Proper Scoring Rules

Yiling Chen

September 17, 2012

Logistics

- Let's introduce ourselves!
- Find a partner for paper presentation
- Email your paper preferences to Mike by 5pm tomorrow (Sep. 18)
- Paper presentations
  - Read the "Presentation Notes" on course website
  - Meet with us roughly one week before your presentation
  - Prepare reading questions
  - Present the paper(s) and lead class discussion

Probabilistic Forecasts

- A random variable with n mutually exclusive and exhaustive outcomes (e.g. Rain, Sun, Snow)

- A probabilistic forecast $\mathbf{p} = (p_1, p_2, \ldots, p_n)$

Calibration

- Ask an expert to predict the daily weather for a year
- If the expert predicts Rain with probability 0.8 for 100 days, what do you expect to observe for these 100 days (assuming the expert is a “good” forecaster)?
- Does good calibration mean accurate forecasts?

Scoring Rules

- A random variable with n mutually exclusive and exhaustive outcomes (e.g. Rain, Sun, Snow)
- A probabilistic forecast $\mathbf{p} = (p_1, p_2, \ldots, p_n)$
- A scoring rule rewards an expert

$$S(\mathbf{p}, \omega)$$

when his prediction is $\mathbf{p}$ and the realized outcome is ω

Scoring Rules

- Let's consider a linear scoring rule

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

- If a risk-neutral expert believes the probabilities for Rain, Sun, and Snow are

$$\mathbf{q} = (0.7, 0.2, 0.1),$$

what should the expert predict?

Strictly Proper Scoring Rules

- A scoring rule is strictly proper iif

$$\mathbf{q} = \arg\max_{\mathbf{p} \in \Delta_n} \sum_{\omega} q_\omega S(\mathbf{p}, \omega)$$

for all $\mathbf{q} \in \Delta_n$.

Examples of Strictly Proper Scoring Rules

- Quadratic scoring rule (Brier score):

$$S(\mathbf{p}, \omega) = -(p_\omega - 1)^2 - \sum_{\omega' \neq \omega} p_{\omega'}^2$$

- Logarithmic scoring rule:

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

- Affine transformation of a strictly proper scoring rule does not change the incentives

- If the expert does not have incentives to lie, would you still use a strictly proper scoring rule? Why or why not?

- What are possible applications?

- How many strictly proper scoring rules do we have?
- How to construct a strictly proper scoring rule?

Savage Characterization

- A proper scoring rule $S(\mathbf{p}, \omega)$ is strictly proper iif

$$S(\mathbf{p}, \omega) = G(\mathbf{p}) - \sum_{\omega'} G'_{\omega'}(\mathbf{p}) p_{\omega'} + G'_{\omega}(\mathbf{p})$$

where $G(\mathbf{p})$ is a strictly convex function, $G'(\mathbf{p})$ is a subgradient of G at $\mathbf{p}$, and $G'_{\omega}(\mathbf{p})$ is its ω-th component.

Savage Characterization

- Expected score function

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

- So far, we have discussed strictly proper scoring rules for eliciting a probability mass function

- What if we have a continuous random variable?

Predicting Continuous Random Variables

- Continuous random variable X
- Expert has a subjective CDF $F(x)$
- Expert reports a CDF $R(x)$
- Are there strictly proper scoring rules that incentivize the expert to report $R(x) = F(x)$?

Construct Strictly Proper Scoring Rules for CDF

- Take a strictly proper scoring rule for binary random variables
- Pick a random divider u that discretizes the continuous space
- Take the prediction $(R(u), 1 - R(u))$ and apply the scoring rule for binary random variables
- Get rid of the dependency on u

[Matheson & Winkler, 76]

- Can we elicit the mean of a random variable?

Predicting Mean of Random Variables

- Random variable X
- Expert has a subjective belief $E(X)$
- Expert reports a predicted mean r
- Are there strictly proper scoring rules that incentivize the expert to report $r = E(X)$?

Construct Strictly Proper Scoring Rules for Eliciting the Mean

- Consider a random variable over [0, 1]
- Take a strictly proper scoring rule for binary random variables
- Take $(r, 1-r)$ as a prediction
- Construct a scoring rule
- Savage'71 provides a characterization using convex functions