Eliciting Information from the Crowd

Part of
the EC'13 Tutorial on Social Computing and User-Generated Content

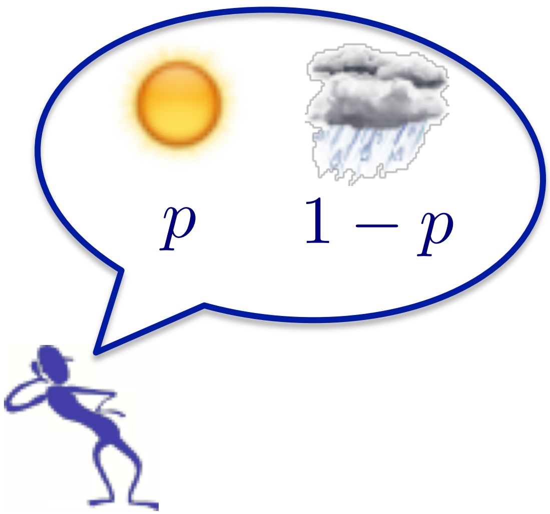
Yiling Chen
Harvard University

June 16, 2013

Roadmap

- Eliciting information for events with verifiable outcomes
  - Proper scoring rules
  - Market scoring rules
- Eliciting unverifiable information
  - Peer prediction
  - Bayesian Truth Serum
  - Robust Bayesian Truth Serum

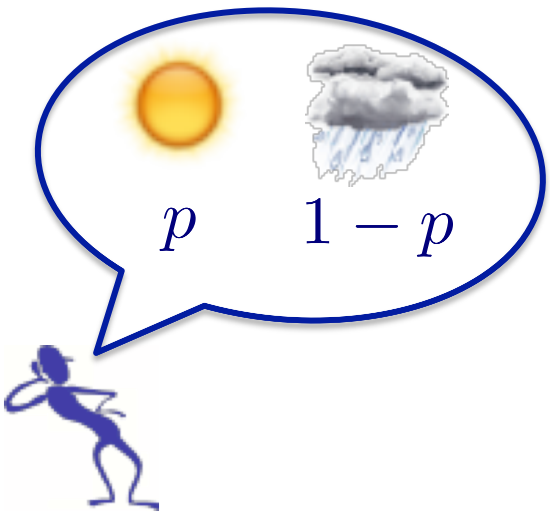
How to Evaluate Weather Forecasts



- Specification of a scale of goodness for weather forecasts
- "… the forecaster may often find himself in the position of choosing to … let it (the verification system) do the forecasting for him by 'hedging' or 'playing the system.' "
- "one essential criterion for satisfactory verification is that the verification scheme should influence the forecaster in no undesirable way"

[Brier 1950]

The Brier Scoring Rule [Brier 1950]



- $S(p,\ \text{sunny}) = 1-(1-p)^2$
  $S(p,\ \text{rainy}) = 1-p^2$
- Expected score of prediction p given belief q

$$\begin{aligned} S(p,q) &= q\,S(p,\ \text{sunny}) + (1-q)S(p,\ \text{rainy}) \\ &= 1-q(1-p)^2-(1-q)p^2 \\ &= 1-q+q^2-(p-q)^2 \end{aligned}$$

- Predicting q maximizes the expected score

$$\arg\max_p S(p,q) = q$$

$S(p,q)$ increases as $|p-q|$ decreases.

The Brier Scoring Rule

- State $\omega \in \{1, 2, \ldots, n\}$
- Prediction $\mathbf{p} = (p_1, p_2, \ldots, p_n)$
- Brier score for prediction $\mathbf{p}$ in state ω is

ω-th indicator vector

$$S(\mathbf{p}, \omega) = a_\omega - b||\mathbf{e}^\omega - \mathbf{p}||^2$$

Parameters, b >0.
Affine transformation does not change incentives.

- Also called the quadratic scoring rule

Other Strictly Proper Scoring Rules

- Logarithmic scoring rule

$$S(\mathbf{p}, \omega) = \log p_\omega$$

- Spherical scoring rule

$$S(\mathbf{p}, \omega) = \frac{p_\omega}{||\mathbf{p}||}$$

Proper Scoring Rules

- Scoring rule $S(\mathbf{p}, \omega)$
- Expected score for prediction $\mathbf{p}$ given belief $\mathbf{q}$:

prediction

$$S(\mathbf{p}, \mathbf{q}) := \sum_{\omega=1}^{n} q_\omega S(\mathbf{p}, \omega)$$

belief

A scoring rule $S(\mathbf{p}, \omega)$ is proper if and only if

$$S(\mathbf{q}, \mathbf{q}) \geq S(\mathbf{p}, \mathbf{q}).$$

It is strictly proper if the inequality is strict unless $\mathbf{q} = \mathbf{p}$.

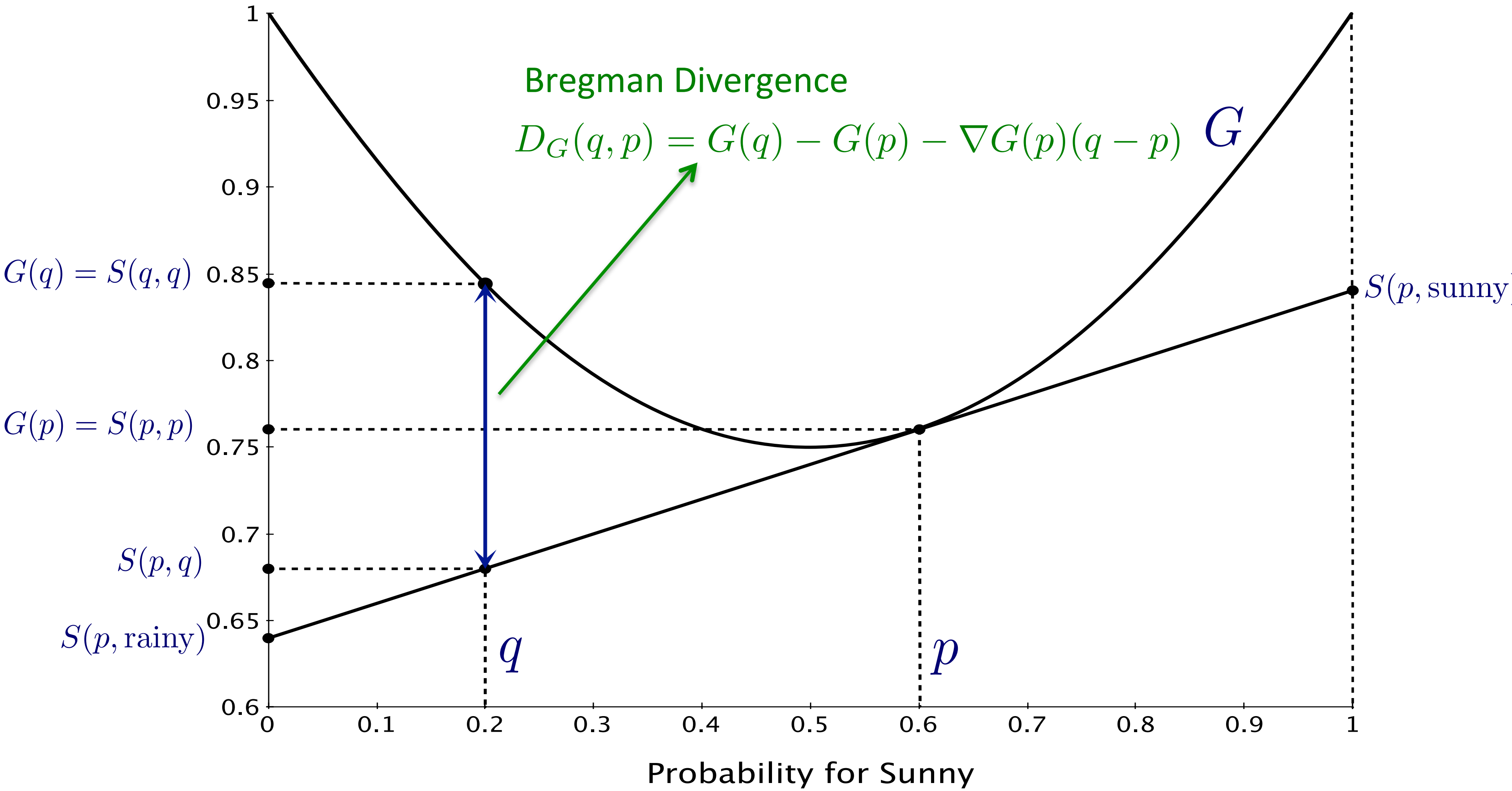
Proper scoring rules are dominant-strategy incentive compatible for risk-neutral agents.

Geometric Interpretation

Brier Scoring Rule

$$S(p,\ \text{sunny}) = 1 - (1-p)^2$$
$$S(p,\ \text{rainy}) = 1 - p^2$$

$$G(p) = S(p,p) = (p - 1/2)^2 + 3/4$$



McCarthy, Savage Characterization

［McCarthy 1956］ [Savage 1971]

A scoring rule $S(\mathbf{p},\omega)$ is (strictly) proper if and only if

$$S(\mathbf{p},\omega) = G(\mathbf{p}) - \nabla G(\mathbf{p}) \cdot \mathbf{p} + \nabla G_\omega(\mathbf{p}) \ ,$$

where G is a (strictly) convex function and $\nabla G(\mathbf{p})$ is a subgradient of G and $\nabla G_\omega(\mathbf{p})$ is its $\omega\text{-th}$ component.

- Expected score

$$S(\mathbf{p},\mathbf{p}) = \sum_\omega p_\omega \left(G(\mathbf{p}) - \nabla G(\mathbf{p}) \cdot \mathbf{p} + \nabla G_\omega(\mathbf{p})\right) = G(\mathbf{p})$$

- In terms of Bregman Divergence, for differentiable G,

$$\begin{aligned} S(\mathbf{p},\omega) &= G(\mathbf{p}) - \nabla G(\mathbf{p}) \cdot \mathbf{p} + \nabla G(\mathbf{p}) \cdot \mathbf{e}^\omega \\ &= G(\mathbf{e}^\omega) - \left(G(\mathbf{e}^\omega) - G(\mathbf{p}) - \nabla G(\mathbf{p}) \cdot (\mathbf{e}^\omega - \mathbf{p})\right) \\ &= G(\mathbf{e}^\omega) - D_G(\mathbf{e}^\omega, \mathbf{p}) \end{aligned}$$

Common Strictly Proper Scoring Rules

- Brier scoring rule $S(\mathbf{p},\omega) = 1 - ||\mathbf{e}^{\omega} - \mathbf{p}||^2$

$$G(\mathbf{p}) = ||\mathbf{p}||^2 \qquad D_G(\mathbf{q},\mathbf{p}) = ||\mathbf{q} - \mathbf{p}||^2$$

- Logarithmic scoring rule $S(\mathbf{p},\omega) = \log p_\omega$

$$G(\mathbf{p}) = \sum_{\omega} p_\omega \log p_\omega \qquad D_G(\mathbf{q},\mathbf{p}) = \sum_{\omega} q_\omega \log \frac{q_\omega}{p_\omega}$$

- Spherical scoring rule $S(\mathbf{p},\omega) = \dfrac{p_\omega}{||\mathbf{p}||}$

$$G(\mathbf{p}) = ||\mathbf{p}|| \qquad D_G(\mathbf{q},\mathbf{p}) = ||\mathbf{q}|| - \frac{\mathbf{q}\cdot\mathbf{p}}{||\mathbf{p}||}$$

Other Work on Scoring Rules

- Proper scoring rules for continuous variables

  [Matheson and Winkler 1976; Gneiting and Raftery 2007]

- Proper scoring rules for properties of distributions (e.g. mean, quintiles)

  [Savage 1971; Lambert et al. 2008; Abernethy and Frongillo 2012]

- Scoring rules for more complex environments
  - Agents can affect the event outcome

    [Shi et al. 2009; Bacon et al. 2012]
  - A decision will be made based on the elicited information

    [Othman and Sandholm. 2010; Chen and Kash 2011; Boutilier 2012]

One Expert → Multiple Experts

Market Scoring Rules (MSR)

[Hanson 2003, 2007]

- Sequential, shared version of proper scoring rules
- Reward improvements on prediction

- Select a proper scoring rule $S(\mathbf{p}, \omega)$
- Market opens with an initial prediction $\mathbf{p}^0$
- Participants sequentially change the market prediction
- Participant who changes the prediction from $\mathbf{p}^{t-1}$ to $\mathbf{p}^t$ receives payment

$$S(\mathbf{p}^t, \omega) - S(\mathbf{p}^{t-1}, \omega)$$

An Example: LMSR Market

$$S(\mathbf{p}, \omega) = 5 \log p_\omega$$

0.5	0.6	0.8	0.4	0.9	
					t
0.5	0.4	0.2	0.6	0.1	

5log(0.9)-5log(0.5)	~~5log(0.6)-~~ 5log(0.5)	~~5log(0.8)-~~ ~~5log(0.6)~~	~~5log(0.4)-~~ ~~5log(0.8)~~	5log(0.9)- ~~5log(0.4)~~
5log(0.1)-5log(0.5)	~~5log(0.4)-~~ 5log(0.5)	~~5log(0.2)-~~ ~~5log(0.4)~~	~~5log(0.6)-~~ ~~5log(0.2)~~	5log(0.1)- ~~5log(0.6)~~

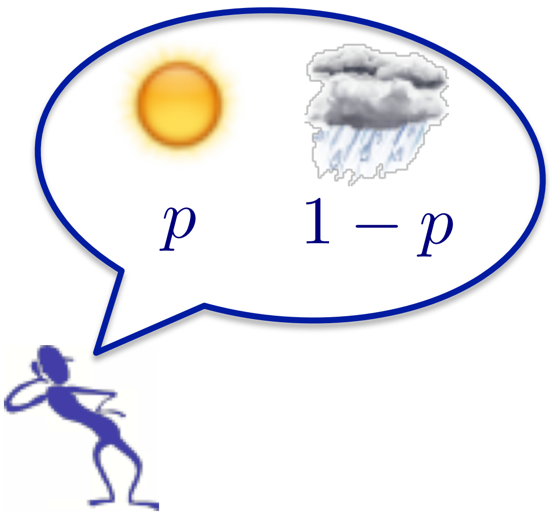
The mechanism only pays the final participant!

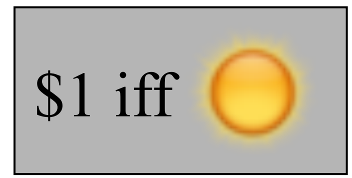
Properties of MSR

- Bounded loss (subsidy)
  - Eg. Loss bound of LMSR with $S(\mathbf{p}, \omega) = b \log p_\omega$ is $b \log n$
- Incentive compatible for myopic participants
- Recent work studies information aggregation in MSR with forward-looking participants

  [Chen et al. 2010; Ostrovsky 2012; Gao et al. 2013]
  - Market as a Bayesian extensive-form game
  - Conditions under which information is fully aggregated in the market

Prediction and Trading





Buy this contract if $\text{price} < p$

Sell this contract if $\text{price} > p$

MSR as Automated Market Makers

[Chen and Pennock 07]

- One contract for each outcome

$$\boxed{\$1 \text{ iff } \omega}$$

- Payments are determined by a cost potential function $C(Q)$
  - Q_i is the current number of shares of the contract for outcome ω that have been purchased
  - Current cost of purchasing a bundle R of shares is

$$C(Q+R) - C(Q)$$

  - Instantaneous prices $p_\omega(Q) = \dfrac{\partial C(Q)}{\partial Q_\omega} \longrightarrow$ Market Prediction

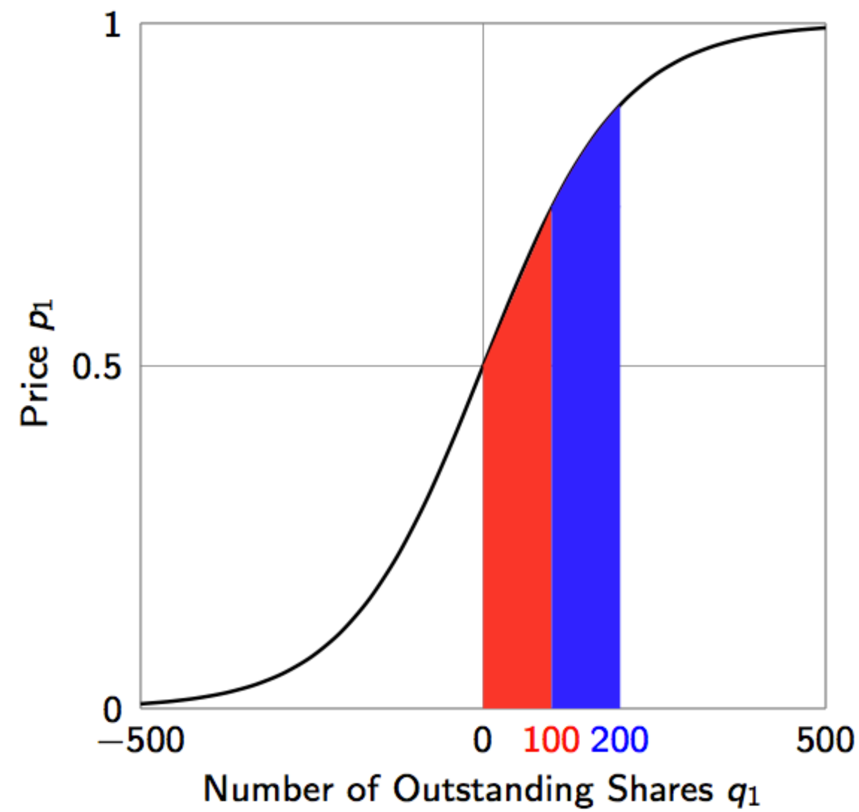
Example

- LMSR with $S(\mathbf{p}, \omega) = b \log p_\omega$
  - Cost function

$$C(Q) = b \log \sum_\omega e^{Q_\omega / b}$$

  - Price functions

$$p_\omega(Q) = \log \frac{e^{Q_\omega / b}}{\sum_{\omega'} e^{Q_{\omega'} / b}}$$



LMSR Equivalence

$$C(Q) = \log \sum_{\omega} e^{Q_\omega}$$

$$p_\omega(Q) = \frac{e^{Q_\omega}}{\sum_{\omega'} e^{Q_{\omega'}}}$$

$$S(\mathbf{p}, \omega) = \log p_\omega$$

- Profit of a participant who changes Q to $\bar{Q}$ in state ω

$$(\bar{Q}_\omega - Q_\omega) - (C(\bar{Q}) - C(Q))$$

$$= (\bar{Q}_\omega - C(\bar{Q})) - (Q_\omega - C(Q))$$

$$= (\log e^{\bar{Q}_\omega} - \log \sum_{\omega'} e^{\bar{Q}_{\omega'}}) - (\log e^{Q_\omega} - \log \sum_{\omega'} e^{Q_{\omega'}})$$

$$= \log \frac{e^{\bar{Q}_\omega}}{\sum_{\omega'} e^{\bar{Q}_{\omega'}}} - \log \frac{e^{Q_\omega}}{\sum_{\omega'} e^{Q_{\omega'}}}$$

$$= \log p_\omega(\bar{Q}) - \log p_\omega(Q)$$

$$= S(\mathbf{p}(\bar{Q}), \omega) - S(\mathbf{p}(Q), \omega)$$

Cost Function MM ⟷ Strictly Proper MSR

[Abernethy et al. 2013]

- An MSR using a strictly proper scoring rule $S(\mathbf{p}, \omega)$
- The expected scoring function $G(\mathbf{p}) = S(\mathbf{p}, \mathbf{p})$
- The corresponding cost function MM

$$C(Q) = \sup_{\mathbf{p} \in \Delta_n} Q \cdot \mathbf{p} - G(\mathbf{p})$$

Conjugate Duality

$$\mathbf{p}(Q) = \nabla G(Q) = \arg\max_{\mathbf{p} \in \Delta_n} Q \cdot \mathbf{p} - G(\mathbf{p})$$

- Profit equivalence holds when $\mathbf{p}(Q)$ is in the relative interior of the probability simplex

Eliciting Unverifiable Information

Unverifiable Information

SurveyMonkey®

Is this website age-appropriate for children?

- Yes
- No

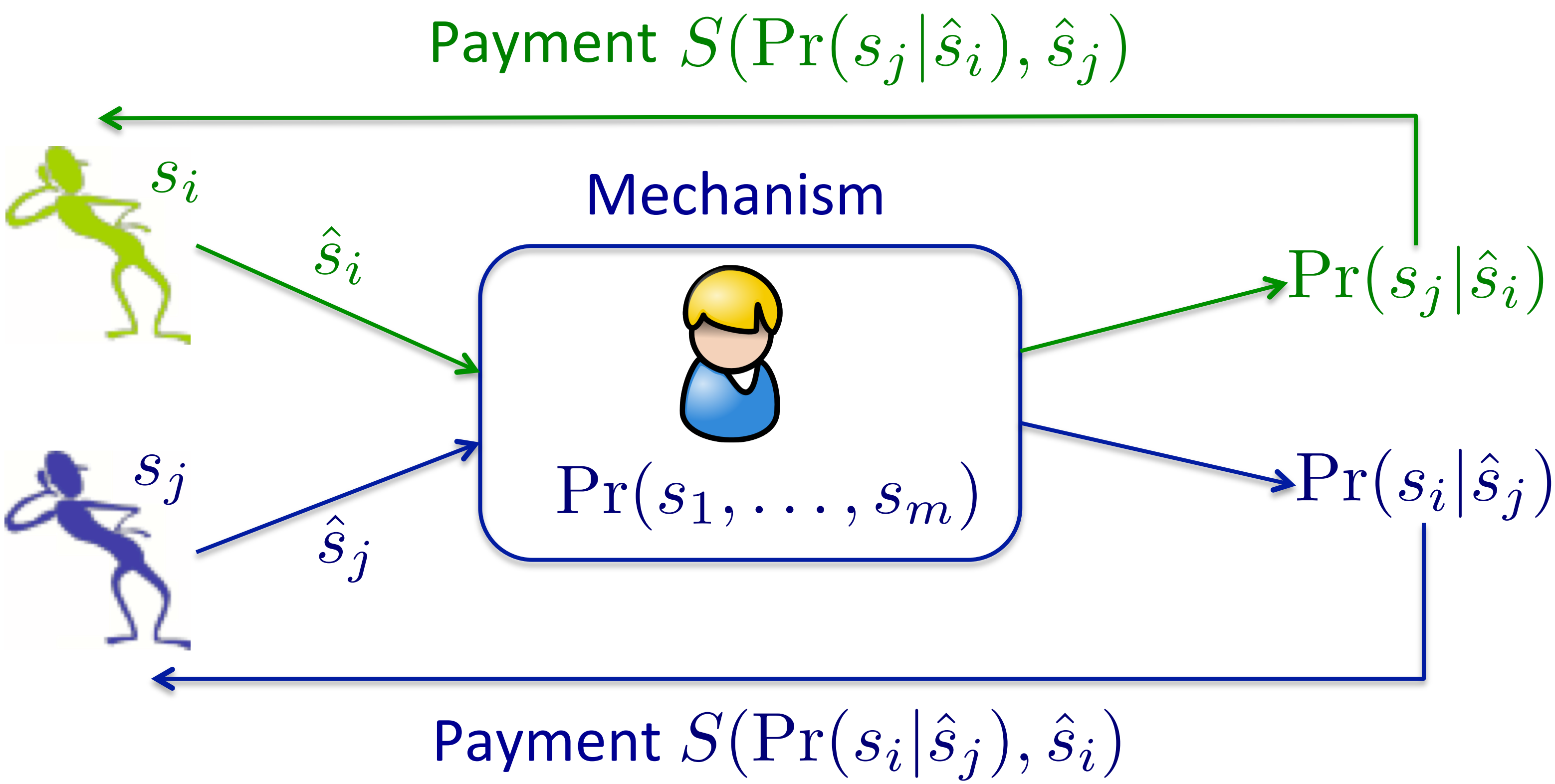
The General Setting

- State $\omega \in \{1, 2, \ldots, n\}$
  - E.g. whether this website is really age-appropriate or not
- $m \geq 2$ agents
- Each agent i receives a private signal $s_i \in \mathcal{S}$
  - E.g. whether agent i thinks the website age-appropriate
- Prior distribution $\Pr(\omega, s_1, \ldots, s_m)$ is common knowledge for all agents
- Goal: truthfully revealing their private signal s_i is a strict Bayesian Nash equilibrium

The Peer Prediction Mechanism

[Miller et al. 2005]

- Require that the mechanism knows the prior
- Each agent has a reference agent



The Peer Prediction Mechanism

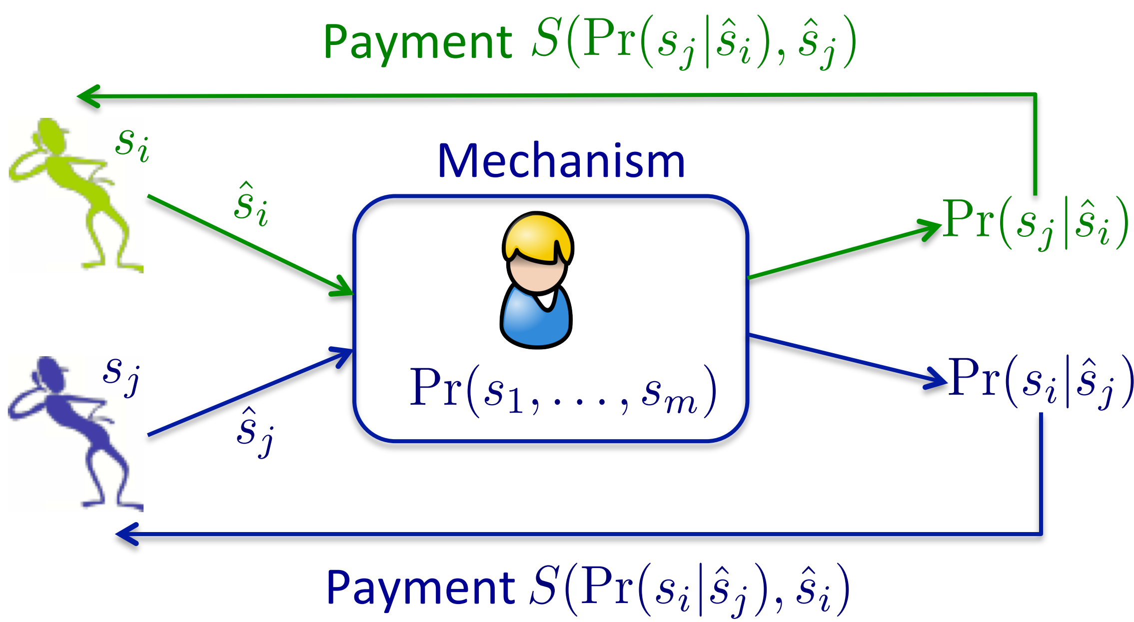
[Miller et al. 2005]

- $S(\mathbf{p}, \hat{s})$ is a strictly proper scoring rule
- Technical assumption: Stochastic Relevance

$$\Pr(s_j | s_i) \neq \Pr(s_j | s_i'), \ \ \forall s_i \neq s_i'$$

Truthful reporting is a strict Bayesian Nash equilibrium.

Requiring the mechanism know the prior is undesirable!



Bayesian Truth Serum (BTS)

[Prelec 2004]

Is this website age-appropriate for children? (Yes, No)

- Each agent is asked for two reports
  - Information report: $\mathbf{x}^{\mathbf{i}}$ is an indicator vector, having value 1 at its $k\text{-th}$ element if agent i reports signal k
  - Prediction report: $\mathbf{y}^{i}$ is agent i's prediction of the frequency of reported signals in the population
- The mechanism calculates

$$\bar{\mathbf{x}}_k = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} {\mathbf{x}^{\mathbf{i}}}_k \qquad \log \bar{\mathbf{y}}_k = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log {\mathbf{y}^{\mathbf{i}}}_k$$

- Score for agent i:

$$\sum_k {\mathbf{x}^{\mathbf{i}}}_k \log \frac{\bar{\mathbf{x}}_k}{\bar{\mathbf{y}}_k} + \sum_k \bar{\mathbf{x}}_k \log \frac{{\mathbf{y}^{\mathbf{i}}}_k}{\bar{\mathbf{x}}_k}$$

Bayesian Truth Serum (BTS)

- Technical assumptions: conditional independence of signals and stochastic relevance

When $n \to \infty$, truthful reporting is a strict Bayesian Nash equilibrium.

For sufficiently large n, truthful reporting is a strict Bayesian Nash equilibrium. But n depends on the prior.

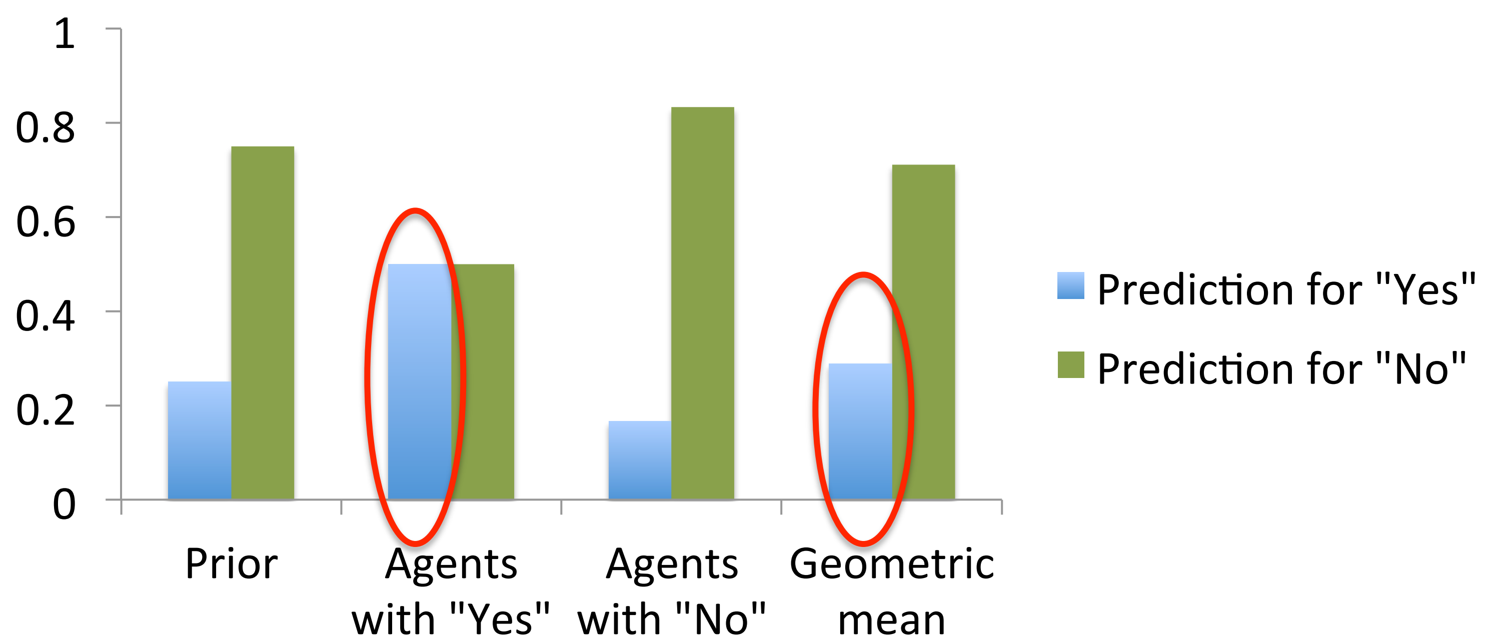
Information Score $\sum_k \mathbf{x^i}_k \log \frac{\bar{\mathbf{x}}_k}{\bar{\mathbf{y}}_k} + \sum_k \bar{\mathbf{x}}_k \log \frac{\mathbf{y^i}_k}{\bar{\mathbf{x}}_k}$ Prediction Score

Log scoring rule!
Truthful prediction reports

BTS: "Surprisingly Common"

- True signal maximizes $\sum_k \mathbf{x^i}_k \log \frac{\bar{\mathbf{x}}_k}{\bar{\mathbf{y}}_k}$ in expectation

Consider an agent with "yes" signal



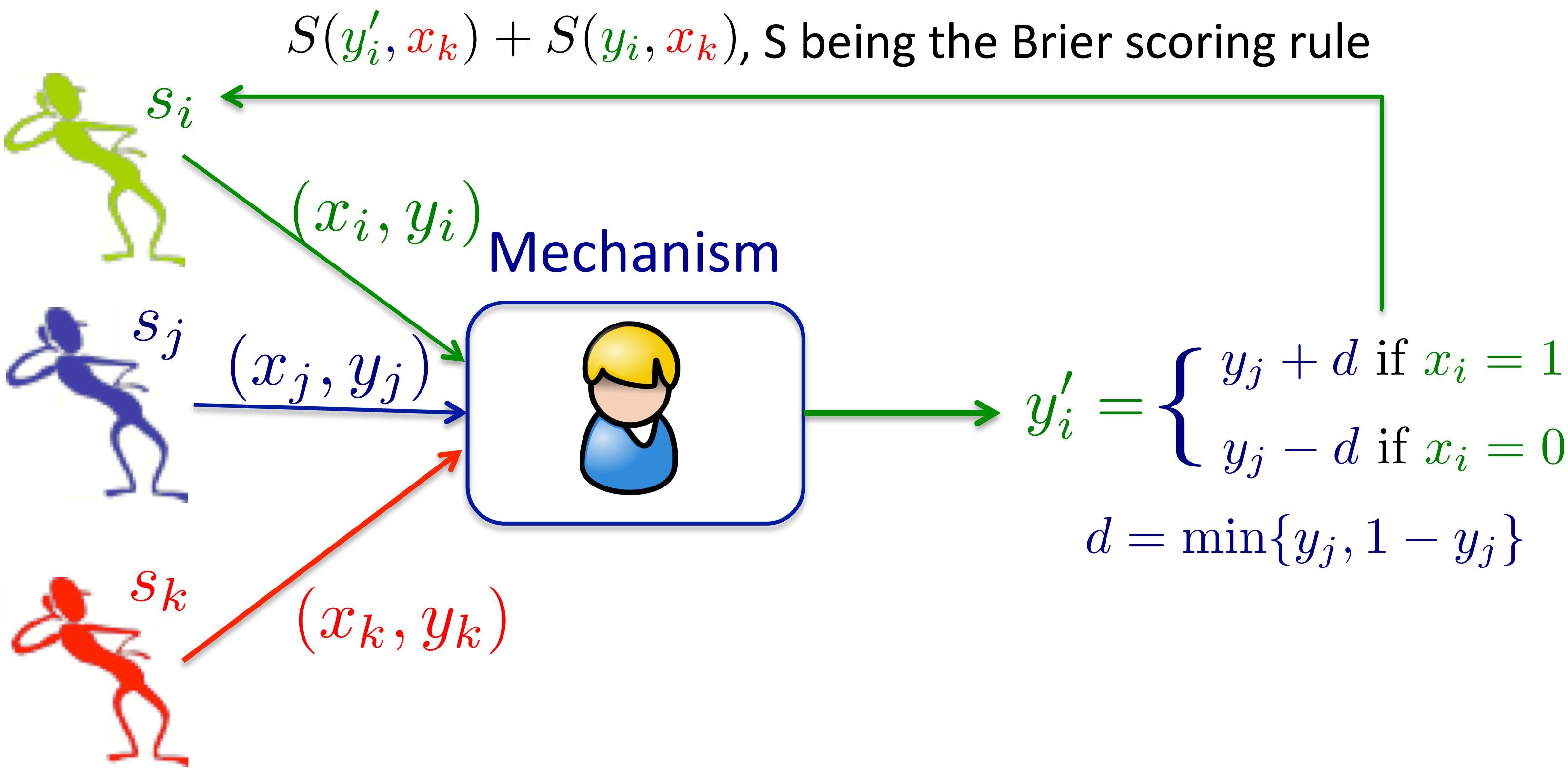
Robust Bayesian Truth Serum (RBTS)

[Witkowski and Parkes 2012a]

- Works for small populations
- RBTS assumes binary signals $\mathcal{S} = \{0, 1\}$
- Each agent i is asked for two reports
  - Information report: reported signal $x_i \in \{0, 1\}$
  - Prediction report: $y_i \in [0, 1]$ is the prediction of the frequency of signal 1
- Each agent has two reference agents

Robust Bayesian Truth Serum (RBTS)

[Witkowski and Parkes 2012a]



Truthful reporting is a strict Bayesian Nash equilibrium for $n \geq 3$.

Shadowing

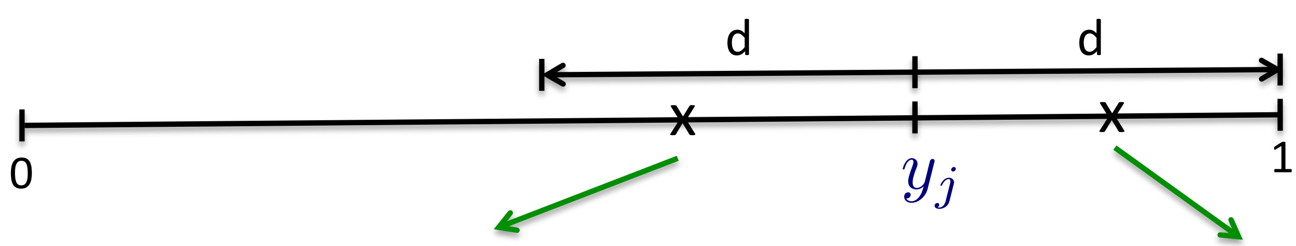
- S is the Brier scoring rule

Information Score $S(y_i', x_k) + S(y_i, x_k)$ Prediction Score

$$y_i' = \begin{cases} y_j + d \text{ if } x_i = 1 \\ y_j - d \text{ if } x_i = 0 \end{cases}$$

Shadowing

$$d = \min\{y_j, 1 - y_j\}$$



Prediction of agent i given signal 0 and y_j

Prediction of agent i given signal 1 and y_j

Other Related Work

- Many improvements on the original peer prediction [Jurca and Faltings 2006, 2007, 2008, 2011]
- Private prior peer prediction [Witkowski and Parkes 2012b]
- RBTS for non-binary signals [Radanovic and Faltings, 2013]

Questions?

yiling@seas.harvard.edu
http://yiling.seas.harvard.edu

References

- J. Abernethy, Y. Chen, and J. W. Vaughan, 2013. Efficient Market Making via Convex Optimization, and a Connection to Online Learning. ACM Transactions on Economics and Computation. Vol. 1, no. 2, pp. 12:1-12:39.
- J. D. Abernethy and R. M. Frongillo. 2012. A Characterization of Scoring Rules for Linear Properties. COLT'12.
- D. F. Bacon, Y. Chen, I. A. Kash, D. C. Parkes, M. Rao, and M. Sridharan. 2012. Predicting Your Own Effort. AAMAS'12.
- C. Boutilier. 2012. Eliciting forecasts from self-interested experts: scoring rules for decision makers. AAMAS'12.
- G. W. Brier. 1950. Verification of forecasts expressed in terms of probability. Monthly Weather Review 78(1): 1-3.
- Y. Chen, S. Dimitrov, R. Sami, D. M. Reeves, D. M. Pennock, R. D. Hanson, L. Fortnow, and R. Gonen. 2010. Gaming prediction markets: equilibrium strategies with a market maker. Algorithmica, 58: 930–969.
- Y. Chen and D. M. Pennock, 2007. A Utility Framework for Bounded-Loss Market Makers. UAI'07.
- Y. Chen and I. A. Kash. 2011. Information elicitation for decision making. AAMAS'11.

References (Cont.)

- X. A. Gao, J. Zhang, Y. Chen. 2013. What You Jointly Know Determines How You Act - Strategic Interactions in Prediction Markets, EC'13.
- T. Gneiting and A. E. Raftery. 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102, 477, 359–378.
- R. D. Hanson. 2003. Combinatorial information market design. Information Systems Frontiers 5(1), 107–119.
- R. D. Hanson. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. Journal of Prediction Markets, 1(1), 1–15.
- R. Jurca and B. Faltings, 2006. Minimum Payments that Reward Honest Reputation Feedback. EC'06.
- R. Jurca and B. Faltings, 2007. Collusion Resistant, Incentive Compatible Feedback Payments. EC'07.
- R. Jurca and B. Faltings.Incentives For Expressing Opinions in Online Polls, 2008. EC'08.
- R. Jurca and B. Faltings 2011. Incentives for Answering Hypothetical Questions. Workshop on Social Computing and User Generated Content at EC'11.
- N. S. Lambert, D. M. Pennock, and Y. Shoham. 2008. Eliciting properties of probability distributions. EC '08.
- J. E. Matheson and R. L. Winkler. 1976. Scoring rules for continuous probability distributions. Management Science 22, 1087–1096.

References (Cont.)

- J. McCarthy. 1956. Measures of the value of information. PNAS: Proceedings of the National Academy of Sciences of the United States of America, 42(9): 654-655.
- N. Miller, P. Resnick, and R. Zeckhauser 2005, Management Science, 51(9): 1359-1373.
- A. Othman and T. Sandholm. 2010. Decision rules and decision markets. AAMAS'10.
- M. Ostrovsky. 2012. Information aggregation in dynamic markets with strategic traders. Econometrica, 80(6): 2595-2648.
- D. Prelec, 2004. A Bayesian Truth Serum for Subjective Data. Science, 306 (5695): 462-466.
- G. Radanovic and B. Faltings, 2013. A Robust Bayesian Truth Serum for Non-Binary Signals. AAAI'13.
- L. J. Savage. 1971. Elicitation of personal probabilities and expectations. Journal of the American Statistical Association, 66(336): 783-801.
- P. Shi, V. Conitzer, and M. Guo. 2009. Prediction mechanisms that do not incentivize undesirable actions. WINE'09.
- J. Witkowski and D. C. Parkes, 2012a. A Robust Bayesian Truth Serum for Small Populations. AAAI'12.
- J. Witkowski and D. C. Parkes, 2012b. Peer Prediction without a Common Prior. EC'12.