Tournament Thresholds

# Summary

The approach is to use the private matches to estimate the cost function for improving Banner. With the cost function, simulated experimental data can be generated, and estimates derived from the simulated data. This approach requires large assumptions on behavior and costs, but that is unavoidable. I use the Signals tournaments as a guide when making a decisions on various needed quantities.

# Assumptions

## Structural Assumptions

* Costs of obtaining a score of are a function of ability () (as measured by ratings):
* The marginal cost function is given by:
* In order to estimate costs and prizes, we need to restrict average marginal cost to be 1:
* Assume the private competitors expended the same amount of effort () for their submissions. Then OLS can be used to estimate the cost parameters ().

## Environmental Assumptions

* Given that the problem is improving a machine learning algorithm, pure marathon ratings should be both the most accurate measure of ability and be (fairly) universal among competitors.
* The distribution of marathon ratings will be similar to that of the Signals challenge. I think the distribution from a recent marathon match will probably be closer to reality.



Figure : Assumed distribution of marathon ratings; from Signals experiment.

# Cost Estimation

* Rescale marathon ratings to the unit interval using the assumed distribution of marathon ratings.

|  |  |  |
| --- | --- | --- |
| Rating | Ability | Score |
| 2296 | 0.73 | 817883.23 |
| 2256 | 0.72 | 833310.61 |
| 1967 | 0.62 | 817602.71 |
| 1878 | 0.59 | 806013.33 |

* Then the cost parameters as estimated by OLS are:

|  |  |
| --- | --- |
| Parameter | Estimate |
|  | 1600171.54 |
|  | 0.07 |
|  | 0.96 |



Figure : Marginal Costs

# Bid Function Estimate

* Assume the prizes are . A 90/10 split of 3 times the effort cost in the private matches. This is a big guess.



Figure : Estimated Bid Function

# Threshold Estimates



Figure : Ability Thresholds based on Score Threshold



Figure : Quantile (of ability) Threshold based on Score Threshold