Exploring America's Pastime with Bayes's Theorem

Baseball Meets Data Science

// FLATIRON SCHOOL

Problem

- Determine best hitters and pitchers in MLB history
- Hitters - BA; Pitchers - WHIP
- Cannot simply rank players according to these stats
  - Some players have very limited experience

Solution: Bayes's Theorem

- Assume to start that players are average
  - This will be our ***Bayesian prior***
- Use Bayes's Theorem:

$$P(h|e) = \frac{P(h)P(e|h)}{P(e)}$$

- P(h) → Prior → Historical Average Statistic
- P(e|h) → Likelihood → Player's Career Numbers
- P(h|e) → Posterior → Adjusted Statistic

Solution (cont.)

Players with lots of experience:

- Career numbers will dominate
- Greater reason to update our belief about whether player is average

Players with little experience:

- Prior will dominate
- Less reason to update our belief about whether player is average