Prediction Markets in ESPR

Nuño Sempere

1 Prediction Markets: A Soft Introduction

A prediction market starts as a sheet of paper with an initial probability estimate about a future event, for example:

Will Donald Trump Win the 2016 US Presidential Elections?

Participant	% Yes	% No	Odds
Player 1	5%	95%	1:19

When another player writes a different probability estimate, she will win or lose points according to the amount of information she adds to the market. For example

Will Donald Trump Win the 2016 US Presidential Elections?

Participant	% Yes	% No	Odds
Player 1	5%	95%	1:19
Suzanne	10%	90%	1:9

If Trump then goes to win, she has added information, because her probability of it happening was higher than the one before. If she had written 1%, or if Trump had lost, she would have taken information from the market, and thus would be penalized.

This is all.

2 Prediction Markets: A Difficult to Hack Scoring Rule

The concept of information is important, because if Peter, Paul and Mary had written:

Will Donald Trump Win the 2016 US Presidential Elections?

Participant	% Yes	% No	Odds
Player 1	5%	95%	1:19
Superpredictor Suzanne	10%	90%	1:9
Peter	10%	90%	1:9
Paul	10%	90%	1:9
Mary	10%	90%	1:9

Then we don't want them to win any points; they're just copying the previous estimate.

Additionally, we also want to incentivize people to write their best estimate of the probability, instead of the ones that net them the most point. That is, we don't want

q = Probability Written Down By(Participant)
and p = Probability Thought In Her Head(Participant)

to be different; we want people to report their real probabilities. A bad scoring rule, which leads to this being the case, might be:

$$\text{Reward}(q) = q^2$$

where q is the probability, in %, written below the event which ended up happening, like above. If that was the scoring rule, even when people's probability estimates were something like 60%, they would expect to get the most points when writing 100%. They could even write 100% - 0% and then 0% - 100% for the other side, to maximize points.
¿ So we want a reward which makes people expect to win the most points when writing down the best probability estimate they can get. A possible one, and the one which we'll be using here, is:

$$\text{Reward}(q) = k \cdot log(\tfrac{q}{b}) = log(q) - log(b)$$

where b is the probability written before. The constant k is 5/log(2); if you double the probability assigned to something and the thing happens, you get 5 tokens, and if you've halved it, you lose 5 tokens.

If you think that an event has a likelihood of p, what is the expected reward if you sneakily write down q?

$$\text{ExpectedReward}(p, q) = log(\tfrac{q}{b}) \cdot p + log(\tfrac{1-q}{1-b}) \cdot (1-p)$$

If one takes the derivative of the reward with respect to q, a maximum is found at p=q! Thus the best option is to write down your best estimate.

3 Prediction Markets: Common Strategies

3.1 Use more optimization power than everyone else

Last year's winner, Q, took some time going around asking people to do push ups when the question was about pushups, and generally spending some time considering each question. Some words of wisdom: Getting a probability directly out of your ass might not result in great estimates.

3.2 Steal other people's optimization power

Suppose that we had the following prediction market:

Will someone be able to trade up from one paperclip to something worth ≥ 100 pounds on adventure day (the last day of camp)?

Participant	% Yes	% No	Odds
Player 1	5%	95%	1:19
Santiago	10	90	1:9
Severus	40	60	2:3
Sybill	99	1	99:1
Sophie	60	40	3:2
Suzanne	?	?	?

Suzanne can now take the average of everyone before her and arrive at 43%, and then extremize it a little, say to 35%. See Philip E. Tetlock's Superforecasting book for details.

Or perhaps she might want to take the geometric, instead of the arithmetic mean, because Sybill thinks she can predict the future and is thus overconfident, which drags the average up a lot. Taking the geometric mean, she'd get a 26%, instead of a 43%, or you might want to take Sybill out of your equation entirely. In any case, taking *somehow* into account what previous people have written is not stupid.

3.3 Arbitrate Between Suckers

Suppose that Sybill thinks she can see the future, but so does Cassandra, and they both predict different things, so they start a prediction war.

Will any student, staff member, or group thereof complete a $\geq 10,000$ words novella during camp?

Participant	% Yes	% No	Odds
Player 1	10%	90%	1:9
Sybill	99%	1%	99:1
Cassandra	1%	99%	1:99
Sybill	99%	1%	99:1

Suzanne notices and starts to arbitrage:

Cassandra	1	99	1:99
Suzanne	**50**	**50**	**1:1**
Sybill	99	1	99:1
Suzanne	**50**	**50**	**1:1**

Suzanne notices, and writes a probability which is in the middle, 50%. If the thing happens, Suzzane gets:

$$\text{k} \cdot [log(\frac{50}{99}) + log(\frac{50}{1})] = 23 \text{ units.}$$

Note that the first factor, $log(\frac{50}{99})$ is negative, whereas the second is positive. And k is, as before $5/log(2)$. At any point, Suzanne *also gets the same amount if the thing doesn't happen.*

3.4 Don't be a sucker

In general, when suckers write $99 : 1$, there isn't usually such a high likelihood, so it's often profitable to bring the probability down a little. Consider:

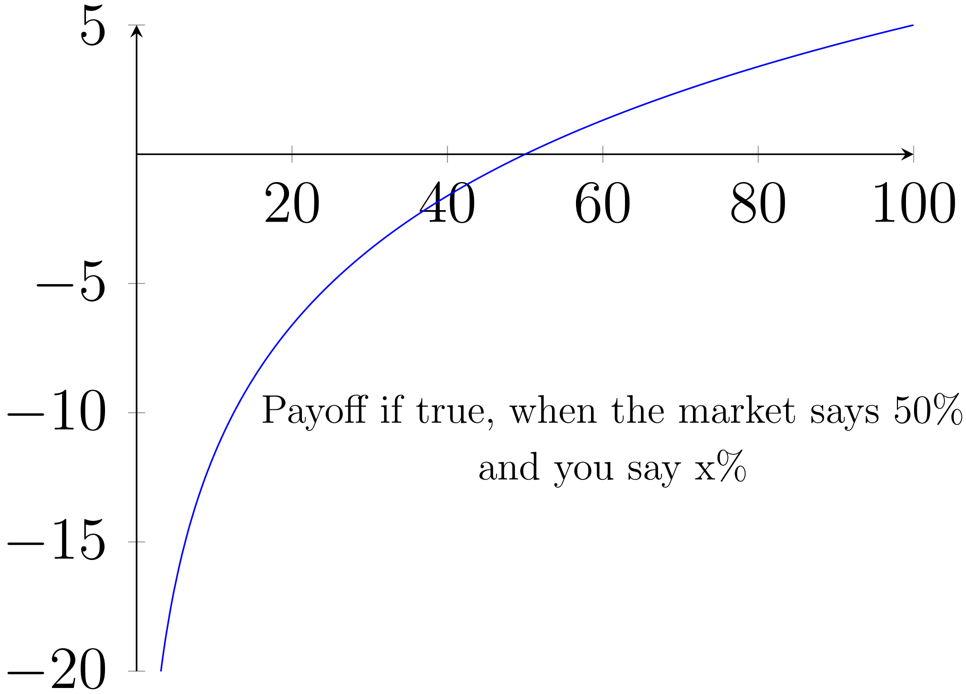
Will 2 randomly selected participants cooperate on a prisoner's dilemma problem with the payoffs in the example at Wikipedia?

Participant	% Yes	% No	Odds
Player 1	50%	50%	1:1
Sybill	99%	1%	99:1
Suzanne	80%	20%	4:1

Suzanne loses 1.5 units if the two randomly selected students cooperate, but wins 21 if they don't cooperate, whereas Meanwhile, Cassandra wins 4.9 units if they do, but loses 28.2 if they don't. We see that the payoffs are very extremized. But more than that, Cassandra could have gotten most of the payoff with less of the downside by writing 90% or 95%.

Cassandra is a sucker, but you don't have to be. Don't be a sucker.

$$f(x) = k \cdot log(x/50)$$



$$f(x) = k \cdot log((100 - x)/50)$$