MS&E 235: Lecture 12
Prof. Ashish Goel
May 12, 2009
Notes by Rio Goodman

Lecture 12: Information/Prediction Markets

In this lecture we will consider the same type of prediction market that was introduced last lecture. We will also see some emerging markets, and discuss their potential.

A Simple Prediction Market

In our simple prediction market, a trader can initially buy one "yes" and one "no" share for \$1. There is then a trading phase, preceding the event. And then, when the event is observed, if the outcome is "yes," then traders are rewarded \$1 for each "yes" share they hold, and if the outcome is "no," traders are rewarded \$1 for each "no" share they hold. In this scheme, a trader breaks even if he/she does not make any trades.

Consider a prediction market where there are N traders. Let p_i denote trader i's estimate of the probability that a "yes" outcome will occur. We will order the p_i's in increasing order so that $p_1 \leq p_2 \leq ... \leq p_N$. Let c_y denote the equilibrium price of a "yes" share.

Scenario 1: Every trader wants to hold exactly[1] two shares.

If $p_i > c_y$ then trader i will want to hold only "yes" shares. For example, if $p_i = 0.8$ and $c_y = 0.6$, trader i expects to make \$0.80 from a "yes" share, so he/she would be happy to buy a "yes" share for \$0.60.

Note that $c_n + c_y = 1$, because c_y and c_n are equilibrium prices, and if $c_n + c_y \neq 1$, there would be arbitrage opportunities. So at our example prices, $c_n = 0.40$.

If $p_j < c_y$, then trader j will want to sell a "yes" share and buy a "no" share, so that he/she holds two "no" shares.

Therefore, the number of traders with $p_i > c_y$ must be equal to the number of traders with $p_i < c_y$. Thus, c_y must be the **median** estimate. Thus a prediction market, by finding the equilibrium prices, actually finds the median estimate of the probability of the event occurring.

Scenario 2: Every trader wants to hold shares in proportion to $|p_i - c_y|$

In this scenario, if a trader's belief is far from the current price, he/she will buy many shares. Whereas if the trader believes that the price is very close to the correct probability, he/she will buy few shares.

In equilibrium,

$$\sum_{i:p_i>c_y} (p_i - c_y) = \sum_{i:p_i<c_y} (c_y - p_i)$$

Therefore

$$\sum_{i:p_i>c_y} (p_i - c_y) - \sum_{i:p_i<c_y} (c_y - p_i) = 0$$

$$\sum_{i:p_i>c_y} (p_i - c_y) + \sum_{i:p_i<c_y} (p_i - c_y) = 0$$

[1]The same results will hold if every trader wants to hold *at most* two shares

$$\sum_i (p_i - c_y) = 0$$

$$\sum_i p_i - N c_y = 0$$

$$c_y = \frac{1}{N} \sum_i p_i$$

Therefore, in scenario 2, c_y is the **mean** estimate.

The aggregation of information to find statistics such as the median and mean is why prediction markets are so useful for predicting outcomes. Prediction markets are also a good metaphor for the stock market.

Other Interesting Markets

We have studied a number of online markets, all of which have been very fleshed out. But there are some interesting emerging markets that are not so fleshed out. There is an Amazon service called Mechanical Turk, that lists a number of low-paying jobs that can be completed (for a few cents) by users. The types of jobs available are "translate this page," or "answer questions about this article." The jobs are called HITs (Human Intelligence Tasks) because they are typically very easy for a human, but difficult for a computer.

There is an online game called "The ESP Game," where players are randomly paired up with a partner that they don't know, who is also playing. The two partners are then shown a picture, and asked to enter words associated with that picture. If the two players enter the same word, then they get points. The advantage of such a game is that the host can collect words associated with the images that are shown. Machines are currently not very good at associating words with images, so this is a novel new way to gain information through a fun game. Since the players are anonymous, there is no way to collude in this game, thus the best strategy is to guess the words that you think best describe the image. This is called "anonymized consensus."

One malicious example of anonymized consensus has to do with new email accounts. When you sign up for an email address, you are asked to translate a funny-looking word or phrase. This is simple for a human, but difficult for a bot that is creating many email accounts (for spam maybe). Imagine that someone with a legitimate online business also wishes to create many email accounts using a bot. If he/she posts the same funny-looking word on their legitimate website, they can have users translate it for them. Whenever more than one user translates an image the same way, the business owner knows that this must be the correct answer.

How could these new markets be combined to form another market? Users could be paid small amounts of money to translate things that are difficult for computers, and then their answers could be cross-examined with other users' answers to determine their legitimacy. For example, the DMV could use such a system for determining whether or not to ticket drivers. Another example is a game very similar to the ESP game, where instead of earning points, users can earn real money.