Discussion of "Inattentive Valuation and Reference Dependent Choice" by Michael Woodford

Alp Simsek

Harvard University

June 2012, Portugal

Ambitious paper with several goals

A departure from the "Chicago man":

1. Rational inattention (RI) with an **alternative cost function.**
   - Emphasis on better conformity with evidence from psychophysics.
   - Particular form of **reference dependence** (related to **priors**).
2. Apply theory to explain various behavioral anomalies.

This discussion:

1. The new cost function vis-a-vis the usual one.
2. Brief discussion of applications.

Recall the rational inattention framework

Rational inattention (Sims):

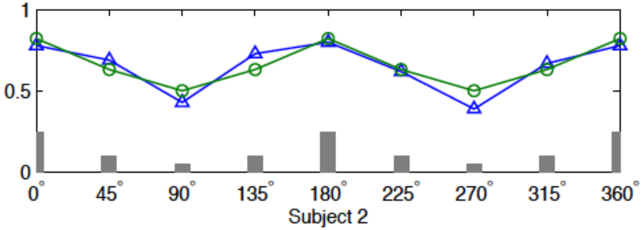
- State of the world, x, with prior $\pi(x)$.
- Representation r with conditional probabilities, $p(r \mid x)$.
- Capacity is measured by **mutual information (MI):**

$$I(r,x) = \underbrace{H(r) - H(r \mid x)}_{\text{entropy reduction}}.$$

- DM chooses r to maximize objective s.t. $I(r,x) \leq C$.

Next: Some examples to illustrate the shortcomings(?) of MI.

Example: The experiment of Shaw-Shaw



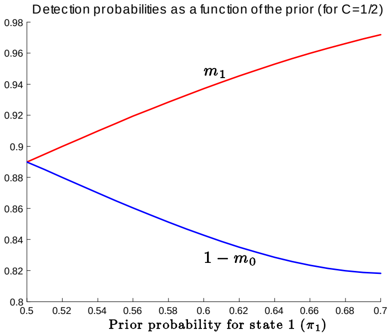
- **Data:** Location matters. Reference dependence.
- **MI criterion:** Location doesn't matter.
- But this example is a bit misleading...

Example: Guess the digit

- Suppose $x \in \{1, 0\}$. Prior $\pi_1 \in [0, 1]$.
- Consider representation, $\hat{x}$. Define $m_x = \Pr(\hat{x} = 1 \mid x)$.
- Choose $\{m_1, m_0\}$ to maximize $E[-\mathbf{1}[\hat{x} \neq x]]$ subject to **the MI constraint**, $I(\hat{x}, x) \leq C$.

What are the optimal detection probabilities m_1 and $1 - m_0$?

Example: Guess the digit



MI criterion also features reference dependence!

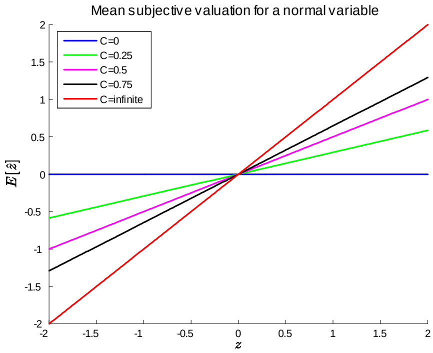
Example: Normal distribution

- Suppose x is Normal with prior $N(0,1)$.
- Choose $\hat{x}$ to maximize $E\left[-(x-\hat{x})^2\right]$ s.t. $I(x,\hat{x}) \leq C$.
- **Solution:** Observe signal $x+\varepsilon$ with $\frac{1}{\sigma_{\varepsilon^2}} = 2^{2C} - 1$.
- Optimal guess is:

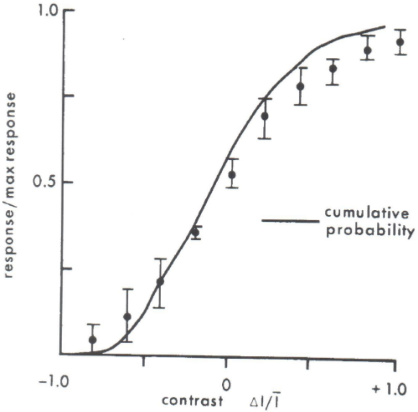
$$\hat{x} = \alpha(x+\varepsilon), \text{ where } \alpha = 1 - \frac{1}{2^{2C}}.$$

Reference dependence again!

Ref. dependence: More accuracy around the mean



So why do we need a different criterion?



- Laughlin: Response to brightness **more sensitive** around the mean.
- **Stronger form of reference dependence** than in MI...

This paper's criterion

- Alternative: **Maximized Mutual Information (MMI):**

$$\max_{\pi} I(r \mid x) \leq C.$$

- Capacity determined by **all possible priors**...

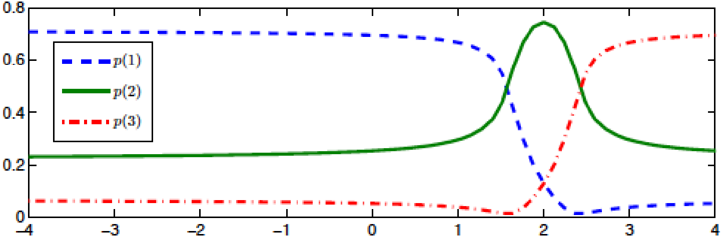
Concern (Motivation): Are we taking MI too seriously as a biological constraint?

Concern (Tractability): MI to the max.

Example: Normal distribution

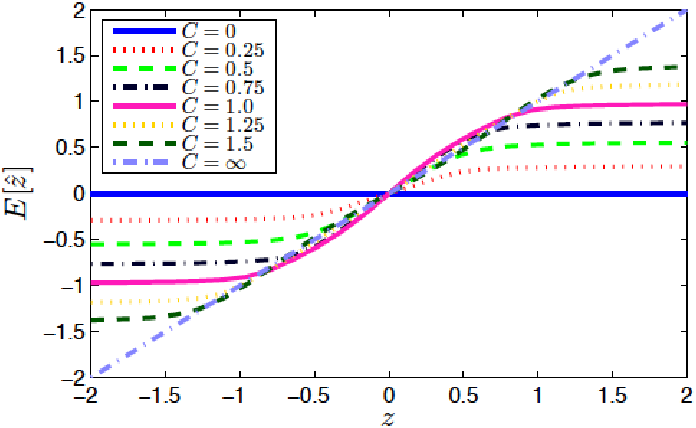
- Let's investigate the optimal signal when $x \sim N(\mu, 1)$.
- The earlier solution (observing $x + \varepsilon$) is not optimal!
- The optimal $\hat{x}$ takes a finite number of values...

Example: Normal distribution



- Generates **stronger form of reference dependence**.

Example: Normal distribution



Value added: Reference dependence in first differences.

Applications

- **Stochasticity:** Same with MI and MMI.
  - Smooth hazard function in sS problems (Woodford, 2009).
- **Focusing effects:** Same with MI and MMI.
  - Optimal capacity increasing function of σ^2.
  - Mackowiak and Wiederholt, "Optimal Sticky Prices..."
- **Context-decoy effects:** Similar with MI and MMI.

Reference effects/Prospect theory: Works with MMI (but not MI).

Kahneman-Tversky (1979)

Problem

In addition to whatever you own, you have been given 1000. You are now asked to choose between (a) winning an additional 500 with certainty, or (b) a gamble with a 50 percent chance of winning 1000 and a 50 percent chance of winning nothing.

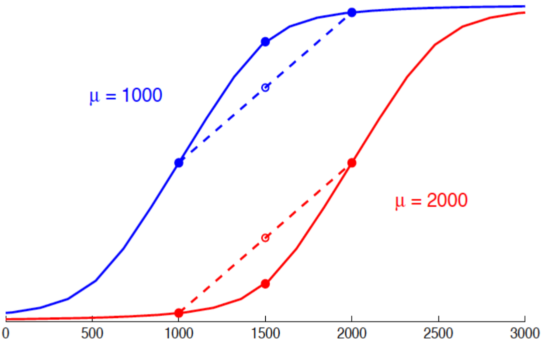
Majority of subjects choose (a)

Problem

In addition to whatever you own, you have been given 2000. You are now asked to choose between (a) losing 500 with certainty, and (b) a gamble with a 50 percent chance of losing 1000 and a 50 percent chance of losing nothing.

Majority of subjects choose (b)

Prospect theory: Explanation with MMI



- Reference dependence through priors (problem statement).
- Deeper explanation for the PT value function.

Comments about applications

Applications need a theory of priors:

- Currently chosen to fit the anomalies. Flexibility or weakness?
- How much are they influenced by **the choice set (or cues)**?
- How much by **history and experience**?

Criterion should be selectively applied:

- Motivated by **humans' biological constraints**.
- **Might be irrelevant for some economic agents** (e.g., firms).
- Observing $x + \varepsilon$ seems simple enough in some applications...

Excellent and thought provoking paper ($\hat{x} = high$).