Incentivizing Truthful Forecasting with Proper Scoring Rules

Aaron Roth

University of Pennsylvania

April 17 2025

Overview

- We've spent a lot of time thinking about auctions: how to allocate goods and extract money.

Overview

- We've spent a lot of time thinking about auctions: how to allocate goods and extract money.
- But what about information?

Overview

- We've spent a lot of time thinking about auctions: how to allocate goods and extract money.
- But what about information?
- This class: How to contract with an expert to incentivize them to report their belief to us about the likelihood of an event we will only observe once.

Example: Betting on the Election

- Suppose we want to know the likelihood that candidate A wins the next presidential election between A and B.

Example: Betting on the Election

- Suppose we want to know the likelihood that candidate A wins the next presidential election between A and B.
- But we don't follow politics and don't have informed beliefs.

Example: Betting on the Election

- Suppose we want to know the likelihood that candidate A wins the next presidential election between A and B.
- But we don't follow politics and don't have informed beliefs.
- Our friend the professional gambler is also a politics wonk. He's got well informed beliefs, but he won't just tell you, he'll only gamble.

Example: Betting on the Election

- Suppose we want to know the likelihood that candidate A wins the next presidential election between A and B.
- But we don't follow politics and don't have informed beliefs.
- Our friend the professional gambler is also a politics wonk. He's got well informed beliefs, but he won't just tell you, he'll only gamble.
- How can we set up a gamble so that if he wants to maximize his payoff he'll tell us his true beliefs?

Example: Betting on the Election

- Attempt 1: "Who do you think will win the election? I'll give you $1 if you get it right."

Example: Betting on the Election

- Attempt 1: "Who do you think will win the election? I'll give you \$1 if you get it right."
  - What will he say to maximize his profits?

Example: Betting on the Election

- Attempt 1: "Who do you think will win the election? I'll give you \$1 if you get it right."
  - What will he say to maximize his profits?
  - If he thinks $\Pr[A] \geq 1/2$ he'll guess A, otherwise he'll guess B.

Example: Betting on the Election

- Attempt 1: "Who do you think will win the election? I'll give you \$1 if you get it right."
  - What will he say to maximize his profits?
  - If he thinks $\Pr[A] \geq 1/2$ he'll guess A, otherwise he'll guess B.
  - But this doesn't tell you his specific belief about $\Pr[A]$ — i.e. can't distinguish $\Pr[A] = 0.51$ from $\Pr[A] = 0.99$.

Example: Betting on the Election

- Attempt 1: "Who do you think will win the election? I'll give you $1 if you get it right."
  - What will he say to maximize his profits?
  - If he thinks $\Pr[A] \geq 1/2$ he'll guess A, otherwise he'll guess B.
  - But this doesn't tell you his specific belief about $\Pr[A]$ — i.e. can't distinguish $\Pr[A] = 0.51$ from $\Pr[A] = 0.99$.
- But we didn't ask the right question...

Example: Betting on the Election

- Attempt 2: "What do you think is the probability p that A will win the election? I'll pay you...

Example: Betting on the Election

- Attempt 2: "What do you think is the probability p that A will win the election? I'll pay you...
- p if A wins and $1 - p$ if B wins."

Example: Betting on the Election

- Attempt 2: "What do you think is the probability p that A will win the election? I'll pay you...
- p if A wins and $1 - p$ if B wins."
  - What will he say to maximize his profits?

Example: Betting on the Election

- Attempt 2: "What do you think is the probability p that A will win the election? I'll pay you...
- p if A wins and $1-p$ if B wins."
  - What will he say to maximize his profits?
  - If he believes that A will win with probability q, then if he reports p his expected profit is:

$$S(p,q) = q \cdot p + (1-q) \cdot (1-p)$$

Example: Betting on the Election

- Attempt 2: "What do you think is the probability p that A will win the election? I'll pay you...
- p if A wins and $1-p$ if B wins."
  - What will he say to maximize his profits?
  - If he believes that A will win with probability q, then if he reports p his expected profit is:

$$S(p,q) = q \cdot p + (1-q) \cdot (1-p)$$

  - If $q \geq 0.5$ this is maximized at $p = 1$. Otherwise it is maximized at $p = 0$...

Example: Betting on the Election

- Attempt 2: "What do you think is the probability p that A will win the election? I'll pay you...
- p if A wins and $1 - p$ if B wins."
  - What will he say to maximize his profits?
  - If he believes that A will win with probability q, then if he reports p his expected profit is:

$$S(p, q) = q \cdot p + (1 - q) \cdot (1 - p)$$

  - If $q \geq 0.5$ this is maximized at $p = 1$. Otherwise it is maximized at $p = 0$...
- So we didn't learn any more than in Attempt 1...

Proper Scoring Rules

1. There is some future event Y that can take value in some finite set $\mathcal{Y}$: for example, $\mathcal{Y} = \{A, B\}$.

Proper Scoring Rules

1. There is some future event Y that can take value in some finite set $\mathcal{Y}$: for example, $\mathcal{Y} = \{A, B\}$.
2. An Agent has a belief $q \in \Delta\mathcal{Y}$ about how the outcome is distributed.

Proper Scoring Rules

1. There is some future event Y that can take value in some finite set $\mathcal{Y}$: for example, $\mathcal{Y} = \{A, B\}$.
2. An Agent has a belief $q \in \Delta\mathcal{Y}$ about how the outcome is distributed.
3. The Agent will report (a possibly different) distribution $p \in \Delta\mathcal{Y}$.

Proper Scoring Rules

1. There is some future event Y that can take value in some finite set $\mathcal{Y}$: for example, $\mathcal{Y} = \{A, B\}$.
2. An Agent has a belief $q \in \Delta\mathcal{Y}$ about how the outcome is distributed.
3. The Agent will report (a possibly different) distribution $p \in \Delta\mathcal{Y}$.
4. Once the outcome $\mathcal{Y} = y$ is realized, the Agent is paid $S(p, y)$, according to a known function (or *scoring rule*) $S : \Delta\mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$.

Proper Scoring Rules

1. There is some future event Y that can take value in some finite set $\mathcal{Y}$: for example, $\mathcal{Y} = \{A, B\}$.
2. An Agent has a belief $q \in \Delta\mathcal{Y}$ about how the outcome is distributed.
3. The Agent will report (a possibly different) distribution $p \in \Delta\mathcal{Y}$.
4. Once the outcome $\mathcal{Y} = y$ is realized, the Agent is paid $S(p, y)$, according to a known function (or *scoring rule*) $S : \Delta\mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$.
5. The Agent will report the distribution p that maximizes their expected payment under their beliefs:

$$p \in \arg\max_{p \in \Delta\mathcal{Y}} \mathrm{E}_{y \sim q}[S(p, y)]$$

Proper Scoring Rules

1. For shorthand, we'll write:

$$S(p; q) = \mathrm{E}_{y \sim q}[S(p, y)] = \sum_{y \in \mathcal{Y}} q(y) S(p, y)$$

   for the Agent's expected payoff of reporting p under belief q.

Proper Scoring Rules

1. For shorthand, we'll write:

$$S(p; q) = \mathrm{E}_{y \sim q}[S(p, y)] = \sum_{y \in \mathcal{Y}} q(y) S(p, y)$$

   for the Agent's expected payoff of reporting p under belief q.

Definition (Proper Scoring Rule)

A scoring rule $S : \Delta\mathcal{Y} \times \mathcal{Y}$ is proper if for every belief q, truthful reporting is a dominant strategy: for every $q, p \in \Delta\mathcal{Y}$:

$$S(q; q) \geq S(p; q)$$

If the inequality is strict for every $p \neq q$, we say that S is a *strictly proper* scoring rule.

An Aside: Convexity

Definition (Convex Set)

A set $C \subseteq \mathbb{R}^d$ is *convex* if it contains the line segment connecting any two points $x, y \in C$. In other words, if for any $x, y \in C$ and any $\alpha \in [0, 1]$:

$$\alpha x + (1 - \alpha) y \in C$$

An Aside: Convexity

Definition (Convex Set)

A set $C \subseteq \mathbb{R}^d$ is *convex* if it contains the line segment connecting any two points $x, y \in C$. In other words, if for any $x, y \in C$ and any $\alpha \in [0, 1]$:

$$\alpha x + (1 - \alpha) y \in C$$

Definition

A function $f : \mathbb{R}^d \to \mathbb{R}$ is convex if $C = \{x : x \geq f(x)\}$ is a convex set. Equivalently, for all $x, y \in \mathbb{R}^d$, and for all $\alpha \in [0, 1]$:

$$f(\alpha x + (1 - \alpha) y) \leq \alpha f(x) + (1 - \alpha) f(y)$$

An Aside: Convexity

An equivalent characterization: a function is convex if and only if every line tangent to the function lies below the function.

An Aside: Convexity

An equivalent characterization: a function is convex if and only if every line tangent to the function lies below the function.

Fact

A differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ is convex if and only if for every $x, y \in \mathbb{R}^d$:

$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y)$$

(See pictures)

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.
2. So $S(p; q) = q \cdot S(p, A) + (1 - q)S(p, B)$.

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.
2. So $S(p; q) = q \cdot S(p, A) + (1 - q)S(p, B)$.
3. Let $f(q) = S(q; q)$.

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.
2. So $S(p; q) = q \cdot S(p, A) + (1 - q)S(p, B)$.
3. Let $f(q) = S(q; q)$.
   3.1 Observation 1: $S(p; q)$ is *linear* in q for all p.

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.
2. So $S(p; q) = q \cdot S(p, A) + (1 - q)S(p, B)$.
3. Let $f(q) = S(q; q)$.
   3.1 Observation 1: $S(p; q)$ is *linear* in q for all p.
   3.2 If S is proper, then for all $q \neq p$, $f(q) \geq S(p; q)$.

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.
2. So $S(p; q) = q \cdot S(p, A) + (1 - q)S(p, B)$.
3. Let $f(q) = S(q; q)$.
   3.1 Observation 1: $S(p; q)$ is *linear* in q for all p.
   3.2 If S is proper, then for all $q \neq p$, $f(q) \geq S(p; q)$.
   3.3 So $f(q) = \max_{p \in [0,1]} S(p; q)$, the maximum of a bunch of linear functions (convex).

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.
2. So $S(p; q) = q \cdot S(p, A) + (1 - q)S(p, B)$.
3. Let $f(q) = S(q; q)$.
   3.1 Observation 1: $S(p; q)$ is *linear* in q for all p.
   3.2 If S is proper, then for all $q \neq p$, $f(q) \geq S(p; q)$.
   3.3 So $f(q) = \max_{p \in [0,1]} S(p; q)$, the maximum of a bunch of linear functions (convex).
   3.4 And for all $p \in [0, 1]$, $S(p; q)$ is the tangent line (gradient) of $f(q)$ at $p = q$, and lies entirely below $f(q)$.

Proper Scoring Rules: Building Intuition

1. Lets consider the binary prediction case: $\mathcal{Y} = \{A, B\}$. We can think of beliefs $p \in \mathbb{R}$, where $p = \Pr[A]$.
2. So $S(p; q) = q \cdot S(p, A) + (1 - q)S(p, B)$.
3. Let $f(q) = S(q; q)$.
   - 3.1 Observation 1: $S(p; q)$ is *linear* in q for all p.
   - 3.2 If S is proper, then for all $q \neq p$, $f(q) \geq S(p; q)$.
   - 3.3 So $f(q) = \max_{p \in [0,1]} S(p; q)$, the maximum of a bunch of linear functions (convex).
   - 3.4 And for all $p \in [0, 1]$, $S(p; q)$ is the tangent line (gradient) of $f(q)$ at $p = q$, and lies entirely below $f(q)$.
4. (See pictures).

Proper Scoring Rules: A Characterization

Theorem

Fix a finite domain $\mathcal{Y}$ with $|\mathcal{Y}| = d$. A scoring rule $S : \Delta\mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ is proper if and only if there exists a convex function $f : \mathbb{R}^d \to \mathbb{R}$ such that:

$$S(p; q) = f(p) + \nabla f(p)(q - p)$$

(In particular $S(p, y) = f(p) + \nabla f(p)(e_y - p)$ where e_y is the unit vector that has a 1 in the y'th component). The function f also satisfies

$$f(q) = S(q; q)$$

Proof

We have two directions to prove. First, if $f : \mathbb{R}^d \to [0, 1]$ is convex, then $S(p, y) = f(p) + \nabla f(p)(e_y - p)$ is proper.

Proof

We have two directions to prove. First, if $f : \mathbb{R}^d \to [0,1]$ is convex, then $S(p, y) = f(p) + \nabla f(p)(e_y - p)$ is proper.

1. We can compute for any p, q:

$$S(p; q) = \mathrm{E}_{y \sim q}[f(p) + \nabla f(p)(e_y - p)] = f(p) + \nabla f(p)(q - p)$$

Proof

We have two directions to prove. First, if $f : \mathbb{R}^d \to [0,1]$ is convex, then $S(p,y) = f(p) + \nabla f(p)(e_y - p)$ is proper.

1. We can compute for any p, q:

$$S(p;q) = \mathrm{E}_{y\sim q}[f(p) + \nabla f(p)(e_y - p)] = f(p) + \nabla f(p)(q - p)$$

2. If $q = p$ then we have:

$$S(q;q) = f(q)$$

Proof

We have two directions to prove. First, if $f : \mathbb{R}^d \to [0,1]$ is convex, then $S(p,y) = f(p) + \nabla f(p)(e_y - p)$ is proper.

1. We can compute for any p, q:

$$S(p;q) = \mathrm{E}_{y \sim q}[f(p) + \nabla f(p)(e_y - p)] = f(p) + \nabla f(p)(q - p)$$

2. If $q = p$ then we have:

$$S(q;q) = f(q)$$

3. So for $p \neq q$, we have $S(q;q) \geq S(p;q)$ exactly when:

$$f(q) \geq f(p) + \nabla f(q)(q - p)$$

Proof

We have two directions to prove. First, if $f : \mathbb{R}^d \to [0, 1]$ is convex, then $S(p, y) = f(p) + \nabla f(p)(e_y - p)$ is proper.

1. We can compute for any p, q:

$$S(p; q) = \mathrm{E}_{y \sim q}[f(p) + \nabla f(p)(e_y - p)] = f(p) + \nabla f(p)(q - p)$$

2. If $q = p$ then we have:

$$S(q; q) = f(q)$$

3. So for $p \neq q$, we have $S(q; q) \geq S(p; q)$ exactly when:

$$f(q) \geq f(p) + \nabla f(q)(q - p)$$

4. Since f is convex, this is always the case! (Tada!)

Proof

In the reverse direction, we need to show that if S is proper, then there is a convex function f such that $S(p, y) = f(p) + \nabla f(p)(e_y - p)$

1. We'll let $f(q) = S(q; q)$

Proof

In the reverse direction, we need to show that if S is proper, then there is a convex function f such that $S(p, y) = f(p) + \nabla f(p)(e_y - p)$

1. We'll let $f(q) = S(q; q)$
2. Recall that for any p:

$$S(p; q) = \sum_{y \in \mathcal{Y}} q(y) S(p, y)$$

   which is linear in q, always lies below $f(q)$, and is tangent to f at $q = p$.

Proof

In the reverse direction, we need to show that if S is proper, then there is a convex function f such that $S(p, y) = f(p) + \nabla f(p)(e_y - p)$

1. We'll let $f(q) = S(q; q)$
2. Recall that for any p:

$$S(p; q) = \sum_{y \in \mathcal{Y}} q(y) S(p, y)$$

   which is linear in q, always lies below $f(q)$, and is tangent to f at $q = p$.
3. So for all p, q we can write:

$$S(p; q) = f(p) + \nabla f(p)(q - p)$$

Proof

In the reverse direction, we need to show that if S is proper, then there is a convex function f such that $S(p, y) = f(p) + \nabla f(p)(e_y - p)$

1. We'll let $f(q) = S(q; q)$
2. Recall that for any p:

   $$S(p; q) = \sum_{y \in \mathcal{Y}} q(y) S(p, y)$$

   which is linear in q, always lies below $f(q)$, and is tangent to f at $q = p$.
3. So for all p, q we can write:

   $$S(p; q) = f(p) + \nabla f(p)(q - p)$$

4. (Since all of f's tangent lines lie below it, it is convex)

Example

1. Let $S(p, y) = \log(p(y))$.

Example

1. Let $S(p, y) = \log(p(y))$.
2. So $S(p; q) = \sum_{y \in \mathcal{Y}} q(y) \log(p(y))$ (Cross entropy)

Example

1. Let $S(p, y) = \log(p(y))$.
2. So $S(p; q) = \sum_{y \in \mathcal{Y}} q(y) \log(p(y))$ (Cross entropy)
3. Lets check our characterization...

Example

1. Let $S(p, y) = \log(p(y))$.
2. So $S(p; q) = \sum_{y \in \mathcal{Y}} q(y) \log(p(y))$ (Cross entropy)
3. Lets check our characterization...
   3.1 $f(q) = S(q; q) = \sum_{y \in \mathcal{Y}} q(y) \log(q(y))$: Negative Shannon Entropy (Convex)

Example

1. Let $S(p, y) = \log(p(y))$.
2. So $S(p; q) = \sum_{y \in \mathcal{Y}} q(y) \log(p(y))$ (Cross entropy)
3. Lets check our characterization...
   3.1 $f(q) = S(q; q) = \sum_{y \in \mathcal{Y}} q(y) \log(q(y))$: Negative Shannon Entropy (Convex)
   3.2 We can recover $S(p, y)$ from our expression:

$$\begin{aligned}
S(p, y) &= f(p) + \nabla f(p)(e_y - p) \\
&= f(p) + \nabla f(p) e_y - \nabla f(p) p \\
&= \sum_{y \in \mathcal{Y}} p(y) \log(p(y)) + (1 + \log p(y)) - 1 - \sum_{y \in \mathcal{Y}} p(y) \log(p(y)) \\
&= \log p(y)
\end{aligned}$$

Closing Remarks: Proper Losses in Machine Learning

1. So cross entropy (a common objective in machine learning) is a proper scoring rule.

Closing Remarks: Proper Losses in Machine Learning

1. So cross entropy (a common objective in machine learning) is a proper scoring rule.
2. So is squared loss...

Closing Remarks: Proper Losses in Machine Learning

1. So cross entropy (a common objective in machine learning) is a proper scoring rule.
2. So is squared loss...
3. Not a coincidence! If you are solving a regression problem to try and learn the probability of a label conditional on some features, the unconstrained optimum will be the true distribution exactly when the loss is proper!

Closing Remarks: Proper Losses in Machine Learning

1. So cross entropy (a common objective in machine learning) is a proper scoring rule.
2. So is squared loss...
3. Not a coincidence! If you are solving a regression problem to try and learn the probability of a label conditional on some features, the unconstrained optimum will be the true distribution exactly when the loss is proper!
4. An important reason why regression models minimize *squared error* rather than e.g. *absolute error*.

Thanks!

See you next class — stay healthy!