story so far...

- why prediction markets
  - initial studies

- market mechanism design
  - LMSR
  - cost function markets

- implementation
  - play money vs real money
  - prediction markets in the wild
road map

• (proper) scoring rules and MSR market makers

• information theoretic properties of log score

• equivalent formulation in terms of cost function
Why LMSR?
properties of
LMSR prediction markets
Uncertain future event
Forecaster predicts $p$
The true state of the world $x$ is revealed
We pay forecaster according to score:
$$S(p; x)$$
(examples we have seen: quadratic, log)
Quadratic Scoring Rule

Tomorrow’s (unknown) weather $\in \{\text{rain, shine}\}$
Forecaster predicts $p = 0.7$
The true state of the world $x \in \{0, 1\}$ is revealed
We pay forecaster according to the **quadratic score**:

$$S(p; x) = -(x - p)^2$$
Log Scoring Rule

Forecaster predicts $p$
The true state of the world $x \in \{0, 1\}$ is revealed
We pay forecaster according to the log score:

$$S(p;x) = \begin{cases} 
\log p & \text{if } x = 1 \\
\log (1-p) & \text{if } x = 0
\end{cases}$$
Logarithmic Scoring Rule is Proper
properties of the logarithmic scoring rule
entropy

\[ H(p) = -\sum_{i=1}^{n} p_i \log p_i \]

For a binary random variable

\[ x = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases} \]

Compute \( H(p) \) when

- \( p = 0.5 \)
- \( p = 0 \)
- \( p = 1 \)
entropy

\[ H(p) = - \sum_{i=1}^{n} p_i \log p_i \]

\[ H(p) = - p \log p - (1-p) \log (1-p) \]
entropy is non-negative

\[ H(p) = - \sum_{i=1}^{n} p_i \log p_i \]

Prove that \( H(p) \geq 0 \)
expected log score

more certain, greater expected score
LMSR: key idea

sequential shared log scoring rule
### LMSR: example

#### Chance of sunshine
- 0.25
- 0.75
- 0.75
- 0.95

#### Pay
- log 0.5
- log 0.25
- log 0.75
- log 0.75

#### Market Maker
- log 0.25/0.5
- log 0.75/0.25
- log 0.75/0.75
- log 0.95/0.75

#### Paid by Market Maker
- log 0.25
- log 0.75
- log 0.75
- log 0.95
relative entropy

\[ D(p\|q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \]

“distance” between distributions
asymmetry

$$D(p||q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}$$

For binary random variables
let $p = 1/2$ and $q = 1/4$
• Compute $D(p||q)$
• Compute $D(q||p)$

In general, $D(p||q) \neq D(q||p)$
Jensen’s inequality

\[ E[f(X)] \leq f(E[X]) \]

\( f \) is concave
\( X \) is rv

Consequence

\[ E[\log(X)] \leq \log(E[X]) \]

\[ D(p \parallel q) \geq 0 \]