Mathematics in Prediction Markets

Ilan Shlain

Want to know what someone will say? You ask them a question. Want to know what they believe? You make a bet with them; money talks. In today's online world, institutions are faced with an overwhelming amount of information pertaining to public opinion about almost everything. This information, while having tremendous potential, is accompanied by an equally significant amount of unfiltered noise. In order to make insights out of the data at hand, institutions spend countless resources filtering out unnecessary factors that would have no significant impact on the true beliefs of the population. A great example of this type of approach would be political polls: organizations use statistical techniques to select a sample to be surveyed, and apply the received sentiments to the population. Aside from being relatively expensive to complete, polling has a speed factor that is causing it to fall behind an emerging alternative in analytics: prediction markets.

At their most basic level, prediction markets are platforms where users can buy and sell event-based contracts. These contracts have a great breadth in terms of the events they can relate to. Presidential elections, prices of commodities, and sports outcomes are just a few of the many areas prediction markets allow people to speculate on. In a similar fashion to the stock market, the prices for these contracts, in theory, are determined by the consensus of independent speculations, which eventually approach a stable true value.

In this paper we strive to show how event contract prices can be mathematically interpreted as probabilities. For this purpose, we will explore the mechanics of prediction markets and try to understand the conditions on which our mathematical models rely.

Do they work?

Before just accepting the idea that contract prices are to be interpreted as probabilities, it may be motivating to see examples in which prediction markets are shown to have an edge. For such examples, we will turn to *The Promise of Prediction Markets* [Arr+08], a *Policy Forum* piece which posits that there is mounting evidence for prediction markets being able to achieve lower prediction error than traditional forecasting methodologies. In particular, we reference their figure on "information revelation through time," which uses data from the Iowa Electronic Markets for the 1988, 1992, 1996, and 2000 presidential elections. The figure plots the average absolute difference between the market prediction and the actual two-party vote share as Election Day approaches. What makes this figure especially useful is that it shows prediction markets not merely identifying who is favored, but producing estimates that move closer to the true outcome over time.

Arrow et al. explain this advantage through the idea that information is often widely dispersed among individuals, and that markets provide a mechanism through which this scattered information can be collected and reflected in price. This gives a practical reason to take prediction markets seriously before moving into the mathematical question of whether their prices can in fact be read as probabilities. Rather than asking individuals directly what they think will happen, prediction markets force participants to act on their beliefs in a setting where being correct has monetary value. In that sense, the market price is meant to condense a broad range of information into a single numerical forecast.

The figure also gives a more concrete sense of this forecasting performance. In the week immediately before the election, the market erred by an average of 1.5 percentage points, compared with 2.1 percentage points for the final Gallup poll. Moreover, even 150 days before the election, the market's average error was only about 5 percentage points, which Arrow et al. note is impressive given that polls interpreted that far in advance tend to perform much worse as forecasts. Thus, the figure suggests that prediction markets are able to aggregate dispersed information in a way that becomes increasingly accurate as more information enters the market. This gives empirical support to the claim that these markets are not merely speculative curiosities, but are capable of serving as serious forecasting tools.

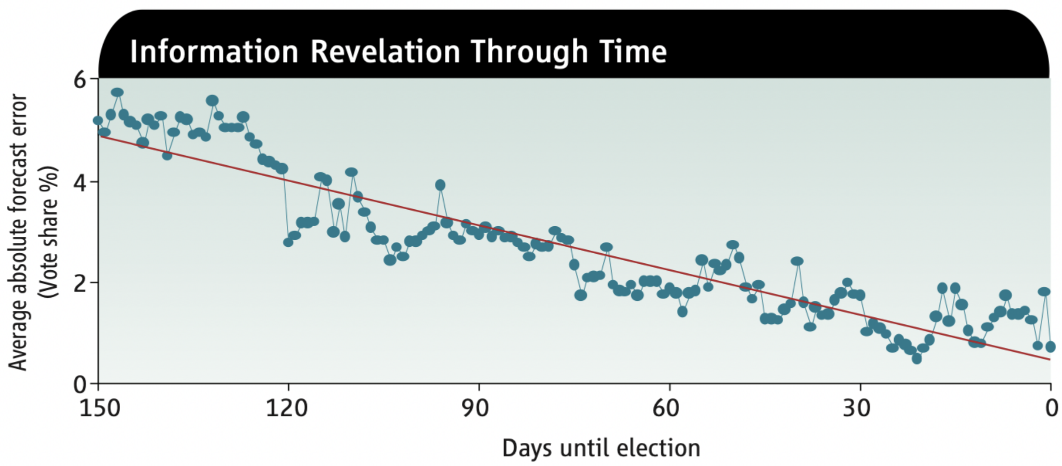

Figure 1: Information revelation through time in the Iowa Electronic Markets. The graph shows the average absolute forecast error in predicting the two-party vote share for the 1988, 1992, 1996, and 2000 U.S. presidential elections. As Election Day approaches, the market prediction moves closer to the true outcome, indicating improved forecasting accuracy over time.

Before moving on, it is worth stating the simplest mathematical interpretation of a prediction market price. If a contract pays

$$X = \begin{cases} 1, & \text{if the event occurs,} \\ 0, & \text{if the event does not occur,} \end{cases}$$

then its expected value is

$$E[X] = 1 \cdot p + 0 \cdot (1 - p) = p,$$

where p is the probability of the event. Under this basic model, the contract price is therefore interpreted as approximately equal to the probability of occurrence:

$$\text{Price} \approx p.$$

Thus, a contract trading at 0.57 would naturally be read as implying about a 57% chance of the event occurring. As the next section will show, however, this interpretation is not always exact.

Why do they work?

Understanding why prediction market prices can generally be interpreted as probabilities is important because, as we've seen, these markets are often used as forecasting tools. If the price of a contract can reliably represent the likelihood of an event, then observers can treat market prices as quantitative predictions rather than simply speculative prices. Wolfers and Zitzewitz, in *Prediction Markets in Theory and Practice*,[WZ06] analyze this question by studying how prediction market prices relate to the beliefs of the traders participating in the market. Their work shows that, under many reasonable assumptions, the market price tends to approximate the average belief among traders about the probability that an event will occur.

To see why this interpretation matters, consider a simple example. Suppose a contract is trading at 0.75, which implies a 75% chance that the event will occur. If you have additional information that leads you to believe the event is actually 90% likely, then buying the contract has positive expected value. In particular, your expected payoff from buying one contract is

$$E[\text{profit}] = 0.90 \cdot (1 - 0.75) + 0.10 \cdot (0 - 0.75) = 0.90 \cdot 0.25 - 0.10 \cdot 0.75 = 0.15.$$

On the other hand, if your belief matches the market's implied probability of 75%, then the expected value becomes

$$E[\text{profit}] = 0.75 \cdot 0.25 - 0.25 \cdot 0.75 = 0.$$

In this case, there is no expected gain from trading, and therefore no incentive to buy or sell the contract.

The mathematical idea behind this result is that each trader enters the market with a personal belief q about the probability of an event. Traders buy contracts when they believe the event is more likely than the market price suggests, and sell contracts when they believe it is less likely. As these trades occur, the market price adjusts until buying and selling balance. At equilibrium, the price reflects the aggregation of all individual beliefs. In the simplest model this leads to the result

$$\pi = \int q f(q)\, dq,$$

meaning that the market price π equals the average belief about the probability of the event.

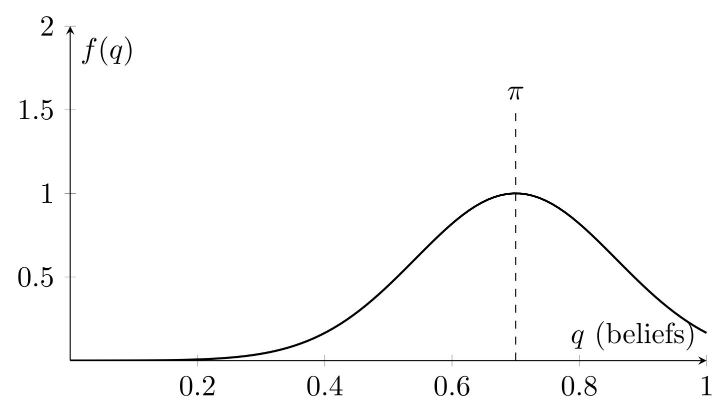

Figure 2: Distribution of trader beliefs $f(q)$ over probabilities $q \in [0, 1]$. The dashed line at π denotes the market price, interpreted as the expected value of the belief distribution. Figure generated by the author.

Are the probabilities always accurate?

Now that we've seen the underlying principles which give prediction markets their utility, it is important to look at the shortcomings of such systems and taking them into account when interpreting probabilities. Here we will present a paper discussing a key limitation that may arise even under relatively ideal market conditions. In particular, this limitation must be considered when interpreting prediction market prices as probabilities.

I selected *Aggregation of Information and Beliefs in Prediction Markets* by Marco Ottaviani and Peter Norman Sørensen (2007) because it directly addresses whether prediction market prices accurately reflect the probability of an event. Much of the existing literature treats market prices as though they naturally correspond to probabilities, but this paper shows that the relationship is more complicated. In particular, the authors discuss what is known as the *favorite–longshot bias*, which is a systematic discrepancy between market prices and true probabilities. What makes this paper especially useful for our purposes is that it explains this discrepancy as the result of an underlying mathematical structure, rather than mere randomness or market noise.

A central mathematical idea in the paper is that traders update their beliefs when new information arrives, and this can be represented through likelihood ratios:

$$\frac{\pi_i}{1-\pi_i} = \frac{q_i}{1-q_i} L,$$

where q_i represents a trader's prior belief, L represents new information, and π_i is the updated belief. While individual traders may update rationally, the market price itself does not necessarily adjust one-for-one with this new information. In other words, the market price moves in the correct direction, but not always by the full amount that a direct probability interpretation would suggest. This underreaction provides a mathematical explanation for the favorite–longshot bias.

The paper illustrates this result with a graph comparing the equilibrium market price to the fully Bayesian posterior probability. The diagonal line represents the benchmark case in which market price and probability are exactly equal. The curved line shows that the actual relationship can depart from this ideal case. For longshots, the market price lies above the posterior probability, meaning that unlikely events are priced too high relative to their true probability. For favorites, the market price lies below the posterior probability, meaning that likely events are priced too low. The figure therefore gives a visual representation of the paper's main point: market prices may contain useful information, but they should not always be interpreted as exact probabilities.

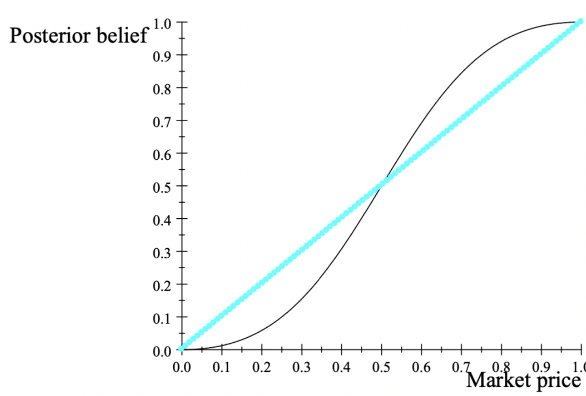

Figure 3: Posterior probability for event E as a function of the market price. The diagonal line represents the benchmark case in which price equals probability, while the curved line shows the fully Bayesian posterior belief [OS07].

Till Technical Assumptions do us Part

Although we have shown instances of prediction markets being incredibly powerful analytical tools, explained the underlying intuitions, and confronted some of their shortcomings, we have only scratched the surface. When it comes to the true nature of prediction markets, there are countless assumptions and much more complicated mathematics at play. Our general assumptions of market efficiency and widely dispersed information take us a long way in explaining the most basic relationships, but they do not capture everything that is going on.

In reality, there are many different types of participants in these markets. Beyond people using outside information to inform their decisions, there are also those studying the structure of prediction markets themselves, as well as phenomena such as the favorite–longshot bias, in order to find opportunities for arbitrage.

This leaves us with a useful but incomplete picture. Prediction markets can often be interpreted as probabilities, and they do a surprisingly good job at aggregating information, but that interpretation depends on assumptions that do not always hold. Understanding where these models work, and where they begin to break down is what makes this topic interesting, and is also where further reading into the structure and behavior of prediction markets is highly encouraged.

References

[Arr+08] Kenneth J. Arrow et al. “The promise of prediction markets”. In: *Science* 320.5878 (2008), pp. 877–878.

[OS07] Marco Ottaviani and Peter Norman Sørensen. “Aggregation of information and beliefs in prediction markets”. In: *The American Economic Review* 97.2 (2007), pp. 64–69.

[WZ06] Justin Wolfers and Eric Zitzewitz. “Prediction markets in theory and practice”. In: *NBER Working Paper No. 12083* (2006).