Bayesian Ranking of Quiz Bowl Teams

Final Presentation for 553.432

Kuleen Sasse

Johns Hopkins University

Table of contents

Research Question

Research Question

Problem:

- Run large tournaments for Quiz Bowl Club on Campus
- Have to seed teams before tournament
- Using intuition currently which is slow and sees little of the data

Question:

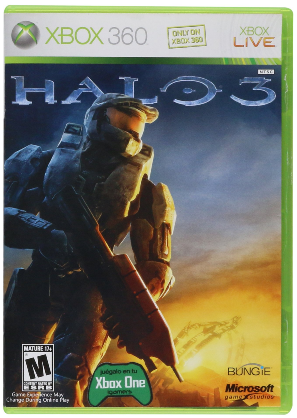
- Who is the current best high school/middle school team?
- What game statistics are the best to determine outcomes?

Literature Review

Literature Review

Ranking for Teams

- Extremely well developed field
- Based around a couple established methods
  - ELO and GLICKO
  - Bradley-Terry
  - Plackett-Luce
  - Davidson Model
  - Thurstone-Mosteller
  - Trueskill and Trueskill 2



Literature Review contd.

Ranking for Quiz Bowl Teams

- GrogerRanks
  - Homespun overfit model
  - Created without much statistical methods
- HSQBRank
  - Defunct
  - Used one indicator of strength

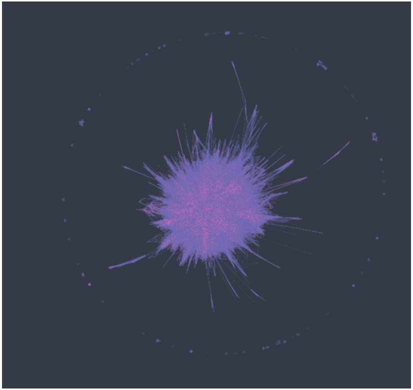
Data

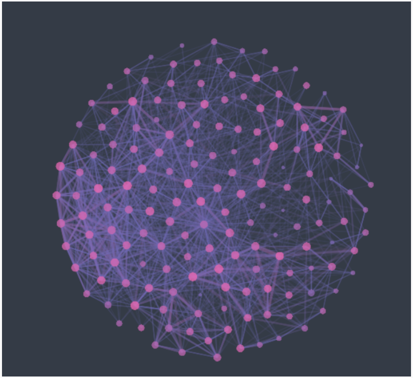
Data Collection and Cleaning

- Scraped 10,000 tournaments over 20 years from the largest host of quiz bowl tournaments
- Over 16,000 teams
- Collected 398,827 matches that included a winner and loser
- Not all matches had scores or statistics

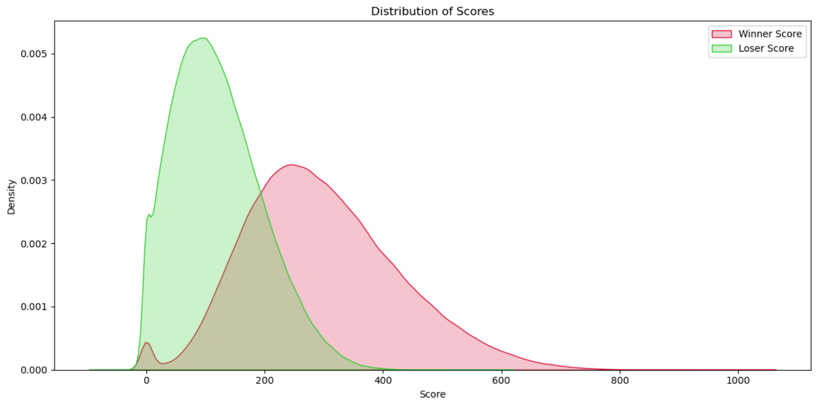
NAQT[SM]
National Academic Quiz Tournaments, LLC

Competitive Network Visualization

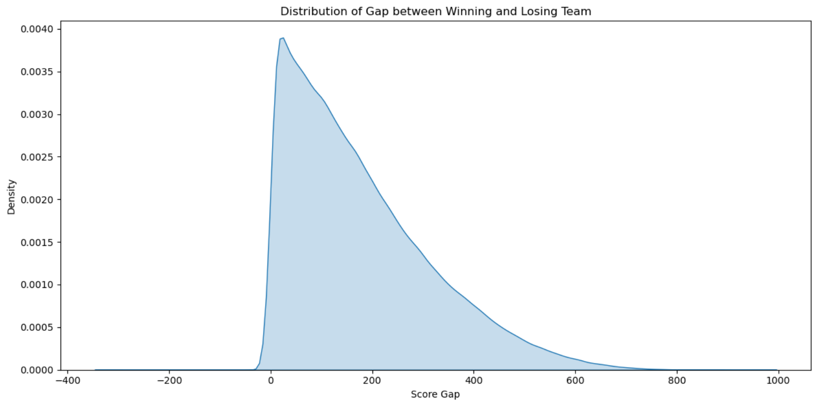




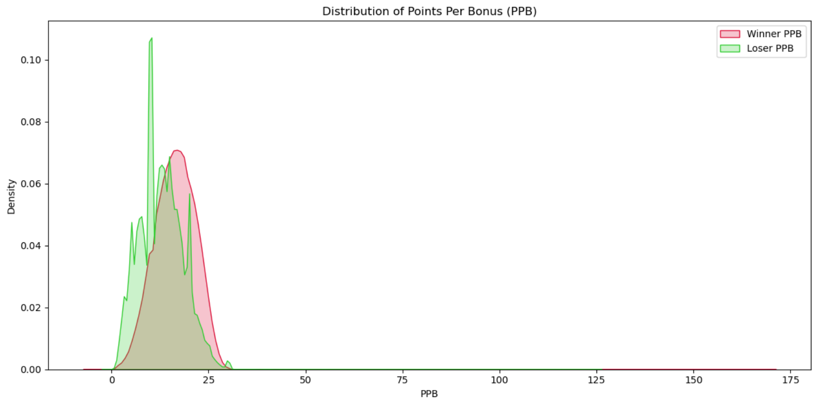
Score Statistics



Score Statistics contd.



Other Statistics



Methods

Likelihood:

$$\mathbb{P}[i \text{ beats } j] = \frac{\exp \lambda_i}{\exp \lambda_i + \exp \lambda_j}$$

$$y_{i,j} \sim \text{Bernoulli}(\mathbb{P}[i \text{ beats } j])$$

Prior:

$$\lambda_i \sim \mathcal{N}(0, \sigma_\lambda^2)$$

$$\sigma_\lambda^2 = 1$$

Generalized Bradley-Terry

Likelihood:

$$\mathbb{P}[i \text{ beats } j] = \frac{\exp \lambda_i}{\exp \lambda_i + \exp \lambda_j}$$

$$\lambda_i = \sum_{i=1}^{p} \beta_i X_i$$

$$y_{i,j} \sim \text{Bernoulli}(\mathbb{P}[i \text{ beats } j])$$

Prior:

$$\beta_i \sim \mathcal{N}(0, \sigma_\beta^2)$$

$$\sigma_\beta^2 = 1$$

Spike and Slab Generalized Bradley-Terry

Likelihood:

$$\mathbb{P}[i \text{ beats } j] = \frac{\exp \lambda_i}{\exp \lambda_i + \exp \lambda_j}$$

$$\lambda_i = \sum_{i=1}^{p} \beta_i X_i$$

$$\beta_i = z_i b_i$$

$$y_{i,j} \sim \text{Bernoulli}(\mathbb{P}[i \text{ beats } j])$$

Prior:

$$b_i \sim \mathcal{N}(0, \sigma_b^2)$$

$$z_i \sim \text{Bernoulli}(p_z)$$

$$\sigma_b^2 = 1$$

$$p_z = 0.5$$

Sampling Methods

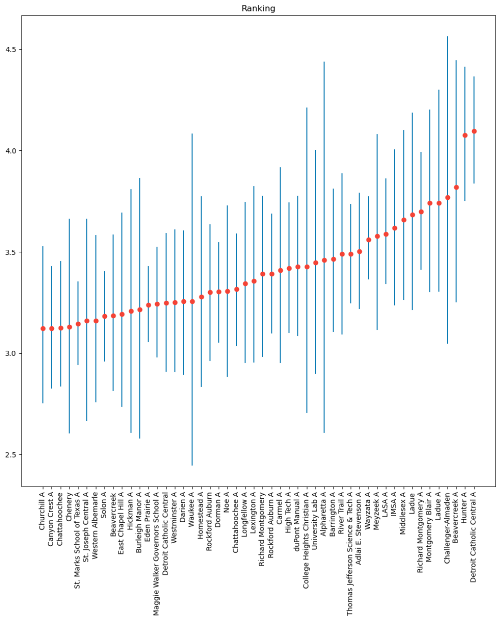
- Uses special sampling algorithm called No-U-Turn Sampler (NUTS)
- Chooses samples based on shape of posterior using using partial derivatives
- Spike and Slab uses NUTS + Gibbs

Preliminary Findings

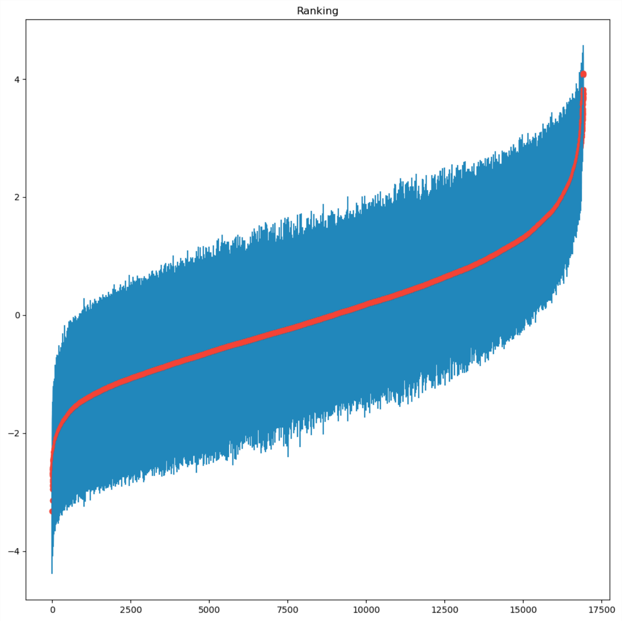
Bayesian Bradley Terry

- Extremely successful
- Outputted average scores with 95% CIs
- Agrees with widely held knowledge

Top 50 Rankings



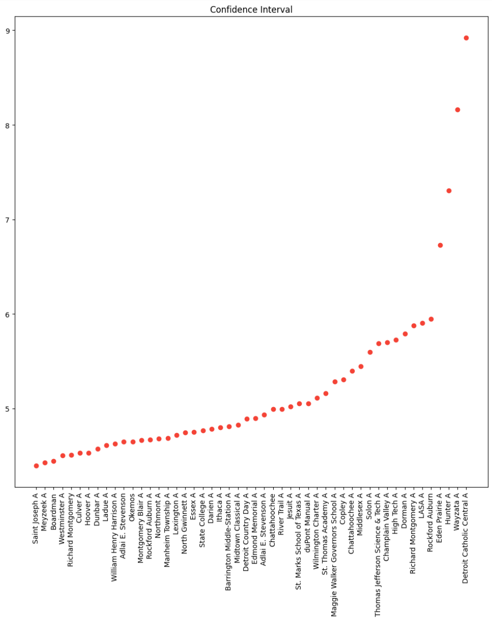
Total Rankings



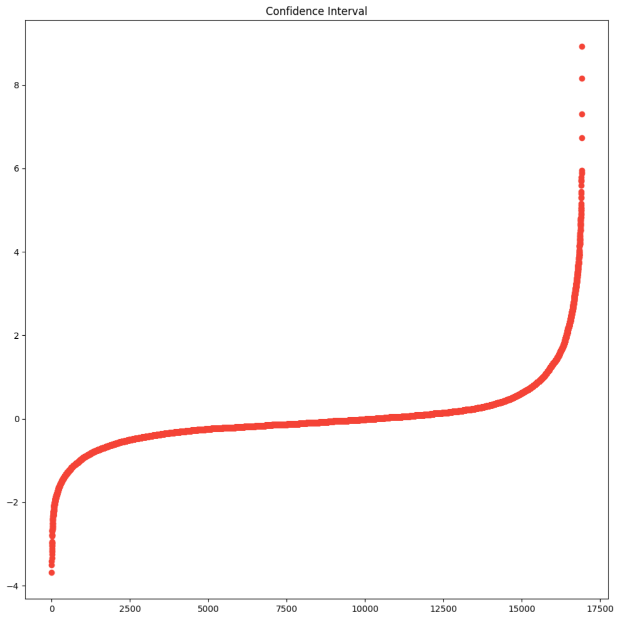
MLE Bradley Terry

- No 95% CIs
  - Could use Fisher Information
  - Extremely taxing to compute
- Possible overfit
- Does not agree as well with widely held knowledge

Top 50 Rankings



Total Rankings



Generalized Bradley Terry

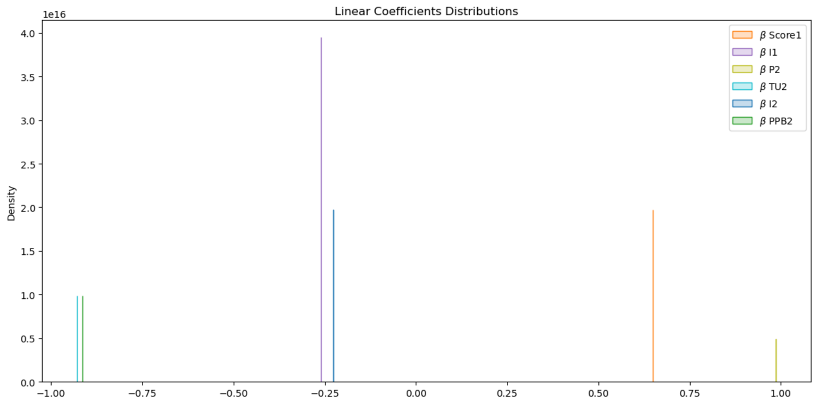
Bayesian methods failed

- Coefficients changed with every run
- Nonsense Coefficients
- Extremely overfit
- Same behavior with Spike and Slab

MLE methods failed

- Too much RAM/VRAM was needed to run it

Generalized Bradley Terry



Generalized Bradley Terry