Summary about Prediction Market

Zhihao Ruan, Jiayu Yi, Naihao Deng
{ruanzh,yijiayu,dnaihao}@umich.edu

August 4, 2019

Contents

1 Exponential Family

2 Scoring Rules

Scoring rules is the simplest form of prediction mechanism. For every agent with some information, a scoring rule evaluates how close the information is from the actual outcome, and pays the agent in return for his/her information.

2.1 Motivations

Taking an event $\mathbf{X}$ with outcome space $\mathcal{X}$, we want to know something about each agent's belief $p(\mathbf{x})$ on the actual outcome $\mathbf{x} \in \mathcal{X}$, compare it with the actual outcome, and see how accurate the agent's prediction is. However, it is impossible to ask agent for his/her entire $p(\mathbf{x})$ probability distribution. What should we do?

We've already known that sufficient statistic is a very suitable parameter to characterize a probability distribution. Therefore, we can use it to represent agent's belief. We just ask each agent for the sufficient statistic $\mathbf{u}(\mathbf{x})$ as a representation of his/her belief $p(\mathbf{x})$. Then we are able to measure how accurate the agent can predict with a scoring rule based on the actual outcome.

Assume $\boldsymbol{\mu} = \mathbb{E}_p[\mathbf{u}(\mathbf{x})]$ is the *expected* sufficient statistic over some $p(\mathbf{x})$ that takes $\mathbf{u}(\mathbf{x})$ as its sufficient statistic and is taken as an estimate of the agent's belief. Assume $\hat{\boldsymbol{\mu}}$ is the agent's report as an estimate of $\boldsymbol{\mu}$. Then, with the actual outcome denoted $\mathbf{x}$, the scoring rule takes the form:

$$S(\hat{\boldsymbol{\mu}}, \mathbf{x}).$$

We can see that the scoring rule is a measure of how close the agent's belief is from the actual outcome.

2.2 Incentive Compatibility

Incentive compatibility is a property of prediction mechanism with which the best strategy for an agent to earn the most profit is to *honestly* report all the information as soon as he/she has it. As a prediction mechanism, a proper scoring rule should leverage **incentive compatibility** in order to get real information from agents.

Assume that we already set the sufficient statistic to be $\mathbf{u}(\mathbf{x})$. Then for each $p \in \mathcal{P}$ that takes this $\mathbf{u}(\mathbf{x})$ as its sufficient statistic, we can calculate its *expected* sufficient statistic $\boldsymbol{\mu} = \mathbb{E}_p[\mathbf{u}(\mathbf{x})]$. A scoring rule is thus called **proper** if it satisfies the following, for all such p, for all $\hat{\boldsymbol{\mu}} \neq \boldsymbol{\mu}$:

$$\mathbb{E}_p[S(\boldsymbol{\mu}, \mathbf{x})] \geqslant \mathbb{E}_p[S(\hat{\boldsymbol{\mu}}, \mathbf{x})].$$

Any scoring rule with this property actually encourages agents to report a probability distribution as close to the actual probability distribution of $\mathbf{X}$ as possible, which is in accordance with the essence of incentive compatibility.

2.3 Logarithmic Scoring Rule

Suppose that we have set a form of $\mathbf{u}(\mathbf{x})$. For some p that takes $\mathbf{u}(\mathbf{x})$ as its sufficient statistic, a classic logarithmic scoring rule takes the form:

$$S(\boldsymbol{\mu}, \mathbf{x}) = \ln p(\mathbf{x}; \boldsymbol{\mu}),$$

where $\boldsymbol{\mu} = \mathbb{E}_p[\mathbf{u}(\mathbf{x})]$ is the expected sufficient statistic over $p(\mathbf{x}; \boldsymbol{\mu})$. **For the following sections, we will be only talking about logarithmic scoring rules.**

2.4 Maximum Entropy Optimization

Now that we have the general expression of logarithmic scoring rules, how can we determine which $p(\mathbf{x}; \boldsymbol{\mu})$ to choose as an estimate of agents' beliefs, given a specific $\mathbf{u}(\mathbf{x})$?

For some particular form of $\mathbf{u}(\mathbf{x})$, denote the set $\mathcal{P}$ which contains all p that takes $\mathbf{u}(\mathbf{x})$ as its sufficient statistic. Then, we formulate the following optimization problem:

$$\begin{aligned} \min_{p \in \mathcal{P}} \quad & F(p) = \int_{\mathbf{x} \in \mathcal{X}} p(\mathbf{x}) \ln p(\mathbf{x}) \, \mathrm{d}h(\mathbf{x}) \\ \text{s.t.} \quad & \mathbb{E}_p[\mathbf{u}(\mathbf{x})] = \boldsymbol{\mu} \end{aligned}$$

where $\boldsymbol{\mu}$ is the expected sufficient statistic over p, and $h(\mathbf{x})$ is the base measure.

It is clear to see that $F(p)$ is actually the negative of entropy of $p(\mathbf{x})$. Therefore, minimizing $F(p)$ is thus maximizing the entropy of p. This actually mean that of all $p \in \mathcal{P}$, we tend to choose a p that shows the most uncertainty of $\mathbf{X}$ as an estimate of agents' beliefs. This estimate is not necessarily exactly the same as agents' beliefs, as the general goal is just to find something that we can use to construct a metric in order to measure how close the agents' beliefs are from the actual outcome.

It has been proved that $S(\boldsymbol{\mu}, \mathbf{x}) = \ln p(\mathbf{x}; \boldsymbol{\mu})$ is **proper** if and only if p is the exponential family. It has also been proved that the solutions to the optimization problem are exponential family distributions which takes the form

$$\begin{aligned} p(\mathbf{x}; \boldsymbol{\eta}) &= h(\mathbf{x}) g(\boldsymbol{\eta}) \exp\{\boldsymbol{\eta} \cdot \mathbf{u}(\mathbf{x})\} \\ &= h(\mathbf{x}) \exp\{\boldsymbol{\eta} \cdot \mathbf{u}(\mathbf{x}) - T(\boldsymbol{\eta})\} \end{aligned}$$

where $\boldsymbol{\eta}$ is the natural parameter of p, $T(\boldsymbol{\eta}) = -\ln g(\boldsymbol{\eta})$ is the log-partition function of p.

Either *expected* sufficient statistic $\boldsymbol{\mu}$ or natural parameter $\boldsymbol{\eta}$ can parametrize a probability distribution p, and actually they are closely related. If we denote the convex conjugate of $T(\boldsymbol{\eta})$ as $G(\boldsymbol{\mu})$, then we have $\boldsymbol{\mu} = \nabla T(\boldsymbol{\eta}), \boldsymbol{\eta} = \nabla G(\boldsymbol{\mu})$.

3 Cost Function based Prediction Market with Bayesian Traders