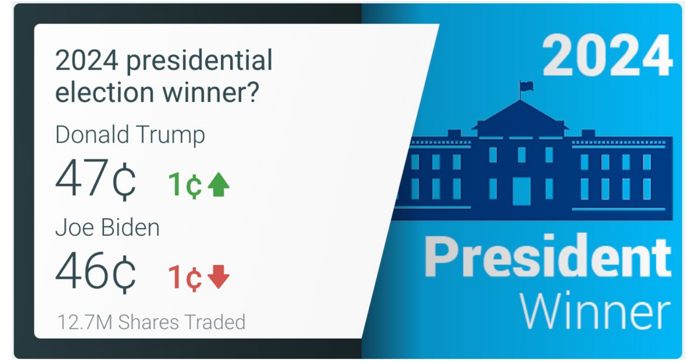
CS 598: AI Methods for Market Design

Lecture 9: Information Elicitation

Xintong Wang

Spring 2024

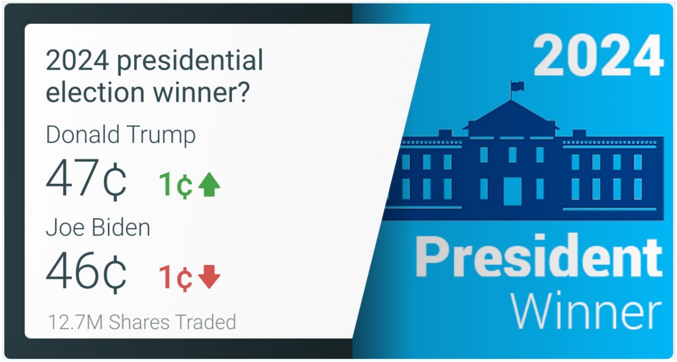
Overview



- Eliciting beliefs about something verifiable in the future
  - E.g., Will Trump be the 2024 presidential election winner?

- Eliciting information without (easy) verification
  - E.g., Does a plumber do high quality work?
    Is a restaurant good for friend gatherings?

Overview

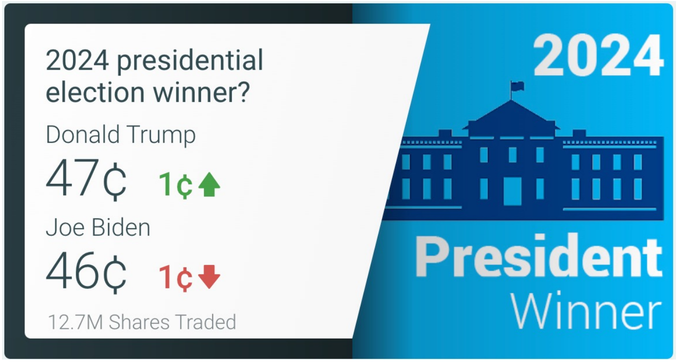


- Eliciting beliefs about something verifiable in the future
  - E.g., Will Trump be the 2024 presidential election winner?

- Eliciting information without (easy) verification
  - E.g., Does a plumber do high quality work?
    Is a restaurant good for friend gatherings?

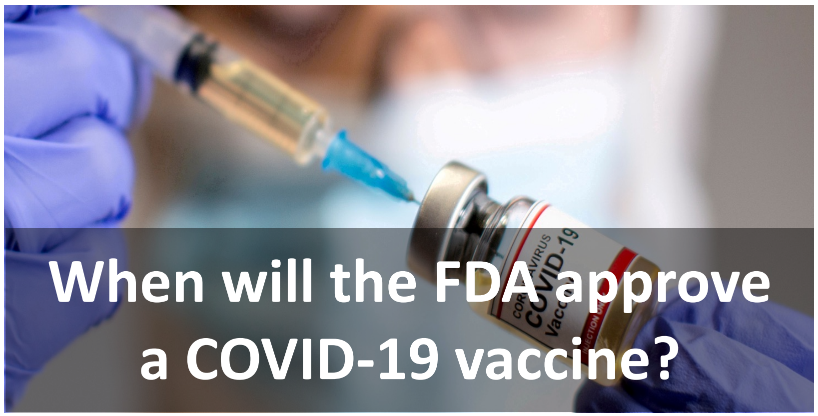
How to crowdsource information to make reliable predictions?

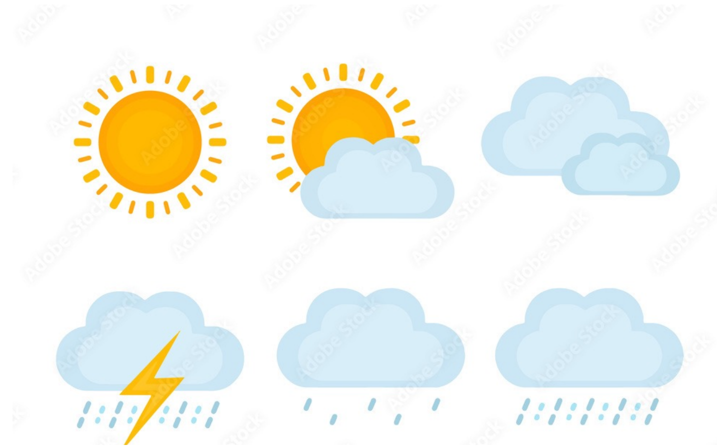
Overview



- Eliciting beliefs about something verifiable in the future
  - E.g., Will Trump be the 2024 presidential election winner?
  - Scoring rules & prediction markets
- Eliciting information without (easy) verification
  - E.g., Does a plumber do high quality work?
    Is a restaurant good for friend gatherings?
  - Peer prediction

Market as a Forecasting Tool





Goal: Produce a forecast based on information dispersed among agents from all sources

Outline

- Scoring Rules

- Peer Prediction
    - Output agreement
    - 1/Prior
    - Scoring-rule based

Outline

- Scoring Rules

- Peer Prediction
  - Output agreement
  - 1/Prior
  - Scoring-rule based

How to Pay a Forecaster

- Possible outcomes $O = \{o_0, \ldots, o_{m-1}\}$, indexed by k
- An agent's *true belief* p
  - E.g., I believe it will rain tomorrow with probability 0.5
- An agent's *belief report* q

How to Pay a Forecaster

- Possible outcomes $O = \{o_0, \ldots, o_{m-1}\}$, indexed by k
- An agent's *true belief* p
  - E.g., I believe it will rain tomorrow with probability 0.5
- An agent's *belief report* q
- A scoring rule pays $s(q, o_k)$ if the outcome is o_k
  - The payment is contingent on the outcome

How to Pay a Forecaster

- Possible outcomes $O = \{o_0, \dots, o_{m-1}\}$, indexed by k
- An agent's *true belief* p
  - E.g., I believe it will rain tomorrow with probability 0.5
- An agent's *belief report* q
- A scoring rule pays $s(q, o_k)$ if the outcome is o_k
  - The payment is contingent on the outcome
- Expected payment

$$E_{o\sim p}[s(q, o)] = \sum_k p_k \cdot s(q, o_k)$$

Example: Linear Scoring Rule

- The weather for tomorrow is a random variable W
- The outcome space is {sun, rain}
- True belief p = Pr(W=rain)
- Reported belief q
- Linear scoring rule: $s_{linear}(q, o_k) = q_k$
  - If it rains, then pay q; if it is sunny, then pay 1-q
- What is the expected payment?

Example: Linear Scoring Rule

- The weather for tomorrow is a random variable W
- The outcome space is {sun, rain}
- True belief p = Pr(W=rain)
- Reported belief q
- Linear scoring rule: $s_{linear}(q, o_k) = q_k$
  - If it rains, then pay q; if it is sunny, then pay 1-q
- What is the expected payment?

p*q + (1-p)*(1-q)

  - Suppose p=0.6. What is the best report?

Example: Linear Scoring Rule

- The weather for tomorrow is a random variable W
- The outcome space is {sun, rain}
- True belief p = Pr(W=rain)
- Reported belief q
- Linear scoring rule: $s_{linear}(q, o_k) = q_k$
  - If it rains, then pay q; if it is sunny, then pay 1-q
- What is the expected payment?

p*q + (1-p)*(1-q)

  - Suppose p=0.6. What is the best report? q=1
  - Based on p, an agent will only report $q \in \{0, 1\}$

Strictly Proper Scoring Rule

- A scoring rule is strictly proper if, for every belief p, the expected payment

$$E_{o\sim p}[s(q,o)] = \sum_k p_k \cdot s(q, o_k)$$

is *uniquely maximized* through truthful report (q=p)

Example: Logarithmic Scoring Rule

- Logarithmic scoring rule

$$s_{log}(q, o_k) = \ln(q_k)$$

- Expected payment under weather forecasting

$$p * \ln(q) + (1-p) * \ln(1-q)$$

- Verify optimality
  - First-order: p/q+1/(q-1)-p/(q-1) = 0 → q=p
  - Second-order derivative is negative
- Logarithmic scoring rule is strictly proper
- Any potential problem?

Example: Quadratic Scoring Rule

- Quadratic scoring rule

$$s_{quad}(q, o_k) = 2q_k - \sum_{k'} q_{k'}^2$$

- Expected payment under weather forecasting

p*(2q-(q^2+(1-q^2)))+ (1-p)*(2(1-q)- (q^2+(1-q^2)))

- Verify quadratic scoring rule is strictly proper
- Any potential problem?

Some Comments

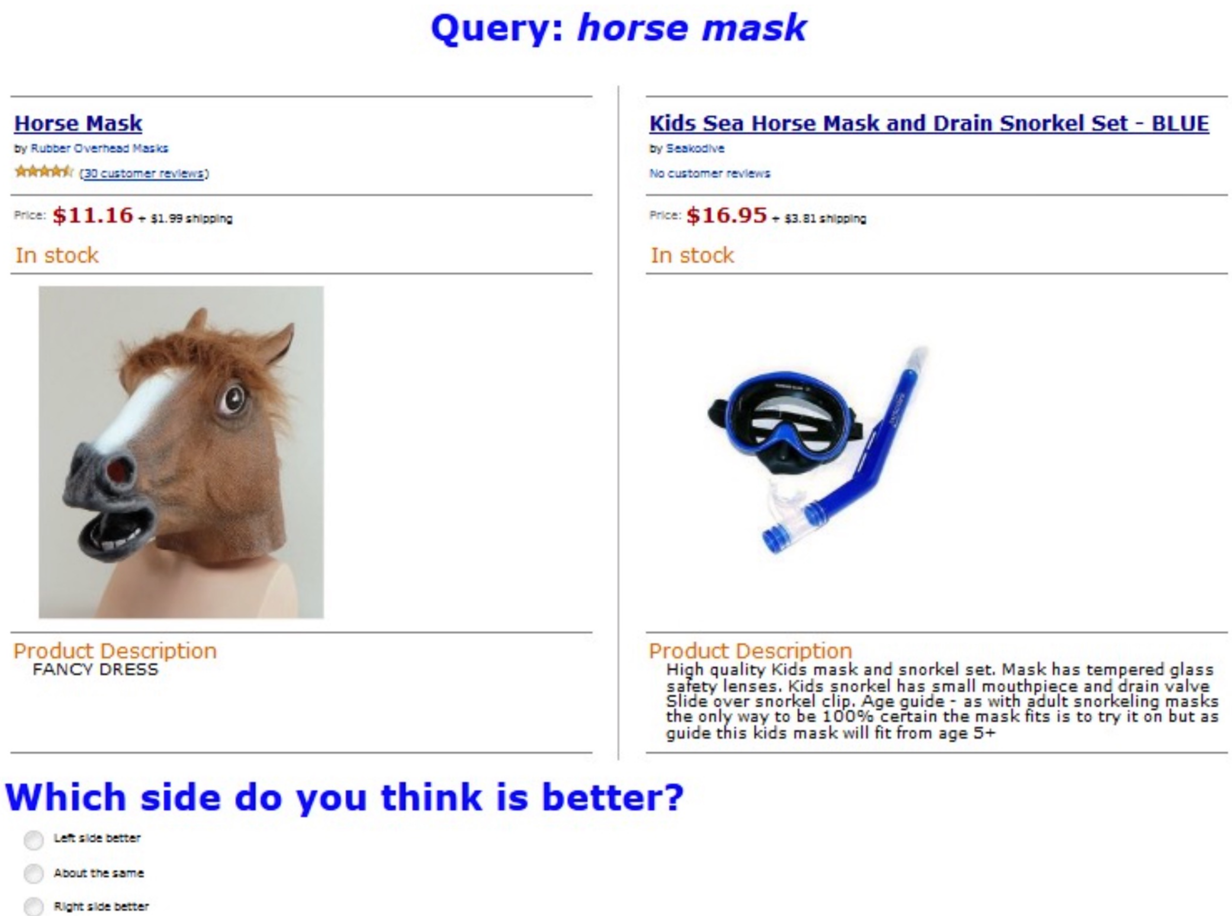
- Scoring rule s'remains strictly proper (for $\beta > 0$) if

$$s'(q, o_k) = \alpha_k + \beta \cdot s(q, o_k)$$

- Simplicity of scoring rules: “local” vs. not “local”
  - Local: $s_{log}(q, o_k) = \ln(q_k)$
  - Not local: $s_{quad}(q, o_k) = 2q_k - \sum_{k'} q_{k'}^2$

How about without Verification?

- Example: which of the two product search results are better?



How about without Verification?

- What grade {*A*, .., *E*} is appropriate for an essay?
- Is a restaurant suitable for large groups?
- Which startup is more likely to succeed?

The only inputs are reports from agents.

No verifications.

Outline

- Scoring Rules

- Peer Prediction
  - Output agreement
  - 1/Prior
  - Scoring-rule based

Peer Prediction Mechanisms

Peer 1 (Signal X_1)
E.g., *A* for the essay

Peer 2 (Signal X_2)
E.g., *B* for the essay

Report r_1

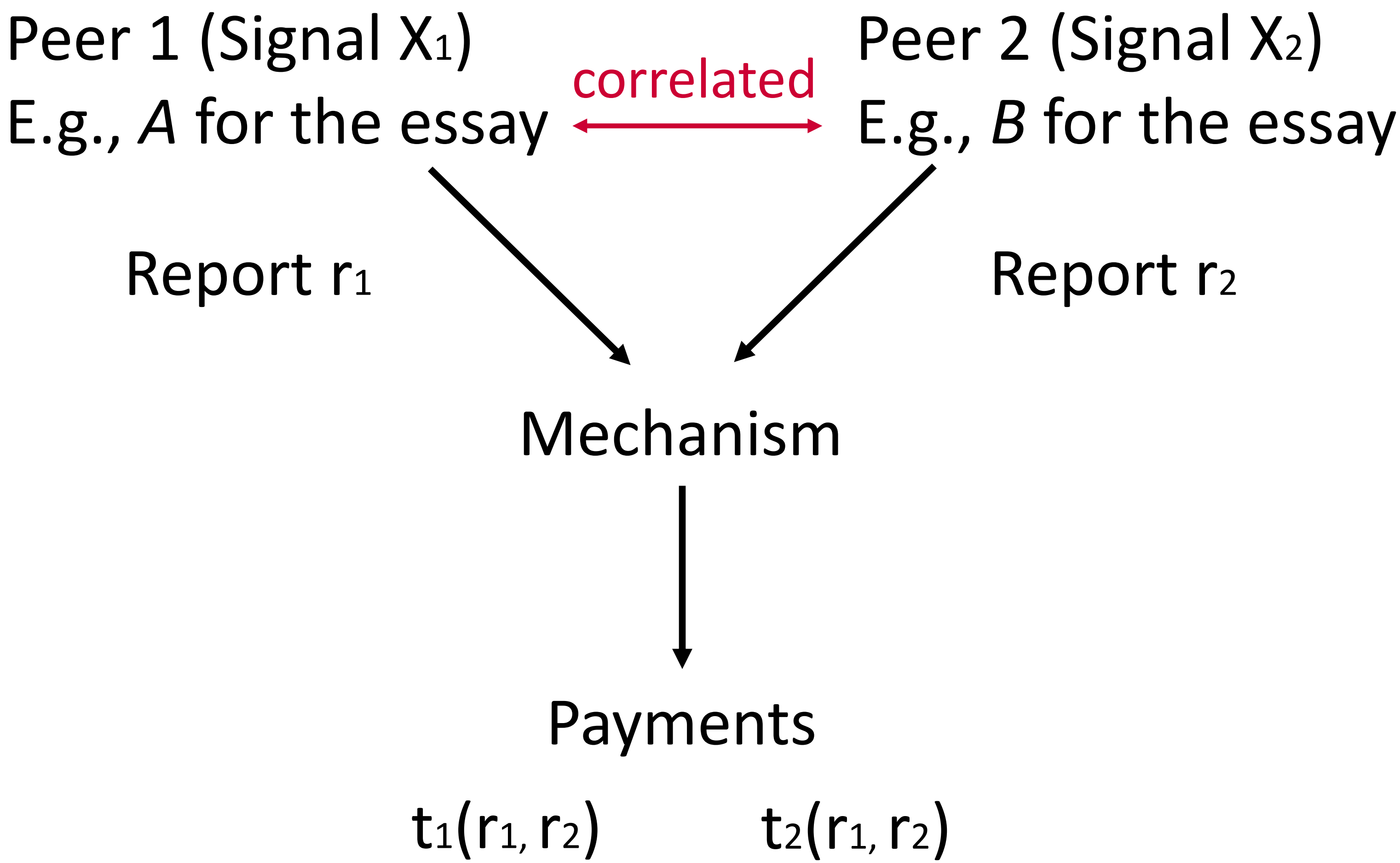
Report r_2

Mechanism

Payments

$t_1(r_1, r_2)$ $t_2(r_1, r_2)$

Peer Prediction Mechanisms



Peer Prediction Mechanisms

Example signal distribution

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

- Symmetric:
  - $P(X_1=1, X_2=0) = P(X_1=0, X_2=1)$
  - Agents are exchangeable (identity does not matter)
  - The marginal probability P(x) of signal x does not dependent on the agent identity, i.e., $P(X_1=1) = P(X_2=1)$

Peer Prediction Mechanisms

- Mechanism: a simultaneous-move game
- Strategy: a mapping from its signal to its report
- Payoffs: the payment rule

Peer Prediction Mechanisms

- Mechanism: a simultaneous-move game
- Strategy: a mapping from its signal to its report
- Payoffs: the payment rule

Goal: Incentivizing truthful reports

Outline

- Scoring Rules

- Peer Prediction
  - Output agreement
  - 1/Prior
  - Scoring-rule based

First Attempt: Output Agreement

Payment rule:

r_1 \ r_2	*0*	*1*
0	*1*	*0*
1	*0*	*1*

Signal distribution (example):

X_1 \ X_2	0	1
0	0.4	0.1
1	0.1	0.4

- Pay each agent \$1 if reports agree, \$0 otherwise
- *Is truthful reporting an equilibrium?*

First Attempt: Output Agreement

Payment rule:

$r_1 \backslash r_2$	*0*	*1*
0	*1*	*0*
1	*0*	*1*

Signal distribution (example):

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

Suppose Agent 2 is truthful. Given X_1=0

- Agent 1's expected payment for reporting 0:

$$0*Pr(X_2=1|X_1=0) + 1*Pr(X_2=0|X_1=0) = 0.8$$

- Agent 1's expected payment for (mis)reporting 1:

$$1*Pr(X_2=1|X_1=0) + 0*Pr(X_2=0|X_1=0) = 0.2$$

First Attempt: Output Agreement

Payment rule:

$r_1 \backslash r_2$	0	1
0	1	0
1	0	1

Signal distribution (example):

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

Suppose Agent 2 is truthful. Given $X_1=0$

- Agent 1's expected payment for reporting 0:

$$0*\Pr(X_2=1|X_1=0) + 1*\Pr(X_2=0|X_1=0) = 0.8$$

- Agent 1's expected payment for (mis)reporting 1:

$$1*\Pr(X_2=1|X_1=0) + 0*\Pr(X_2=0|X_1=0) = 0.2$$

Truthful!

Strictly Proper Peer Prediction

- A peer prediction mechanism with payment rule (t_1, t_2) is strictly proper if truthful reporting is a strict correlated equilibrium:

$$\mathbf{E}_{X_2 \sim P(X_2|X_1=j)}[t_1(j, X_2)] > \mathbf{E}_{X_2 \sim P(X_2|X_1=j)}[t_1(j', X_2)]$$

for all signals *j* of Agent 1, all misreports j' (with roles of 1 and 2 switched)

First Attempt: Output Agreement

Payment rule:

$r_1 \backslash r_2$	*0*	*1*
0	*1*	*0*
1	*0*	*1*

Signal distribution (example):

$X_1 \backslash X_2$	0	1
0	0.5	0.2
1	0.2	0.1

How about this case? *Is truthful reporting an equilibrium?*

First Attempt: Output Agreement

Payment rule:

$r_1 \setminus r_2$	0	1
0	1	0
1	0	1

Signal distribution (example):

$X_1 \setminus X_2$	0	1
0	0.5	0.2
1	0.2	0.1

Suppose Agent 2 is truthful. Given $X_1=1$

- Agent 1's expected payment for (mis)reporting 0:

  $$1*Pr(X_2=0|X_1=1) + 0*Pr(X_2=1|X_1=1) = 2/3$$

- Agent 1's expected payment for reporting 1:

  $$1*Pr(X_2=1|X_1=1) + 0*Pr(X_2=0|X_1=1) = 1/3$$

First Attempt: Output Agreement

Payment rule:

$r_1 \backslash r_2$	*0*	*1*
0	*1*	*0*
1	*0*	*1*

Signal distribution (example):

$X_1 \backslash X_2$	0	1
0	0.5	0.2
1	0.2	0.1

Suppose Agent 2 is truthful. Given X_1=1

- Agent 1's expected payment for (mis)reporting 0:

$$1*Pr(X_2=0|X_1=1) + 0*Pr(X_2=1|X_1=1) = 2/3$$

- Agent 1's expected payment for reporting 1:

Not truthful!

$$1*Pr(X_2=1|X_1=1) + 0*Pr(X_2=0|X_1=1) = 1/3$$

Property for OA to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "strongly-diagonalization", i.e., diagonal entries larger than other entries

$$P(X_2 = j \mid X_1 = j) > P(X_2 = j' \mid X_1 = j)$$

Meaning if agent 1 has signal j, then it is more likely that agent 2 has signal j than any other signal

Property for OA to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "strongly-diagonalization", i.e., diagonal entries larger than other entries

$$P(X_2 = j \mid X_1 = j) > P(X_2 = j' \mid X_1 = j)$$

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

YES!

$X_1 \backslash X_2$	0	1
0	0.5	0.2
1	0.2	0.1

NO!

Property for OA to be Strictly Proper

Need "strongly-diagonalization"

- Proof: Suppose Agent 2 is truthful. Given $X_1=j$

$$
\begin{aligned}
& \mathbf{E}_{X_2 \sim P(X_2|X_1=j)}[t_1(j, X_2)] > \mathbf{E}_{X_2 \sim P(X_2|X_1=j)}[t_1(j', X_2)] \\
\Leftrightarrow \quad & \sum_{\ell \in [m]} P(X_2 = \ell \mid X_1 = j) \cdot t_1(j, \ell) > \sum_{\ell \in [m]} P(X_2 = \ell \mid X_1 = j) \cdot t_1(j', \ell) \\
\Leftrightarrow \quad & P(X_2 = j \mid X_1 = j) \cdot 1 + \sum_{\ell \in [m],\ \ell \neq j} P(X_2 = \ell \mid X_1 = j) \cdot 0 > \\
& \qquad P(X_2 = j' \mid X_1 = j) \cdot 1 + \sum_{\ell \in [m],\ \ell \neq j'} P(X_2 = \ell \mid X_1 = j) \cdot 0 \\
\Leftrightarrow \quad & P(X_2 = j \mid X_1 = j) > P(X_2 = j' \mid X_1 = j),
\end{aligned}
$$

First Attempt: Output Agreement

Payment rule:

$r_1 \backslash r_2$	*0*	*1*
0	*1*	*0*
1	*0*	*1*

Signal distribution (example):

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

- Pay each agent $1 if reports agree, $0 otherwise
- *Any other potential problem?*

First Attempt: Output Agreement

Payment rule:

$r_1 \backslash r_2$	0	1
0	1	0
1	0	1

Signal distribution (example):

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

- Pay each agent $1 if reports agree, $0 otherwise
- *Any other potential problem?*
- (0, 0) and (1, 1) are NE. Uninformative with payoff dominates truthful reporting!

Outline

- Scoring Rules
- Peer Prediction
  - Output agreement
  - 1/Prior
  - Scoring-rule based

1/Prior Mechanism

- Use knowledge of marginal probabilities

r_1 \ r_2	0	1
0	$\frac{1}{P(S_i=0)}$	0
1	0	$\frac{1}{P(S_i=1)}$

- Provide a higher payment for agreement on signals that are a priori less likely

1/Prior Mechanism

- Use knowledge of marginal probabilities

$X_1 \backslash X_2$	0	1
0	0.5	0.2
1	0.2	0.1

signal distribution

$r_1 \backslash r_2$	0	1
0	$(1/0.7, 1/0.7)$	$(0, 0)$
1	$(0, 0)$	$(1/0.3, 1/0.3)$

payoff matrix

- Provide a higher payment for agreement on signals that are a priori less likely

Property for 1/Prior to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "self predicting", i.e., the conditional probability of Agent 2 having signal *j* is maximized by Agent 1 having signal *j:*

$$P(X_2 = j | X_1 = j) > P(X_2 = j | X_1 = j')$$

Property for 1/Prior to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "self predicting", i.e., the conditional probability of Agent 2 having signal *j* is maximized by Agent 1 having signal *j:*

$$P(X_2 = j | X_1 = j) > P(X_2 = j | X_1 = j')$$

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

YES!

$X_1 \backslash X_2$	0	1
0	0.5	0.2
1	0.2	0.1

YES!

Property for 1/Prior to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "self predicting", i.e., the conditional probability of Agent 2 having signal *j* is maximized by Agent 1 having signal *j:*

$$P(X_2 = j | X_1 = j) > P(X_2 = j | X_1 = j')$$

$X_1 \backslash X_2$	0	1
0	0.4	0.1
1	0.1	0.4

YES!

$X_1 \backslash X_2$	0	1
0	0.5	0.21
1	0.21	0.08

NO!

Property for 1/Prior to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "self predicting", i.e., the conditional probability of Agent 2 having signal *j* is maximized by Agent 1 having signal *j:*

$$P(X_2 = j | X_1 = j) > P(X_2 = j | X_1 = j')$$

- *HW:* verify that the 1/Prior peer prediction mechanism is strictly proper if and only if the signal distribution is self predicting

Outline

- Scoring Rules

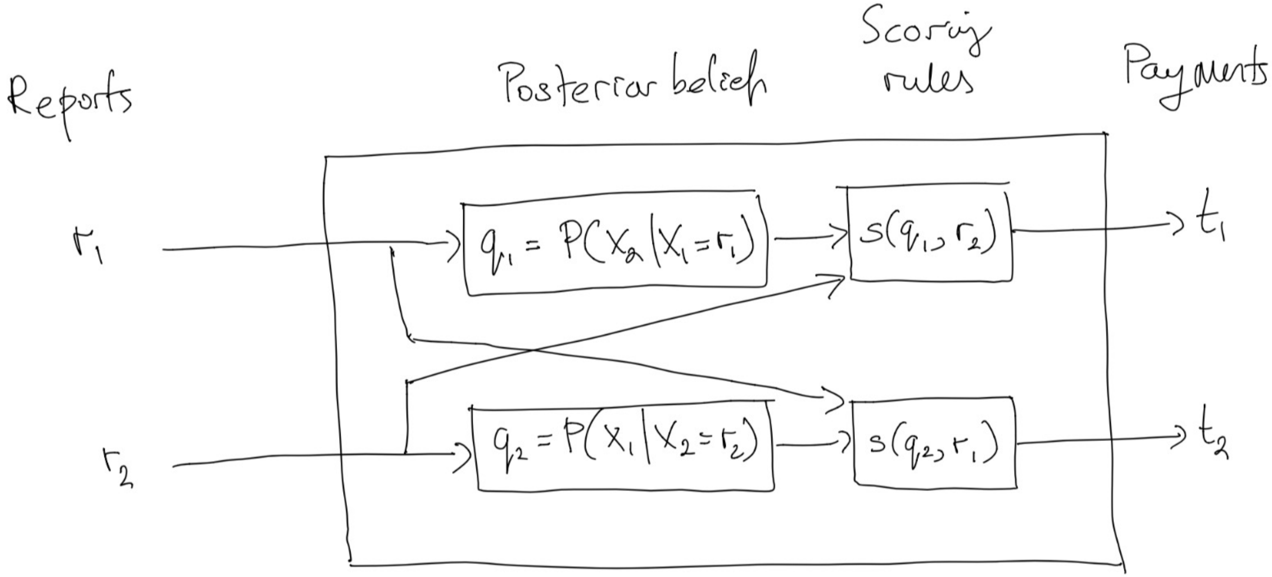
- Peer Prediction
  - Output agreement
  - 1/Prior
  - Scoring-rule based

Scoring-Rule Based Mechanisms

- Use knowledge of the joint distribution
- Based on report r_1, compute *posterior*

$$q=P(X_2|X_1=r_1)$$

- Score posterior "reported belief" q against "outcome" r_2, using **a strictly proper scoring rule**



Scoring-Rule Based Mechanisms

- Use knowledge of the joint distribution
- Based on report r_1, compute *posterior*

$$q=P(X_2|X_1=r_1)$$

- Score posterior "reported belief" q against "outcome" r_2, using a **strictly proper scoring rule**
- Example: using logarithmic scoring rule

$$s_{log}(q, r_k) = \ln(q_k)$$

$X_1 \backslash X_2$	0	1
0	0.5	0.2
1	0.2	0.1

signal distribution

agent \ peer	$r_2 = 0$	$r_2 = 1$
$r_1 = 0$	$\ln(5/7)$	$\ln(2/7)$
$r_1 = 1$	$\ln(2/3)$	$\ln(1/3)$

payoff matrix

Property to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "stochastic relevance"*:*

$$P(X_2 = j | X_1 = j) \neq P(X_2 = j | X_1 = j')$$

Meaning the signal of one agent always carries some information about the signal of the other

- This is a much weaker condition

Property to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "stochastic relevance":

$$P(X_2 = j | X_1 = j) \neq P(X_2 = j | X_1 = j')$$

- Proof:

Outcome: peer's report

$$\mathbf{E}_{\ell \sim P(\ell \mid j)}[t_1(j, \ell)] > \mathbf{E}_{\ell \sim P(\ell \mid j)}[t_1(j', \ell)]$$

$$\Leftrightarrow \quad \mathbf{E}_{\ell \sim P(\ell \mid j)}[s(q, \ell)] > \mathbf{E}_{\ell \sim P(\ell \mid j)}[s(q', \ell)]$$

P(X2 | X1 = j) P(X2 | X1 = j')

  - Inequality holds as a misreport leads to a different signal-conditioned belief (**by stochastic relevance**)
  - Therefore, a lower expected payment than truthful reporting j (**by strict properness of scoring rule**)

Property to be Strictly Proper

What condition should the signal distribution satisfy?

- Need "stochastic relevance"*:*

$$P(X_2 = j | X_1 = j) \neq P(X_2 = j | X_1 = j')$$

- Proof:

Outcome: peer's report

$$\mathbf{E}_{\ell \sim P(\ell \mid j)}[t_1(j, \ell)] > \mathbf{E}_{\ell \sim P(\ell \mid j)}[t_1(j', \ell)]$$

$$\Leftrightarrow \quad \mathbf{E}_{\ell \sim P(\ell \mid j)}[s(q, \ell)] > \mathbf{E}_{\ell \sim P(\ell \mid j)}[s(q', \ell)]$$

P(X2 | X1 = j) P(X2 | X1 = j')

Scoring-rule based mechanisms → peer "prediction":

An agent's expected payment is higher when its signal leads to a more accurate belief about the signal reported by the peer!

Summary

- Scoring rules promote truthful belief elicitation when there is a verifiable, future outcome to score against (next lecture: prediction market!)
- Peer prediction promotes truthful elicitation of information without (easy) verification

Peer-prediction mechanism	Knowledge required on the part of the design	Domain property required for strict-properness
OA	nothing	strong diagonalization
1/prior	signal prior	self predicting
scoring-rule based	signal conditionals	stochastic relevance

Announcements

- Two paper presentations this week, one next week
  - Peer evaluation for paper presentation
- Next week for class project feedback. Sign up a slot to discuss your (group) project
- Midterm survey summary

Midterm Survey

- Less theory more applied material and examples
- Techniques / core intuition behind proofs
- Pre-class CQs on Chapters as guidelines for reading
- Zoom office hours

Midterm Survey

More about pre-class readings:

- Familiarize yourself with the topic
- Get to know the major contributions of a research paper
- No need to delve into proofs or experiment details
- Refresh yourself on mathematical tools used in the Chapter / paper