Bayesball

Steven Tsai

KIN482D: *Computational Modeling of Human Sensorimotor Control and Learning*

Dr. Hyosub Kim

University of British Columbia

April 10 2024

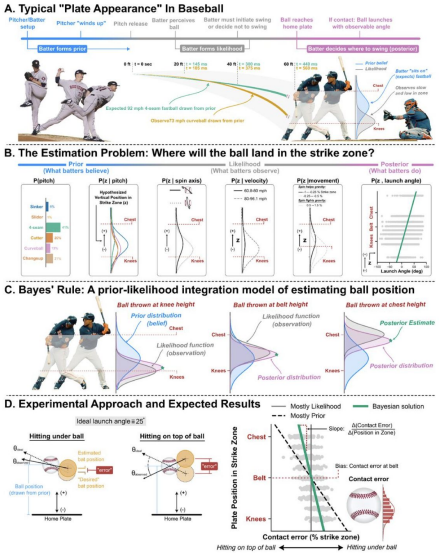
Source: Chat GPT (only for this picture I promise)

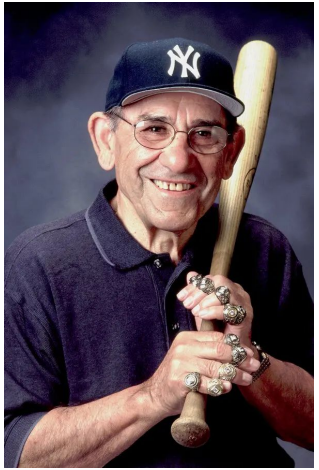
Brantley & Körding (2022)

- “You can’t think and hit at the same time”

  — Yogi Berra

- Would a batter be Bayesian?
  - Making decision with posterior that comes from integrating prior and likelihood



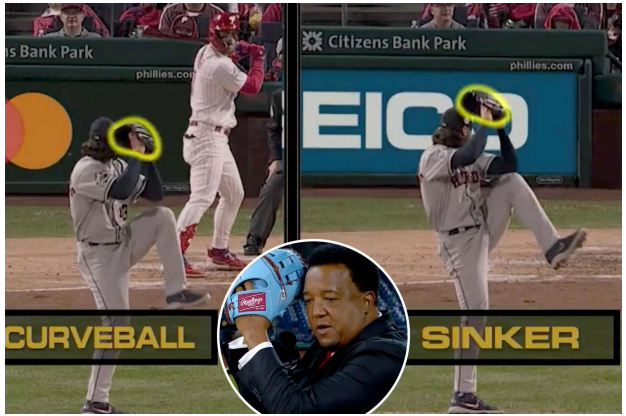


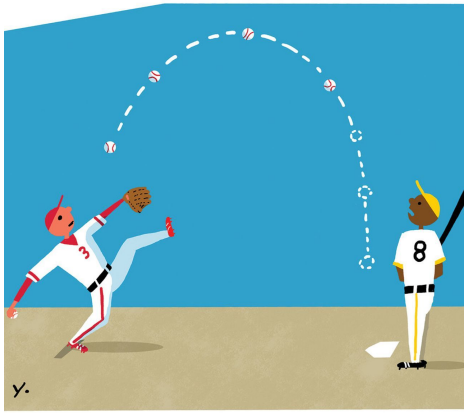
Link to the paper: https://www.biorxiv.org/content/10.1101/2022.10.12.511934v1.full

Three special cases

Brantley & Körding (2022) discussed three special cases where we may have different prior and/or likelihood:

1. Pitch tipping
    - Narrower prior -> narrower posterior
2. Knuckleball
    - Wide prior, wide likelihood
    - Likelihood is still more informative than prior
    - Posterior relied on likelihood
3. Eephus
    - Wide prior, narrow likelihood -> narrow posterior

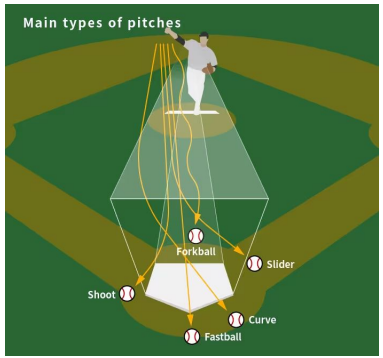




Knuckleball (source: Tread Athletics)

The issues

- Uncommon! (of these three cases)
- Only considering the vertical distance above the plate (z-axis)



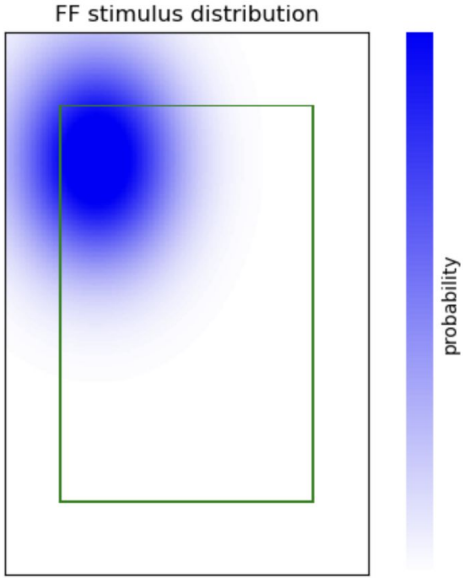
Motivation/Goals

1. Try to visualize the decision-making process on the strike zone plane (both x-axis and z-axis)
2. “What kind of pitch is coming?”
   - Uncertainty in the pitch type of the incoming pitch
   - Explore it using bayesian framework

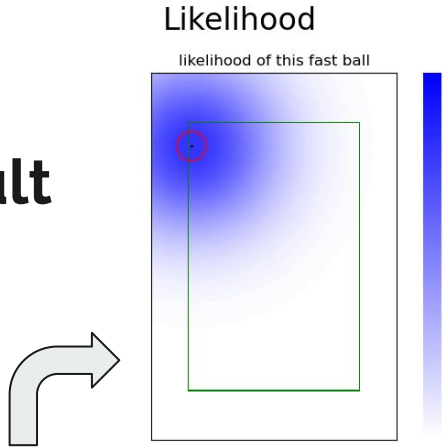
Case 1: only fastball

$$p(x) \sim \mathcal{N}(\mu_x, \sigma^2_{s,x})$$
$$p(z) \sim \mathcal{N}(\mu_z, \sigma^2_{s,z})$$
$$p(x, z) = p(x) * p(z)$$

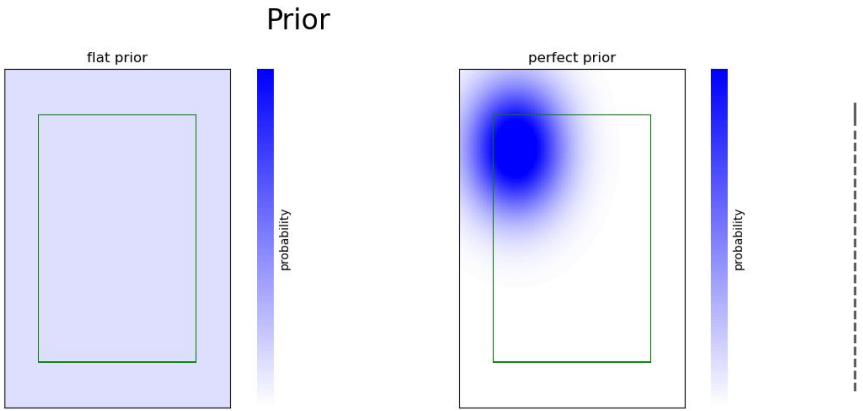
- In-lab (batting cage) experiment
  - With technologies to track movements of the ball & the swing
- A finely calibrated pitching machine
  - Gives us access to the stimulus distribution
- Feed the stimulus distribution to the batter
  - Artificially create his prior
  - In reality, we won't have access to the batter's prior (Brantley & Körding, 2022)

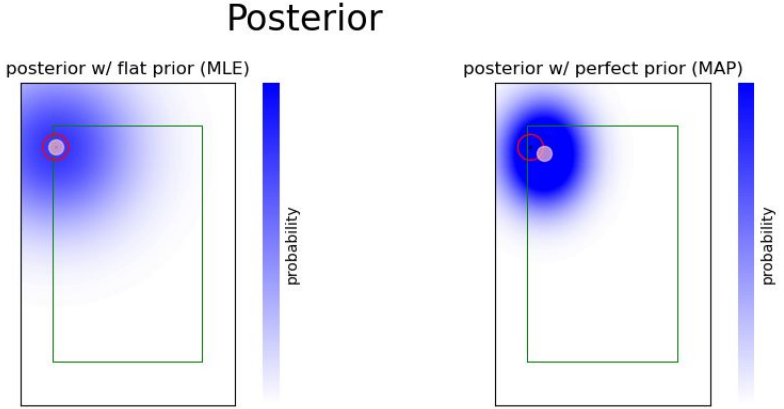


Case 1: result



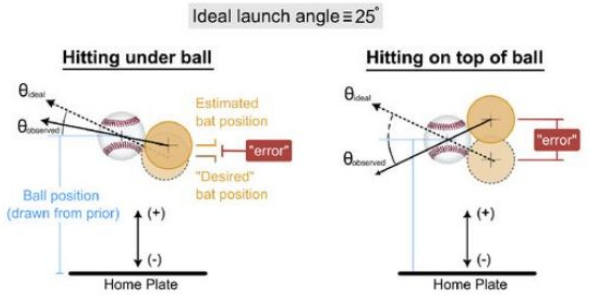
$$f(x,z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-((x-x_{obs})^2+(z-z_{obs})^2)}{2\sigma^2}}$$





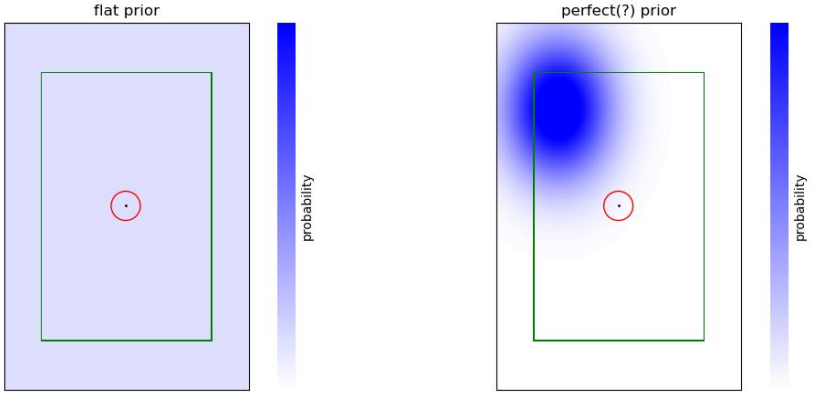
Case 1: find sigma (batter's measurement param)

- Assumptions:
  - Batter uses stimulus distribution as his prior
  - Batter aims to hit at 25deg (optimal launch angle for a homerun)
  - Contact error (Brantley & Körding, 2022) $\mathbf{e}_{\text{contact}} = -\mathbf{r}_{\text{baseball}} \times \left(\sin(\theta_{\text{optimal}}) - \sin(\theta_{\text{contact}})\right)$
  - Collect data -> Calculate the max log likelihood of sigma! (textbook C.5)
- Concerns:
  - Motor noise!
  - Z-axis only?
    - Contact happens in a 3D space



Case 1: surprise pitch

Prior, with a ball landing at the very centre

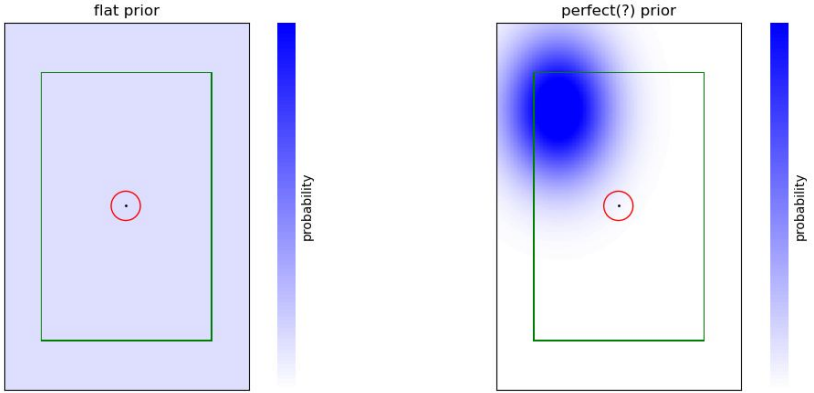


OR

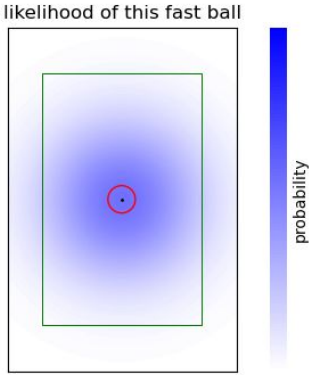
?

surprise pitch - result

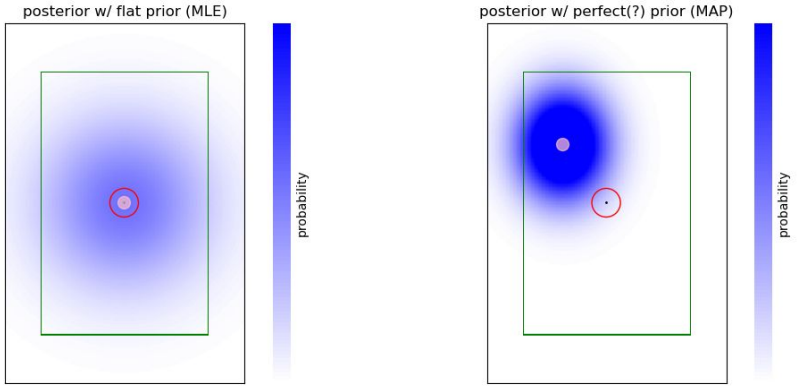
Prior, with a ball landing at the very centre



Likelihood

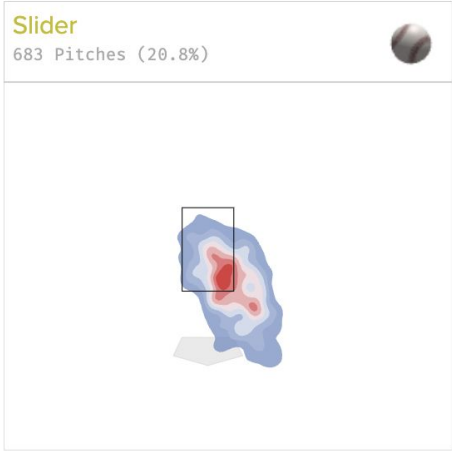


Posterior

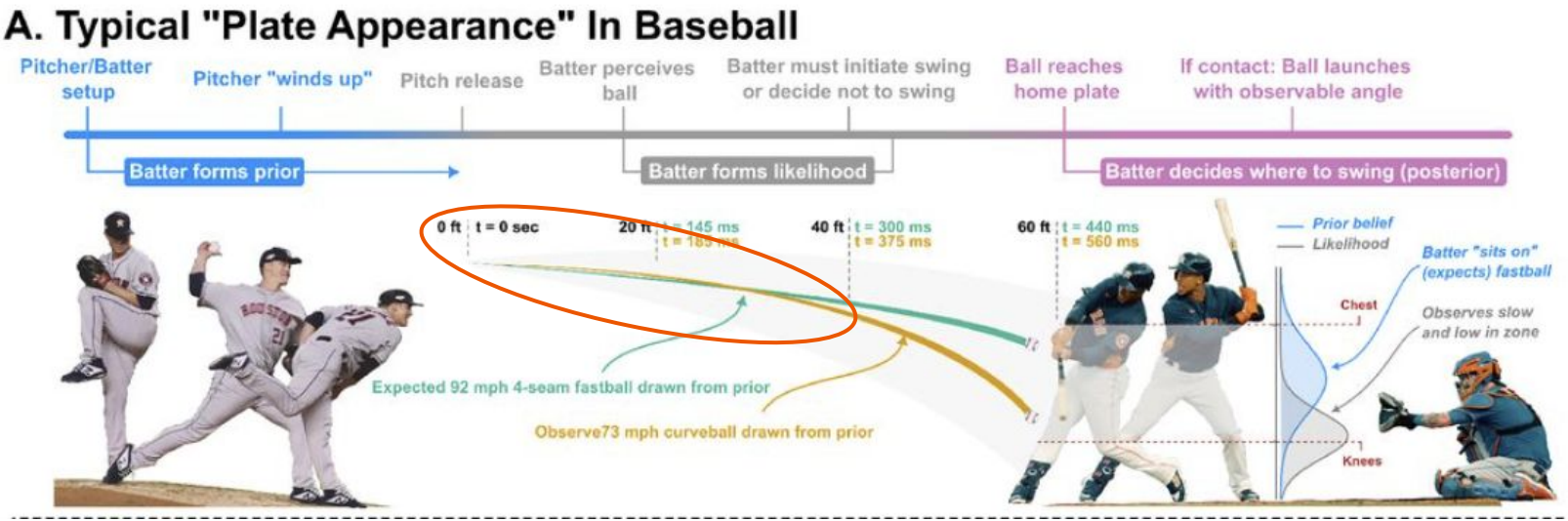


About Prior

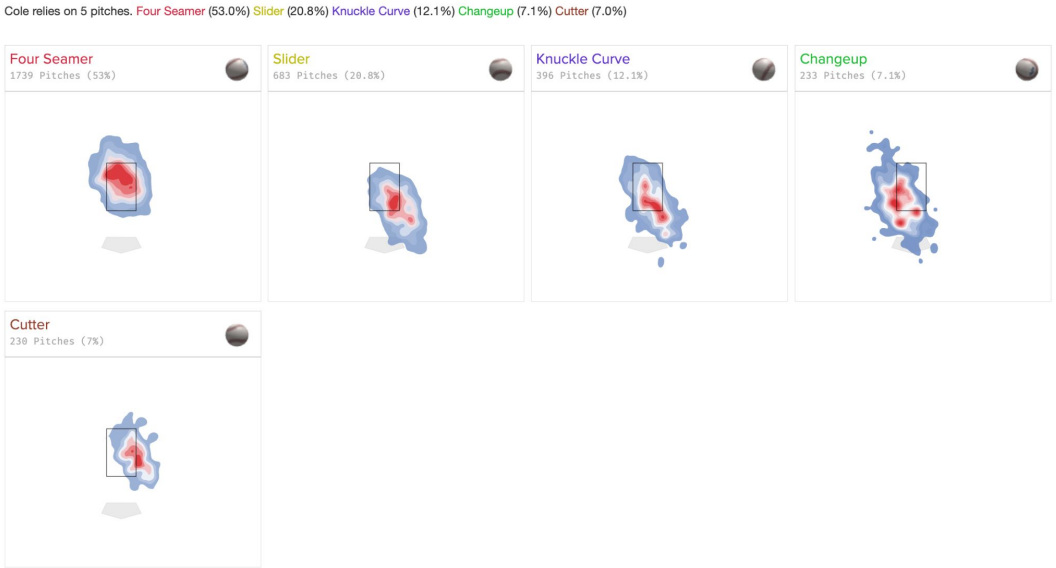
- Powerful when there is some level of measurement uncertainty
  - But can be misleading if incorrect
- Actually, if only fastballs, an elite batter may have a lot narrower likelihood
  - Pulls the posterior closer to the observation even if it is a surprise pitch
  - Would be interesting to run in the lab
- In reality, is not necessarily Gaussian
  - Can just be a probability matrix —>
  - Do the pointwise multiplication w/ likelihood
  - Happens in the batter's brain - not measurable



Timeline of a fastball (Brantley & Körding, 2022)



Case 2: uncertainty in pitch type



Case 2: Generative Model (tentative)

Step 1: infer pitch type

Prior: p(pitch)

- Fastball (50%)
- Slider (35%)
- Changeup (15%)

Observation *from the 1st half of the trajectory:

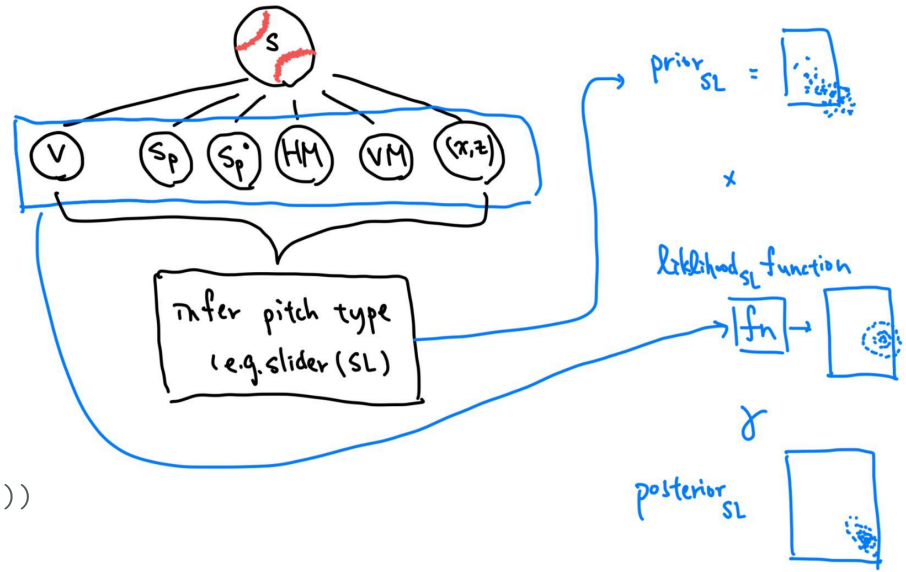
- V: speed
- Sp: spin
- Sp°: spin angle
- HM: horizontal movement
- VM: vertical movement
- x,z: horizontal & vertical positions

Cue Combination:

$$p(pitch|V, Sp, \ldots) \propto p(pitch)p(V|pitch)\ldots p(x, z|pitch)$$

Find the pitch type:

```
pitch=np,max(p(FF|obs), p(SL|obs), p(CH|obs))
```



(some of the) Concerns/Future Directions

- This model is confined to a 2D plane
  - Good: easy to visualize & understand
  - Bad:
    - Cannot consider swing timing
    - inference happens during the flight of the pitch; not on the strike zone plane
    - Contact not only happens on the strike zone plane; inaccurate when calculating response distribution/utility
  - A 3D model can be built if possible
- Data gathering
  - Using technologies e.g. HitTrax, high speed camera, radar
- When using launch angle to find the sigma (using max log likelihood)
  - Motor noise is not differentiated from it
  - How significant is the motor noise, given that swinging involves movement from multiple joints

Reference & GitHub

Brantley, J. A., & Körding, K. P. (2022). Bayesball: Bayesian Integration in Professional Baseball Batters. BiorXiv. doi: https://doi.org/10.1101/2022.10.12.511934. https://www.biorxiv.org/content/10.1101/2022.10.12.511934v1.full

GitHub to this project: https://github.com/Stevey8/Bayesball