1 Introduction

Over the past decade, I have been actively playing in the competitive scene of a trading card game known as Yu-Gi-Oh. This game involves building a deck from a selection of trading cards created by a company called Konami. Players then compete in tournaments against others with their strategically assembled decks. At the competitive level, decks typically revolve around specific archetypes— a collection of cards sharing the same name and strategy. For instance, well-known archetypes include "Blue-Eyes White Dragon" and "Dark Magician," meaning players might build decks predominantly featuring these archetypes.

A critical question for competitive players is what archetype should they play? In making this decision, players must consider several factors, including the power of each deck, its consistency, and how it fares against other popular decks. These considerations are part of what is known as meta-game analysis.

1.1 Meta-Game Analysis

Meta-game analysis involves examining the game at the tournament level to determine the best performing decks. While players can rely on fundamental analysis for meta-game insights, the complex nature of selecting decks and the potential for human bias make this challenging. Hence, why players will also perform technical analysis, manifested in tournament results.

Tournament results detail the top N players, the tournament's name, archetype names, player placements, and tournament sizes. The value of N varies based on tournament size but typically includes the top 4, 8, 16, 32, and 64. By analyzing tournament outcomes, a player can gauge the standing of archetypes in the meta by noting how frequently a deck "tops," indicating its presence in the tournament results. However, only major tournaments are reported to platforms like https://ygoprodeck.com/ ("YGOPRODECK"), and there are only a handful of tournaments each season. This method currently underpins the technical analysis of the meta, influencing both players' deck choices and even Konami's decisions on which decks to nerf (which means to make weaker.)

1.2 Bayesian Use Case

The existing approach of counting the number of times a deck tops offers basic statistics about deck performances within the meta but fails to consider the placement, the tournament size in the "tops," the power and consistency with which each deck tops, or the reliability of these predictions. It also struggles to adapt to the dynamic nature of the meta, particularly when cards are banned/unbanned or new cards are introduced in new sets. For these reasons, I propose adopting a Bayesian hierarchical model, where each deck's performance is modeled using a beta distribution.

2 Analysis

2.1 The Data

I have hand made data from 160 tournament results from several Yu-Gi-Oh Championship Series tournaments (YCS) held from September to October 2023 that were reported on ygoprodeck.com. This dataset includes the tournament name, the archetype, the placement within the tournament, and its size. During this period, a ban list was issued that nerfed the Kashtira archetype. These developments are mirrored in the tournament outcomes, with a noticeable decline in Kashtira's popularity over time.

2.2 Preprocessing

In the preprocessing step, I normalized each deck's placement to fit within the (0,1) range required for the Beta distribution, using the formula: $\text{data['normalized\_placement']} = 1 - \frac{\text{data['Place']}}{\text{data['Size']}}$. This normalization works well because it also assigns higher values to wins in larger tournaments. For instance, securing 1st place in a 100-person tournament yields a normalized placement of 0.99, whereas winning a 2-person tournament results in a 0.5.

2.3 Bayesian Hierarchical Model

This model assesses the strengths of different Yu-Gi-Oh archetypes by examining individual deck performances across various tournaments.

- Let d denote individual decks, where $d = 1, \ldots, N_d$.
- Let a denote the archetypes, with $a = 1, \ldots, N_{a,d}$.
- p_{ad} represents the normalized placement of the archetype a which contains d decks, serving as a performance metric in a tournament.

The normalized placement p_{ad} for each deck observation d within archetype a follows a Beta distribution:

$$p_{ad} \sim \text{Beta}(\alpha, \beta) \tag{1}$$

The α and β parameters for each archetype, shaped by base priors, are modeled using Half-Normal distributions to ensure positivity:

$$\alpha_{\text{prior}} \sim \text{HalfNormal}(\sigma = 1) \tag{2}$$
$$\beta_{\text{prior}} \sim \text{HalfNormal}(\sigma = 1) \tag{3}$$
$$\alpha \sim \text{HalfNormal}(\sigma = \alpha_{\text{prior}}, \text{shape} = N_{\text{archetypes}}) \tag{4}$$
$$\beta \sim \text{HalfNormal}(\sigma = \beta_{\text{prior}}, \text{shape} = N_{\text{archetypes}}) \tag{5}$$

Key model components include:

- α_{prior} and β_{prior}, the priors for archetype strength parameters.
- α and β, the archetype-specific strength parameters determining the shape of the Beta distribution for deck performances within each archetype. This controls for independence of decks playing against each other at the same tournament and also brings together the results of each deck into their respective archetype.

Bayesian updating on every deck observation allows us to find the posterior distributions of α and β for each archetype, enabling the calculation of the posterior mean and variance for each archetype. The posterior mean (μ) and variance (σ^2) of a Beta distribution, representing the expected performance and variability of an archetype, are given by:

$$\mu = \frac{\alpha}{\alpha + \beta} \tag{6}$$

and

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{7}$$

where α and β denote the shape parameters associated with successful and weaker performances, respectively. The posterior mean (μ) estimates the average performance of an archetype, with values nearer to 1 indicating stronger performance. The posterior variance (σ^2) reflects the uncertainty in performance estimates, with higher variance denoting greater unpredictability.

3 Results

Below are two tables. The one on the left represents the output of the Bayesian hierarchical models, and the one on the right represents the tier list from the frequentist approach.

The comparison between these two results is subjective because determining which decks are the strongest from the metagame is challenging and has no clear metric. Therefore, I will analyze the two lists based on my knowledge of the metagame and observations.

Deck	Power	Consistency
Rescue-Ace	0.970196	0.001647
Unchained	0.968252	0.001474
Tearlaments	0.967532	0.001808
Purrely	0.960406	0.002019
Floowandereeze	0.959069	0.002695
Vanquish Soul	0.957854	0.004068
Kashtira	0.957842	0.002440
Rikka	0.956833	0.004392
Mannadium	0.956192	0.002995
Dragon Link	0.954695	0.002436
Despia	0.953344	0.003160
Labrynth	0.952309	0.003531
Spright	0.948642	0.004221
Exosister	0.948056	0.005961
Chimera	0.944069	0.004562
Horus	0.943666	0.006330
Infernoble Knight	0.940234	0.006617
Mikanko	0.938906	0.006805
Tri-Brigade	0.937728	0.006795
Runick	0.936032	0.007072
Virtual World	0.931374	0.007559
Mathmech	0.931089	0.006880
Fur Hire	0.930656	0.007661
Dinosaurs	0.923060	0.008679

Table 1: Power and Consistency of Deck Archetypes Using Bayesian Modeling

Deck	Count
Unchained	27
Purrely	18
Rescue-Ace	16
Dragon Link	14
Tearlaments	13
Kashtira	12
Floowandereeze	11
Despia	9
Mannadium	9
Labrynth	6
Spright	4
Chimera	4
Vanquish Soul	3
Rikka	3
Mathmech	2
Tri-Brigade	1
Runick	1
Horus	1
Mikanko	1
Fur Hire	1
Virtual World	1
Exosister	1
Dinosaurs	1
Infernoble Knight	1

Table 2: Count of Deck Archetypes

Regarding the aspects my Bayesian model handled well, it accurately identified the Rescue-Ace archetype as the strongest during the time frame of this tier list, closely followed by Unchained. It also highly ranked archetypes with less representation, such as Vanquish Soul and Rikka, which I believe deserve their high ranking. Another strength of the model is its report on the high consistency numbers for each of these decks, indicating that while their average performance is strong, they are inconsistent in the tournament setting. These insights are invaluable when discussing the metagame.

However, my model struggled to identify the strength of Dragon Link and Purrely accurately. This oversight became apparent upon reviewing the data, as these are both affordable decks in Yu-Gi-Oh. Typically, players who invest in top-tier decks like Unchained or Rescue-Ace are fully committed to the game and, on average, are more skilled. Conversely, more affordable archetypes attract a broader range of player skill levels. This leads to a higher frequency of lower-skilled players finishing near the bottom of tournament results. This negatively impacts their beta distribution and unfairly penalizes these decks for lower placements, which is not an ideal outcome compared to decks with less representation.

4 Conclusion

Overall, my method contributes important data to determine archetypal strengths within the meta. Calculating the power and consistency of decks is crucial for uncovering strong strategies, particularly when analyzing archetypes that are under represented. I enjoyed working on this project and plan to further develop it beyond this assignment. I welcome feedback, particularly in areas such as:

- Feature Engineering: How I can improve on my data going in to create a more accurate model?
- Model Choice: What other Bayesian approaches can I do that could result in a more accurate analysis?
- How to incorporate more of the data I can gather, such as controlling by tournament or using the contents of each deck.

5 References

1. “YGOPRODECK ” YGOPRODECK, ygoprodeck.com/.