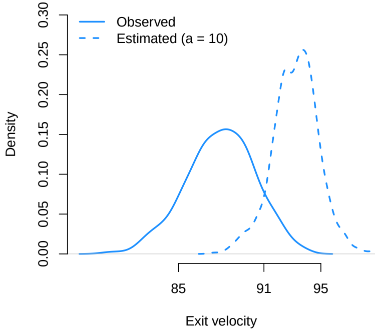
Exit velocity model

$$\log(130 - s_i) = \alpha + \omega \cos(a_i - 10) + \beta_{B_i} + \gamma_{P_i} + \delta_{S_i} + \zeta H_i + \theta O_i + \epsilon_i$$

$$\epsilon_i \overset{\text{ind.}}{\sim} \text{Normal}(0, \sigma_s^2)$$

- Fit via ridge regression

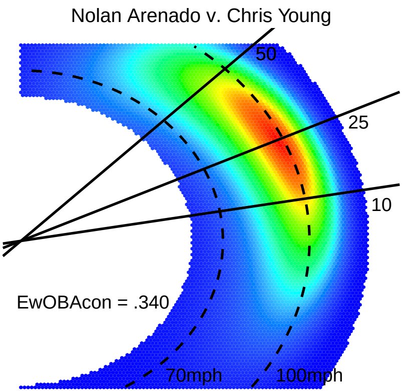
Mean exit velocity results (10° launch angle)

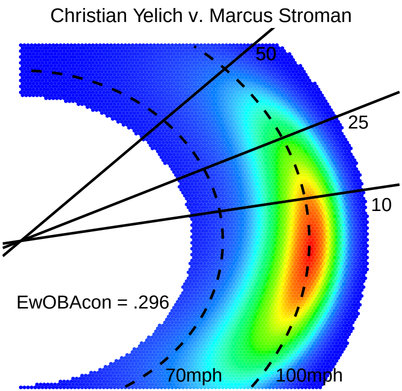


Player	$\mathbb{E}[v\|a = 10]$
Giancarlo Stanton	99.0
Mark Trumbo	98.9
Nelson Cruz	98.5
Matt Holliday	98.0
Ryan Zimmerman	97.6
...	
Jarrod Dyson	88.9
Jose Iglesias	88.7
Dee Gordon	88.5
Billy Hamilton	88.2
Billy Burns	87.5

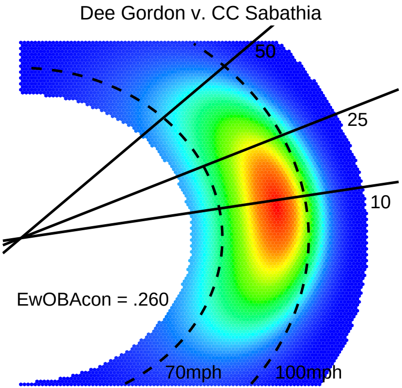
Part III:
Putting it all together

Extreme examples: Launch angle





Extreme examples: Exit velocity



Comments

- Optimal launch angle likely depends on batter
- Pitch types/locations could be included in model
- Next steps:
  - Model validation
  - Incorporating batted ball direction
  - Including non-batted balls (e.g. swing and miss)

Thank you for your attention!

Scott Powers
sspowers@stanford.edu
@saberpowers

`github.com/saberpowers/trajectory-distribution`

Scott Powers, "Toward a probability distribution over batted ball trajectories," *The Hardball Times*, August 19, 2016, hardballtimes.com/toward-a-probability-distribution-over-batted-ball-trajectories