

Yoopick: A Combinatorial Sports Prediction Market

Sharad Goel and David Pennock and Daniel Reeves and Cong Yu

Yahoo! Research, 111 W. 40th Street, New York, NY, 10018

Abstract

We describe Yoopick, a combinatorial sports prediction market that implements a flexible betting language, and in turn facilitates fine-grained probabilistic estimation of outcomes.

Introduction

Several methods have been developed for predicting outcomes of sporting events, including simple polling, machine learning techniques, and Las Vegas style betting exchanges. Studies show that betting markets and related *prediction markets*, designed to forecast everything from elections to flu outbreaks, yield remarkably accurate probability estimates (Berg *et al.* 2001; Wolfers & Zitzewitz 2004).

Two of the most common types of sports bets are *money line* bets and *point spread* bets. In money line bets, agents bet on which team will win a given game, with payoffs based on the likelihood of that team winning. In particular, betting on the favorite team pays less than betting on the underdog. In spread betting, gamblers wager on whether the final point difference between the competing teams will be greater than or less than the value specified by a bookmaker. If the bookmaker sets the spread at 3.5, for example, a bet on the favorite team pays if the favorite wins by more than 3.5 points. As bets are placed, the bookmaker continually adjusts the spread so that gamblers are equally likely to bet on either side of the spread.

With money line and point spread bets, bookmakers manually set the odds. A traditional prediction market automates that process by creating a single security that agents buy and sell. In the case of money line bets, for example, the security would pay \$1 if a given team wins a particular game. The current price of the security is the market probability of the team winning. Point spread bets correspond to a security that pays at even odds if the actual spread is on one side of a threshold, where traders negotiate the threshold using a double auction instead of the price. The threshold in this case is the market-derived median point spread.

Money line and point spread markets produce, respectively, probability estimates of a given team winning and the expected point difference. One may desire, however, detailed information regarding the entire point spread proba-

bility distribution (i.e., the probability Team *A* beats Team *B* by exactly *k* points). An approach for eliciting this information is to create separately traded securities, one for each possible point difference. Each security would pay \$1 precisely when its corresponding point difference were achieved, and its current price would indicate the market probability of that outcome. The significant disadvantage of this mechanism is that information does not automatically propagate between securities even though they are clearly related. As an extreme illustration, the asset prices do not necessarily even sum to \$1, creating potential arbitrage opportunities.

Yoopick is a *combinatorial prediction market* implementation that directly allows estimation of the entire point spread probability distribution within in a single unified market (Chen *et al.* 2007; Chen, Goel, & Pennock 2008; Hanson 2003). Agents select a point interval, and bet on the final point difference landing in that interval. Odds for intervals are calculated via the logarithmic market scoring rule (Hanson 2007). In particular, bets on disjoint intervals affect one another. This combinatorial market maker mechanism provides agents with a flexible language for injecting information into the market. Agents can mimic money line bets by selecting the interval $[0, \infty)$, and point spread bets by selecting the interval $[m, \infty)$ where *m* is the median of the market distribution. This flexibility in turn facilitates fine-grained predictions. Yoopick was designed as a Facebook application, and is available for beta trial at <http://apps.facebook.com/yoopick>.

Market Scoring Rules

Prediction markets are speculative markets designed to elicit forecasts. Because the goal is to gather information, an automated market maker that expects to lose a small and bounded amount of money on average may actually help attract informed traders.

Asset prices in Yoopick are determined by Hanson's logarithmic market scoring rule market maker (Hanson 2007), which has several advantages over call market designs. Agents trade with the market maker, who sets asset prices and who accepts all buy and sell orders at these prices. In particular, market makers help reduce both the *thin market* and *irrational participation* problems that affect stock market style exchanges. The thin market problem arises

when agents have to coordinate which assets they will trade with each other, as is the case in call markets. In Yoopick, since agents are allowed to bet on any point interval, the liquidity added by a market maker is essential to support the large number of bet types. ‘No-trade’ theorems for zero-sum games (Milgrom & Stokey 1982) state that rational agents, after hedging their risks, will no longer trade with each other even when they hold private information. Market makers avoid this irrational participation issue by, in essence, subsidizing the market.

A market scoring rule maintains a probability distribution over an outcome space Ω which reflects a consensus estimate of the likelihood of any event. Market scoring rules may be implemented as market-maker driven exchanges in which traders buy and sell securities of the form “Pays \$1 if ω occurs”. All transaction costs are paid to a market maker who agrees to honor the contracts. Let $q : \Omega \mapsto \mathbb{R}$ indicate the number of outstanding shares on each state. If a trader wishes to change the number of outstanding shares from q to \tilde{q} (i.e., wants to buy or sell shares) the cost of the transaction under the logarithmic market scoring rule (Hanson 2007) is $C(\tilde{q}) - C(q)$ where

$$C(q) = b \log \sum_{\tau \in \Omega} e^{q(\tau)/b}.$$

The parameter b is the liquidity, or depth, of the market. When b is large, it becomes more expensive for any particular agent to move the market distribution. If there are q outstanding shares, the spot price for shares on a given outcome ω is

$$P_q(\omega) = \frac{d}{dq(\omega)} C(q) = \frac{e^{q(\omega)/b}}{\sum_{\tau \in \Omega} e^{q(\tau)/b}}$$

which is interpreted as the aggregate, market-generated probability estimate for ω .

Moving outstanding shares from q to \tilde{q} results in a (state-dependent) net pay out to the agent of

$$\begin{aligned} & [\tilde{q}(\omega) - q(\omega)] - [C(\tilde{q}) - C(q)] \\ &= b \log \left(\frac{e^{\tilde{q}(\omega)/b}}{\sum_{\tau \in \Omega} e^{\tilde{q}(\tau)/b}} \right) - b \log \left(\frac{e^{q(\omega)/b}}{\sum_{\tau \in \Omega} e^{q(\tau)/b}} \right) \\ &= b \log P_{\tilde{q}}(\omega) - b \log P_q(\omega). \end{aligned}$$

In other words, for moving the prices from P_q to $P_{\tilde{q}}$, an agent receives payment $b \log P_{\tilde{q}}$ in exchange for agreeing to pay the last agent who interacted with the market.

The market maker opens the market with an initial distribution of shares on states (e.g., $q_0 \equiv 0$). If the market closes with a distribution of shares \tilde{q} , the market maker makes a net payment to traders of

$$\begin{aligned} & [\tilde{q}(\omega) - q_0(\omega)] - [C(\tilde{q}) - C(q_0)] \\ &= b \log P_{\tilde{q}}(\omega) - b \log P_{q_0}(\omega) \\ &\leq b \log(1/P_{q_0}(\omega)). \end{aligned}$$

In particular, for $q_0 \equiv 0$, the market maker’s maximum loss is bounded by $b \log |\Omega|$.

Yootles Currency

The Yoopick mechanism itself is currency-agnostic but the name derives from the fact that bets are placed in *yootles* (Reeves, Soule, & Kasturi 2006; Ghosh *et al.* 2007) using the Yootles API (<http://yootles.com/api>) and is being launched with a partner Facebook application, <http://apps.facebook.com/yootles>.

Future Work

Yoopick aggregates information regarding the entire point spread probability distribution. Furthermore, its flexible betting language encourages agents to offer subtle information. We plan to compare the predictions made by Yoopick to those of traditional single-asset prediction markets and professional bookmakers.

The net payment of the market maker to agents depends on the opening market distribution. Effectively, the market pays agents for improving upon the initial distribution. We currently open markets with a mean zero normal distribution and sport-specific variance inferred from historical data. We plan to experiment with more sophisticated machine learning methods for generating this initial distribution. By charging traders a modest commission, our goal is to design a system that produces accurate predictions while operating at zero net loss.

References

- Berg, J. E.; Forsythe, R.; Nelson, F. D.; and Rietz, T. A. 2001. Results from a dozen years of election futures markets research. In Plott, C. A., and Smith, V., eds., *Handbook of Experimental Economic Results* (forthcoming).
- Chen, Y.; Fortnow, L.; Nikolova, E.; and Pennock, D. M. 2007. Betting on permutations. In *ACM Conference on Electronic Commerce*. New York: ACM.
- Chen, Y.; Goel, S.; and Pennock, D. M. 2008. Pricing combinatorial markets for tournaments. In *Symposium on Theory of Computing*.
- Ghosh, A.; Mahdian, M.; Pennock, D. M.; Reeves, D. M.; and Fugger, R. 2007. Mechanism design on trust networks. In *Workshop on Internet and Network Economics*.
- Hanson, R. D. 2003. Combinatorial information market design. *Information Systems Frontiers* 5(1):107–119.
- Hanson, R. D. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets* 1(1):1–15.
- Milgrom, P., and Stokey, N. L. 1982. Information, trade and common knowledge. *Journal of Economic Theory* 26(1):17–27.
- Reeves, D. M.; Soule, B. M.; and Kasturi, T. 2006. Yootopia! *SIGecom Exchanges* 6:1–26.
- Wolfers, J., and Zitzewitz, E. 2004. Prediction markets. *Journal of Economic Perspective* 18(2):107–126.