

Prediction Markets for Exponential Families

Sindhu Kutty

Department of Computer Science and Engineering
University of Michigan, Ann Arbor

Rahul Sami

School of Information
University of Michigan, Ann Arbor

August 6, 2013

Abstract

This paper identifies and develops new connections between two approaches to aggregating information from multiple sources: statistical learning techniques and prediction markets that rely on profit-seeking traders to update prices. We present a new rich class of prediction markets for infinite outcome spaces that are based on parametric distribution families. For these markets, we prove that a sequence of traders maximizing their expected profit will result in price sequences analogous to sequential maximum likelihood estimation from statistics. Further, we show that even when some traders are adversarial attackers rather than strategic profit-maximizers, these prediction markets continue to have attractive learning properties: The damage that the attackers can cause is bounded by their initial budgets, and unlike attackers causing damage, informative traders' budgets will grow in expectation as they trade in more markets.

1 INTRODUCTION

Prediction markets are aggregation mechanisms that allow market prices to be interpreted as predictive probabilities on an event. Each trader in the market is assumed to have some private information that he uses to make a prediction on the outcome of the event. The traders are assumed to be strategic in the sense that every trader wants to maximize his profit by participating in the market. Further, since the trades are done sequentially, the trader is allowed to observe all past trades in the market and update his private information based on this information.

Traders are allowed to report their beliefs on the outcome of the event by allowing them to buy and sell securities whose value depends on the outcome of this future event. This will effect the state of the market, thus updating the predictive probabilities for the event. For example, the Iowa Electronic Market predicts the outcome of a presidential election by allowing one to trade on securities that pay off only if the Democratic candidate wins, or only if the Republican candidate wins. Traders can see

the past history of trades, so the price at which a current trader is willing to buy and sell these securities can be interpreted as an aggregate “market probability” of a particular candidate winning the election.

Currently, most deployed prediction markets have binary or finite discrete outcome spaces. This means that the outcome of every event can be characterized by a finite set of mutually exclusive exhaustive outcomes. The securities in these market are usually tied to each outcome so that buying (respectively, selling) a security related to an outcome, indicates a belief in an increased (respectively, decreased) probability of occurrence of the outcome.

One form of prediction markets, which is rapidly gaining in popularity, is the market scoring rule [Hanson, 2003]. A market scoring rule considers all trades as a single chronological sequence. Traders earn rewards proportional to the incremental reduction in prediction loss caused by their trades in comparison to the previous trade. In other words, their rewards depend on the change in market probabilities caused by their trade, as well as on the eventual outcome. Thus, each trader has an incentive to minimize the prediction loss. In this format, the *market maker* who runs the market can suffer an overall loss, but Hanson [Hanson, 2003] showed that, for market scoring rules on finite outcome spaces, the loss of the market maker can be bounded.

Directly applying techniques used for binary outcome spaces to large outcome spaces is computationally infeasible. In fact, it has been shown that computing and updating prices for combinatorial outcomes spaces is $\#P$ -hard [Chen et al., 2008]. Infinite outcome spaces pose a particular challenge in prediction markets since it is not clear how to define securities in such spaces. The most direct way of discretizing the space may lead to loss of information from the traders. Extending prediction market techniques to this space has been a subject of recent interest. Various techniques to bound market maker loss in both a combinatorial and infinite outcome space have been studied (see, for instance, [Abernethy et al., 2011, Othman and Sandholm, 2011]).

In this paper, we use a powerful tool from statistics for handling infinite outcome spaces: the use of parametric distributions. Specifically, we consider markets for possibly infinite outcome spaces, with the restriction that the true distribution over the outcomes belongs to an exponential family of distributions.

For such parametric families, we show that we can design a prediction market such that, when traders are strategically maximizing their own expected profit but are not adversarial, the market state can provably be interpreted as the maximum likelihood estimate of the natural parameters of the distribution. Also, this market has the desirable property of being *arbitrage-free*: There is no trade or sequence of trades that traders can exploit to make a guaranteed profit.

Next, we consider a more realistic situation in which there are both adversarial and strategic traders. We present a budget limited version of the prediction market that can be shown to suffer limited damage under concerted attacks by adversarial traders. We also exploit an information-theoretic interpretation of the particular form of the scoring rule that we use to show that an informative trader will eventually be unconstrained and he will be allowed to carry out an unrestricted trade. Taken together, we will have shown that while the market suffers limited damage from malicious traders, it is also able to make use of all the information from informative traders in the long run.

Apart from its value as a new form of prediction market, we believe that this con-

nection between prediction markets and exponential families can be exploited further to design high performance online learning algorithms. In section 6, we suggest future work in this direction.

1.1 Related Work

Designing prediction markets to handle a large outcome space is an active area of research. In [Chen et al., 2008], the authors use a restricted betting language to design efficient markets for a combinatorial outcome space. This technique is generalized by [Pennock and Xia, 2011]. [Gao et al., 2009] consider extending various automated market makers to an infinite outcome space. For the logarithmic market scoring rule they show that unbounded market maker loss can result in this setting. [Abernethy et al., 2011] and [Othman and Sandholm, 2011] specify frameworks under which they design cost function based markets that satisfy the desirable property of bounded market maker loss even in infinite outcome spaces.

The connection between machine learning and prediction markets has been studied previously. [Chen and Vaughan, 2010] and [Abernethy et al., 2011] have previously explored the connection to learning algorithms to inform the design and understanding of prediction markets. In particular, [Chen and Vaughan, 2010] consider the correspondence between prediction markets with market scoring rules and the Follow the Regularized Leader algorithm proposed by [Kalai and Vempala, 2005] and thus provide insight into the aggregation mechanism of a prediction market.

Independently of this work, Beygelzimer et al. [Beygelzimer et al., 2012] have shown that, for a particular form of binary prediction markets and traders with log-utility, the long-run dynamics of trading activity and budgets over many prediction markets lead the markets to satisfy a bounded regret property with respect to the best single trader. Our results here, and the future work we have suggested, form a program to prove bounded regret properties of budget-limited prediction markets under much more general conditions.

2 EXPONENTIAL FAMILIES

A number of popular distributions are members of a class of distributions called the exponential family. For instance, exponential, beta, Bernoulli and Poisson distributions are all members of the exponential family of distributions. In this work, we are concerned with outcomes that are drawn from an exponential family distribution.

A member of an exponential family \mathcal{F} , $P_{\beta} \in \mathcal{F}$ assigns a probability density (with respect to some base measure) over x defined over a space \mathcal{X} and has the following form:

$$P_{\beta}(x) = e^{\beta \cdot \phi(x) - \psi(\beta)}$$

Note that \mathcal{X} could be a multi-dimensional space.

Here, the set of natural parameters β is drawn from some space \mathcal{B} . $\psi(\beta)$ is known as the *log-partition function* and is defined as

$$\psi(\beta) \stackrel{\text{def}}{=} \log \int e^{\beta \cdot \phi(x)} dx$$

The vector of sufficient statistics $\phi(x)$ is some function of the data and is defined differently for different members of the exponential family.

For example, the Bernoulli distribution is a discrete probability distribution that is a member of the exponential family. It takes value 1 with probability p and value 0 with probability $1 - p$. In this case, the vector of sufficient statistics, natural parameters and log-partition function is given by:

$$\begin{aligned}\phi(x) &= x \\ \beta &= \log \frac{p}{1-p} \\ \psi(\beta) &= -\log(1-p)\end{aligned}$$

3 PREDICTION MARKETS

A prediction market is a market where traders bet on the outcome of an event and are paid according to the accuracy of their forecast. The current market state indicates the net prediction on the probability of an event. Let $i \in \{1, \dots, N\}$ be the possible (mutually exclusive, exhaustive) outcomes of the event. Let \mathbf{p} indicate the current market state. Traders in the market influence the value of \mathbf{p} and hence the predictive probability of the event.

Market scoring rules were developed as a way to allow sequential participation in a prediction market. In the Log Market Scoring Rule (LMSR), a trader who moves the current market estimate of an actual outcome i from p_{i-1} to p_i receives a (possibly negative) payoff of $\log \frac{p_i}{p_{i-1}}$.

In a cost-function based prediction market, a trader pays for the movement of the market probability upfront [Hanson, 2003]. This movement is associated with the purchase of securities. The cost of purchase is determined by a cost function. The trader is then paid according to a payoff function. Ideally, the cost function is such that it covers any possible loss that the traders may incur on the trade, thereby preventing any bankruptcy problems at the time of payoff settlement.

LMSR can also be specified as a cost-function based prediction market. Let $\{s_i\}$ be the securities issued in this market. Then q_i represents the number of securities of type s_i held by all the traders in the market and \mathbf{q} is the vector of all securities held. The cost function $C(\mathbf{q})$ is a differentiable function that depends on the market state \mathbf{q} . A trader pays $C(\mathbf{q}') - C(\mathbf{q})$ to move the market probabilities from \mathbf{q} to \mathbf{q}' . The partial derivative of $C(\mathbf{q})$ with respect to q_i represents the instantaneous price of buying an infinitesimally small quantity of security s_i at the current market state \mathbf{q} . Typically, the payoff of each security is fixed at \$1 for an outcome that occurs.

A necessary condition for the market to be arbitrage-free is that the sum of the instantaneous price of all securities be exactly equal to the payoff of each security (in this case \$1) for all states of the market. To see this we note that if this sum is less than 1 (respectively, greater than 1), then a trader could buy (respectively, sell) a bundle of securities for a guaranteed profit. Thus, arbitrage-freeness is a desirable property of a market.

We will now formalize the cost-function based market for a general payoff function $\Phi(x)$. Let x be an event whose outcome is to be determined. Let us define securities s_i whose payoff is given by $\phi_i(x)$; let $\Phi(x)$ be the payoff vector. Let q_i be the number of s_i securities purchased; let \mathbf{q} be the securities vector.

For a purchase of \mathbf{q} securities, the cost function for LMSR is defined as

$$C(\mathbf{q}) \stackrel{\text{def}}{=} \log \int e^{\mathbf{q}^T \Phi(x)} dx$$

and the instantaneous price of the securities is given by

$$p_i(\mathbf{q}) \stackrel{\text{def}}{=} \frac{\partial C(\mathbf{q})}{\partial q_i}$$

Note that a perfectly rational, risk-neutral trader with infinite budget will purchase securities until the instantaneous price of the securities is equal to the payoff he believes he will obtain.

4 AGGREGATING BELIEFS OVER A CONTINUOUS OUTCOME SPACE

In this section, we will set up a prediction market that aggregates beliefs from traders where the outcome is a continuous random variable. In particular, we assume that the outcome is drawn from an exponential family distribution. Each of the traders has access to a series of points drawn from this distribution. In other words, the traders have access an empirical mean of the sufficient statistics of the exponential family. Every trader has infinite budget so that the current market price after a trader has traded in the prediction market, exactly reflects his beliefs.

For $x_i \sim P_{\beta}$, recall that the likelihood function for independently drawn data x_1, \dots, x_n is given by $\prod_{i=1}^n P_{\beta}(x_i)$. The maximum likelihood estimate of the natural parameters β is the value of the natural parameters that maximizes the likelihood function.

We will now set up a prediction market with log market scoring rule (LMSR) and infinite budget traders such that the market state represents the maximum likelihood estimate (MLE) of the natural parameters of an exponential family distribution.

For a given exponential family distribution, the prediction market is defined as follows:

Traders The prediction market will simulate a trader i corresponding to expert i . This trader processes all information samples available to her directly or inferred from previous trades, and forms a belief distribution such that the believed means of the sufficient statistics matches the empirical means of the sufficient statistics from the information samples. The trader trades in the market to maximize her expected payoff under her believed distribution.

Securities and their payoff For each $i = 1, 2, \dots, k$, we define a security s_i with payoff $\phi_i(x)$ where x is the ultimate outcome and $\Phi(\cdot)$ defines the vector of

sufficient statistics of the exponential family distribution according to which x is drawn. We define an additional security s_0 with payoff $\phi_0(x) := a - \sum_{i=1}^k \phi_i(x)$ where a is an appropriately chosen constant dependent on the range of Φ so that the payoff of s_0 is non-negative.

We also note that since we want non-negative payoffs, we restrict the exponential families under consideration so that the sufficient statistics are lower bounded. If we have a constant lower bound on the sufficient statistics, then without loss of generality we can add a constant to each sufficient statistic without changing the exponential family in any way, as the constants will be absorbed in the log-partition function.

Arbitrage-free Property We can easily show that the prediction market we have defined is arbitrage-free: there is no sequence of trades that guarantees the trader a profit under all conditions. First, note that the incremental cost and payoff function are additive over multiple trades, so the net profit of a sequence of trades depends only on the initial and final market position, and is independent of the actual path along which the trade takes place.

Now, consider any trade (or sequence of trades) that moves the market from an initial state of \mathbf{q} to a final state of $\mathbf{q} + \mathbf{q}'$. The cost of this trade is $C(\mathbf{q} + \mathbf{q}') - C(\mathbf{q})$. The cost function of our market is the log-partition function of an exponential family, and thus, $C(\cdot)$ is a strictly convex function [Wainwright and Jordan, 2008]. Thus, we have the inequality:

$$C(\mathbf{q} + \mathbf{q}') - C(\mathbf{q}) > \mathbf{q}' \nabla C|_{\mathbf{q}} = \mathbf{q}' \mu,$$

where μ is the mean sufficient statistics vector for the distribution with parameters \mathbf{q} . The payoff due to this trade, with outcome x , is $\mathbf{q}' \cdot \phi(x)$. Now, we observe that under the distribution with parameters \mathbf{q} , the expected payoff is $\mathbf{q}' E[\Phi(x)] = \mathbf{q}' \mu$. Thus, under this distribution over outcomes x , the expected profit is strictly less than 0. This is only possible if the realized profit is less than 0 for at least one outcome x . As this statement is true for every \mathbf{q} and \mathbf{q}' , and the profits are path-independent, the market is arbitrage-free.

Market State and Natural Parameters We define the interpretation function

$$\begin{aligned} I(\mathbf{q}) &= (q_1 - q_0, q_2 - q_0, \dots, q_k - q_0) \\ &= (\beta_1, \beta_2, \dots, \beta_k) = \boldsymbol{\beta} \end{aligned} \tag{1}$$

This allows us to interpret the state of the market in terms of a prediction on the natural parameters of the distribution.

Theorem 1 *The natural parameter vector corresponding to the interpretation of the market state given by Equation 1 is the vector of their maximum likelihood estimates. Further, this interpretation is unique.*

Proof: We observe that for $i = 1, \dots, n$,

$$\begin{aligned}
\mu_i &= \left(\frac{\partial C(\mathbf{q})}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log \int \exp(\mathbf{q}^T \Phi(x)) dx}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log \int e^{q_0 \phi_0(x) + \sum_{i=1}^n q_i \phi_i(x)} dx}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log \int e^{q_0(a - \sum_{i=1}^n \phi_i(x)) + \sum_{i=1}^n q_i \phi_i(x)} dx}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log \int e^{q_0 a} e^{\sum_{i=1}^n (q_i - q_0) \phi_i(x)} dx}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \log(e^{q_0 a} \int e^{\sum_{i=1}^n (q_i - q_0) \phi_i(x)} dx)}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial(q_0 a)}{\partial q_i} \frac{\partial \log(\int e^{\sum_{i=1}^n (q_i - q_0) \phi_i(x)} dx)}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \psi(\boldsymbol{\beta})}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \psi(\boldsymbol{\beta})}{\partial \beta_i} \frac{\partial \beta_i}{\partial q_i} \right)_{\mathbf{q}=\mathbf{q}^*} \\
&= \left(\frac{\partial \psi(\boldsymbol{\beta})}{\partial \beta_i} \right)_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}
\end{aligned}$$

Thus, this choice of parameters satisfies $\frac{\partial \psi(\boldsymbol{\beta})}{\partial \beta_i} = \mu_i$. We will now argue that this vector is unique.

If the exponential family is represented so that there is a unique parameter vector associated with each distribution, the representation is said to be minimal. The Bernoulli, Gaussian, and Poisson distributions all have minimal representations. Now, for an exponential family whose representation is minimal, the gradient is an injection [Wainwright and Jordan, 2008, Prop. 3.2]. Thus, there is a unique parameter vector $\boldsymbol{\beta}$ that satisfies $\frac{\partial \psi(\boldsymbol{\beta})}{\partial \beta_i} = \mu_i$, for each $i \in \{1, \dots, n\}$. Thus, if the μ_i 's correspond to the empirical means, since our choice of $\boldsymbol{\beta}$ satisfies the equality, it must also be the vector corresponding to the MLE. ■

Thus, we have shown that this prediction market (along with its strategic traders) aggregates trader information in a way that the market state can be interpreted as a predictive distribution (the maximum likelihood distribution) over an infinite outcome space.

It is worth noting here that if we can bound the parameter space and the sufficient statistics of the distribution, the market maker defined here has bounded worst case

loss: The probability density, and hence log loss, would then be bounded for all market states that could be reached and all outcomes. In fact, for many practical applications, a distribution with unbounded parameter space is an approximation for a true space which itself may be bounded. In this case, worst case loss would thus be bounded.

5 AGGREGATING BELIEFS OVER A CONTINUOUS OUTCOME SPACE UNDER ADVERSARIAL CONDITIONS

In the previous section, we saw that we can define a cost-function based prediction market so that the aggregated belief of the traders represents the maximum likelihood estimate of the natural parameters of the true exponential family distribution.

In this section, we consider the the prediction market setup with traders that may be either informative or malicious. The malicious traders may want to inject faulty information into the market. The informative traders on the other hand receive points drawn from the true distribution on which they base their beliefs.

We will show that if we are able to impose finite initial budgets on the traders and control the market prices based on these budgets, then it possible to set up the market so that it is prohibitive for damaging traders to participate in the market. Further, the informative traders can be shown to have expected growth in budget so that they are eventually able to move the market prices without restriction.

5.1 Budget-limited Aggregation

Imposing budget limits on the traders will allow us to control the amount of influence any one trader can have on moving the market prices. We will also satisfy an additional requirement that no trader has negative budget at any point of participation in the market. This is achieved by restricting the movement of the market and hence influencing the cost incurred by the trader. Recall that the payoff in this market is non-negative and hence the only adverse influence on a trader's budget is the cost of movement of the market state.

In this section, we assume that the budget of each trader is known to the market maker, and that the market maker can directly limit the allowed trades based on a trader's budget. Let α be the budget of a trader in the market. Suppose with infinite budget, the trader would have moved the market state from \mathbf{q}_{init} to \mathbf{q} , where \mathbf{q} represents his true beliefs. Suppose further that $\alpha < C(\mathbf{q}) - C(\mathbf{q}_{init})$. In this case, we want to budget-limit the trader's influence on the market state.

We define the budget-limited final market state as $\tilde{\mathbf{q}}$. Here, we consider a specific functional form of $\tilde{\mathbf{q}}$:

$$\tilde{\mathbf{q}} = \lambda \mathbf{q} + (1 - \lambda) \mathbf{q}_{init}$$

where

$$\lambda = \min \left(1, \frac{\alpha}{C(\mathbf{q}) - C(\mathbf{q}_{init})} \right)$$

We first show that this trade is feasible given the trader's budget:

Theorem 2 *Let the current market state be given by the vector \mathbf{q}_{init} and $\tilde{\mathbf{q}} = \lambda \mathbf{q} + (1 - \lambda) \mathbf{q}_{init}$. For $\lambda = \min \left(1, \frac{\alpha}{C(\mathbf{q}) - C(\mathbf{q}_{init})} \right)$, the cost to the trader to move the market state from \mathbf{q}_{init} to $\tilde{\mathbf{q}}$ is at most his budget α .*

Proof: From the convexity of C , we have

$$C(\tilde{\mathbf{q}}) \leq (1 - \lambda)C(\mathbf{q}_{init}) + \lambda C(\mathbf{q})$$

Now

$$\begin{aligned} C(\tilde{\mathbf{q}}) - C(\mathbf{q}_{init}) &\leq (1 - \lambda)C(\mathbf{q}_{init}) \\ &\quad + \lambda C(\mathbf{q}) - C(\mathbf{q}_{init}) \\ &= \lambda (C(\mathbf{q}) - C(\mathbf{q}_{init})) \end{aligned}$$

Thus, $C(\tilde{\mathbf{q}}) - C(\mathbf{q}_{init}) \leq \alpha$. ■

We note that moving to $\tilde{\mathbf{q}}$ as defined may not be the optimal trade for a rational trader maximizing her expected profit. In general, the inequality above is strict, and so a trader does not fully exhaust her budget by moving to $\tilde{\mathbf{q}}$. Our results below will continue to hold in the case that strategic informative traders move to a position closer to their beliefs \mathbf{q} .

5.2 Damage Bound

Given this restriction on the initial budget of traders, we can now show that the total loss that can be induced by malicious traders is bounded. We define the loss function for \mathbf{q} shares held as:

$$L(\mathbf{q}, x) = -\log(p_{\beta}(x)) = \log \frac{\int e^{\mathbf{q}^T \Phi(x)} dx}{e^{\mathbf{q}^T \Phi(x)}}$$

with the correspondence between β and \mathbf{q} as defined earlier.

Suppose the prediction market runs over multiple rounds t . Let \mathbf{q}_0^t be the initial number of shares of each security that are held. Let $\tilde{\mathbf{q}}_k^t$ be the final values corresponding to the market state after the traders have made their reports. Let us assume that at this point we receive the value of the random variable x^t .

Over multiple instances of the prediction market, we can track the change in budget of each trader. The change in budget for trader i is

$$\begin{aligned} \alpha_i^t - \alpha_i^{t-1} &= C(\tilde{\mathbf{q}}_{i-1}^t) - C(\tilde{\mathbf{q}}_i^t) - (\tilde{\mathbf{q}}_{i-1}^t - \tilde{\mathbf{q}}_i^t)^T \Phi(x^t) \\ &= L(\tilde{\mathbf{q}}_{i-1}^t, x^t) - L(\tilde{\mathbf{q}}_i^t, x^t) \end{aligned}$$

Define the myopic impact of a trader i in segment t as

$$\Delta_i^t := L(\tilde{\mathbf{q}}_{i-1}^t, x^t) - L(\tilde{\mathbf{q}}_i^t, x^t)$$

Thus, the myopic impact captures incremental gain due to the trader in a round. Note that the myopic impact caused by trader i at round t is equal to the change in his budget in that round.

The total myopic impact due to all k active traders is given by

$$\Delta^t = L(\mathbf{q}_0^t, x^t) - L(\tilde{\mathbf{q}}_k^t, x^t)$$

Thus $-\Delta^t$ captures the incremental loss due to the predictive probability after aggregation of all k traders.

Theorem 3 *A coalition of b malicious traders can at most cause loss bounded by their initial budgets.*

Proof: Consider the myopic impact of a single trader i after participating in the market T times. Since the market evolves so that the budget of any trader never falls below zero, the total myopic impact in T rounds caused due to trader i is:

$$\Delta_i := \sum_{t=1}^T \Delta_i^t = \sum_{t=1}^T (\alpha_i^t - \alpha_i^{t-1}) = \alpha_i^T - \alpha_i^0 \geq -\alpha_i^0$$

Thus, any coalition of b adversaries $\{1, \dots, b\}$ can cause at most $\sum_{i=1}^b \alpha_i^0$ damage. ■

This means that if it can be made prohibitively expensive for an attacker to generate clones, we can set up the prediction market with mostly informative traders.

In Section 5.4 we show that for an informative trader in every round, his budget increases in expectation. We provide an information-theoretic justification for this claim in Section 5.3. The intuition behind this claim is that his prediction moves the input moves the market probability closer to the true probability distribution resulting in net expected profit.

5.3 An Information-Theoretic Interpretation

We observe a useful alternative view of the market scoring rule prediction market for exponential family learning. We connect the cost, payoff and profit function to information-theoretic quantities associated with the exponential family.

The following result has been previously pointed out by Amari [Amari, 2001].

Theorem 4 (Profit Decomposition): *Consider an exponential family \mathcal{F} of distributions over some set of statistics $\phi(x)$, with natural parameters β . Let $\pi, \rho \in \mathcal{F}$ be any two probability distributions in the family. We use β_π to denote the natural parameters of π and μ_π to denote the expected value of the sufficient statistics under π . Likewise, we can define β_ρ and μ_ρ . We abuse notation slightly and let $\psi(\rho)$ indicate the log partition function of ρ which technically depends on its natural parameters. Let $H(\pi)$ denote the entropy of the distribution π , and $K(\pi||\rho)$ denote the KL-divergence of ρ relative to π . Then, the following equality holds:*

$$K(\pi||\rho) + H(\pi) = \psi(\rho) - \beta_\rho \cdot \mu_\pi \quad (2)$$

Proof:

$$\begin{aligned}
K(\pi||\rho) + H(\pi) &= \int_x \pi(x) \log \frac{\pi(x)}{\rho(x)} dx \\
&\quad - \int_x \pi(x) \log \pi(x) dx \\
&= - \int_x \pi(x) \log \rho(x) dx \\
&= - \int_x \pi(x) [\beta_\rho \cdot \phi(x) - \psi(\rho)] dx \\
&= -\beta_\rho \cdot \int_x \pi(x) \phi(x) dx \\
&\quad + \psi(\rho) \int_x \pi(x) dx \\
&= \psi(\rho) - \beta_\rho \cdot \mu_\pi
\end{aligned}$$

■

Equation 2 gives us an alternative view of the market scoring rule construction. Assume that π is the true distribution, and consider two arbitrary distributions $\rho_1, \rho_2 \in \mathcal{F}$. Note that $\psi(\rho)$ is independent of π , and $\beta_\rho \cdot \mu_\pi$ is linear in the probabilities $\pi(x)$. If we want to measure loss by the KL-divergence, we can do so (in expectation) by setting a cost function that captures the first term, and defining security quantities (β) and payoffs (ϕ) to capture the second term. In particular, in a market with cost function ψ , if the market price initially implies a distribution ρ_1 , and a trader moves the market to price that implies a distribution ρ_2 , then the cost she incurs is $\psi(\rho_2) - \psi(\rho_1)$. The number of securities bought to make this trade is given by the vector $(\beta_{\rho_2} - \beta_{\rho_1})$, and the expected payoff of the securities are given by μ_π . Thus, by equation 2, the *net profit* of the trader is equal to $K(\pi||\rho_1) - K(\pi||\rho_2)$, *i.e.*, the reduction in KL-divergence with respect to the true distribution. We note one useful property of this construction: For a fixed vector β of purchased securities, the cost is independent of the outcome (and outcome distribution π), while the payoff is independent of the initial market state in which these securities were purchased.

5.4 Budget of Informative Traders

Given the information-theoretic interpretation of the cost-function based prediction market, we now show that the informative trader in the prediction market defined above increases his budget in a round in expectation *under his own belief distribution*.

We now characterize the expected change in budget for an informative trader. The following result holds for any round t ; for simplicity, we have therefore dropped the superscript from the notation.

Theorem 5 *Suppose that each informative trader gets a random sample of data, resulting in a sequence of trader beliefs \mathbf{q}_i , and hence a sequence of budget-limited market positions $\tilde{\mathbf{q}}_i$, before a final outcome x . Then, the expectation, over trader i 's belief*

distribution \mathbf{q}_i , of trader i 's realized profit is greater than zero whenever her budget is positive and her belief differs from the previous market position $\tilde{\mathbf{q}}_{i-1}$.

Proof: Trader i 's believed distribution is the distribution parametrized by \mathbf{q}_i . Therefore, the expected profit, over possible outcome values x , for a trader i in a given round is given by

$$\begin{aligned} & E[C(\tilde{\mathbf{q}}_{i-1}) - C(\tilde{\mathbf{q}}_i) - (\tilde{\mathbf{q}}_{i-1} - \tilde{\mathbf{q}}_i)\Phi(x)] \\ &= K(\mathbf{q}_i|\tilde{\mathbf{q}}_{i-1}) + H(\mathbf{q}_i) - [K(\mathbf{q}_i|\tilde{\mathbf{q}}_i) + H(\pi)] \\ &= K(\mathbf{q}_i|\tilde{\mathbf{q}}_{i-1}) - K(\mathbf{q}_i|\tilde{\mathbf{q}}_i) \end{aligned}$$

We recall that $\tilde{\mathbf{q}}_i$ can be expanded as:

$$\tilde{\mathbf{q}}_i = \lambda \mathbf{q}_i + (1 - \lambda)\tilde{\mathbf{q}}_{i-1}$$

where λ is strictly greater than 0, but no more than 1. It is a standard result from information theory that $K(\cdot, \cdot)$ is a convex function, and therefore, we have:

$$K(\mathbf{q}_i|\tilde{\mathbf{q}}_{i-1}) \leq \lambda K(\mathbf{q}_i|\mathbf{q}_i) + (1 - \lambda)K(\mathbf{q}_i|\tilde{\mathbf{q}}_i)$$

The first term on the right hand side is zero, and thus, we get:

$$K(\mathbf{q}_i|\tilde{\mathbf{q}}_{i-1}) - K(\mathbf{q}_i|\tilde{\mathbf{q}}_i) \geq \lambda K(\mathbf{q}_i|\tilde{\mathbf{q}}_{i-1})$$

■

For continuous distributions with a density, the probability that a trader with private information will form exactly the same beliefs as the current market position is 0, and thus, each trader will have positive expected profit on almost all sequences of observed samples and beliefs. This result suggests that, eventually, every informative trader will have the ability to influence the market state in accordance with his beliefs, without being budget limited.

We note one important aspect of Theorem 5: The expectation is taken with respect to each trader's belief at the time of trade, rather than with respect to the true distribution. This is needed because we have made no assumptions about the optimality of the traders' belief updating procedure; depending on the distribution family and the prior distribution over parameter values, maximum likelihood estimation might not optimize the true expected score. If we assume that the traders' belief formation is optimal, then this growth result will extend to the true distribution as well.

6 CONCLUSIONS AND FUTURE WORK

We have defined a new class of cost-function based prediction markets for a continuous random variable distributed according to an exponential family. Our first result shows that if traders in this market believe sufficient statistics drawn from true distribution, then the market price provably represents the MLE of natural parameters. Under some restrictions on the distribution, the market maker will also have limited worst case loss.

We then show that if we can budget limit the traders so that their influence on the market is constrained by their budgets, malicious traders can induce only bounded loss. We also show that eventually informative agents will be able to fully influence the market state in accordance with their beliefs.

This last result provides a technique that can be applied to constructing bounded regret learning algorithms. Online learning algorithms based on prediction markets are attractive for particular domains. Specifically, expert forecasts and recommendations in internet settings present the following challenges: Some forecasters may provide best-effort forecasts, but in some cases the experts may have vested interests in manipulating the system. Second, advice from different experts often arrive haphazardly over time allowing later experts to clone the predictions of earlier experts. Current expert advice algorithms cannot distinguish between genuine forecasters and clones, even though the clones are forced to make forecasts later than the genuine forecasters they copy. We believe that prediction markets can be naturally applied to the sequential forecasting setting to design high-performance algorithms in this domain.

However, a scoring rule that only considers the impact of the expert on the predicted probability distribution on the target variable is bound to fail in the following way. An adversarial entity may control multiple experts that clone the reports of every honest expert. Consider such an expert who arrives after some honest expert in the sequence. Cloning the honest report will not lead to an increase in his score since he hasn't affected the prediction. However, although his imitation of this expert may not immediately lead to an expected increase in his score, he effectively reduces the impact on the prediction of any honest and informative expert arriving later in the sequence. Thus, if the adversarial entity now clones this later honest expert, he can siphon off some of the increase in budget due to her.

This attack is possible because the experts are scored based on their impact on the prediction. Further, they are not penalized if they adversely affect the impact of later (honest) experts in the sequence, so long as they do not move the prediction. Thus, this possible non-myopic attack needs to be considered in the assigned scores. We leave the details to future work.

References

- [Abernethy et al., 2011] Abernethy, J., Chen, Y., and Wortman Vaughan, J. (2011). An optimization-based framework for automated market-making. In *Proceedings of the 12th ACM conference on Electronic commerce, EC '11*, pages 297–306, New York, NY, USA. ACM.
- [Amari, 2001] Amari, S.-I. (2001). Information geometry on hierarchy of probability distributions. *Information Theory, IEEE Transactions on*, 47(5):1701–1711.
- [Beygelzimer et al., 2012] Beygelzimer, A., Langford, J., and Pennock, D. M. (2012). Learning performance of prediction markets with kelly bettors. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 3, AAMAS '12*, pages 1317–1318, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

- [Chen et al., 2008] Chen, Y., Goel, S., and Pennock, D. M. (2008). Pricing combinatorial markets for tournaments. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, STOC '08, pages 305–314, New York, NY, USA. ACM.
- [Chen and Vaughan, 2010] Chen, Y. and Vaughan, J. W. (2010). A new understanding of prediction markets via no-regret learning. In *Proceedings of the 11th ACM conference on Electronic commerce*, EC '10, pages 189–198, New York, NY, USA. ACM.
- [Gao et al., 2009] Gao, X., Chen, Y., and Pennock, D. M. (2009). Betting on the real line. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, WINE '09, pages 553–560, Berlin, Heidelberg. Springer-Verlag.
- [Hanson, 2003] Hanson, R. (2003). Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119.
- [Kalai and Vempala, 2005] Kalai, A. and Vempala, S. (2005). Efficient algorithms for online decision problems. *J. Comput. Syst. Sci.*, 71:291–307.
- [Othman and Sandholm, 2011] Othman, A. and Sandholm, T. (2011). Automated market makers that enable new settings: Extending constant-utility cost functions. In *Proceedings of the Second Conference on Auctions, Market Mechanisms and Their Applications (AMMA)*, pages 19–30.
- [Pennock and Xia, 2011] Pennock, D. and Xia, L. (2011). Price updating in combinatorial prediction markets with bayesian networks. In *Proceedings of the Twenty-Seventh Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-11)*, pages 581–588, Corvallis, Oregon. AUAI Press.
- [Wainwright and Jordan, 2008] Wainwright, M. J. and Jordan, M. I. (2008). Graphical models, exponential families, and variational inference. *Found. Trends Mach. Learn.*, 1:1–305.