Causality and Causal Misperception in Dynamic Games

Sungmin Park

The Ohio State University

November 24, 2024

Motivation

Limited observation of reality $\Rightarrow$ Varying perceptions of causality

- People have different perceptions about how actions affect outcomes

Smoking Education Police size Social media

- Subjects in lab experiments look at the same data and tell different causal narratives (Kendall and Charles, 2022)

- Yet, most applications of game theory continue to assume Rational Expectations (RE)

What I do

Question What is a useful solution concept to incorporate people's misperceptions about causality in extensive-form games?

Answer Let each player best respond to a belief about Nature and others' strategies consistent with observed outcomes

Even better + let each player's belief be the simplest explanation consistent with observation

"Maximum-entropy Observation-consistent Equilibrium" (MOE)

What I do

Question What is a useful solution concept to incorporate people's misperceptions about causality in extensive-form games?

Answer Let each player best respond to a belief about Nature and others' strategies consistent with observed outcomes

Even better + let each player's belief be the simplest explanation consistent with observation

"Maximum-entropy Observation-consistent Equilibrium" (MOE)

What I do

Question What is a useful solution concept to incorporate people's misperceptions about causality in extensive-form games?

Answer Let each player best respond to a belief about Nature and others' strategies consistent with observed outcomes

Even better + let each player's belief be the simplest explanation consistent with observation

"Maximum-entropy Observation-consistent Equilibrium" (MOE)

What I do

Question What is a useful solution concept to incorporate people's misperceptions about causality in extensive-form games?

Answer Let each player best respond to a belief about Nature and others' strategies consistent with observed outcomes

Even better + let each player's belief be the simplest explanation consistent with observation

"Maximum-entropy Observation-consistent Equilibrium" (MOE)

Main Results

Does it Exist?

Every finite extensive-form game with perfect recall and observational constraint has an MOE

Is it Useful?

MOE captures common causal misperceptions such as

- Correlation neglect
- Omitted-variable bias (selection neglect)
- Simultaneity bias (reverse causality bias)

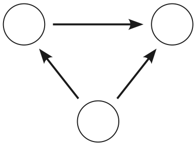
Is it Compatible with RE?

If agents have perfect observation of outcomes,

- OE $\Leftrightarrow$ Self-confirming equilibrium
- MOE $\Leftrightarrow$ Perfect Bayesian Equilibrium (PBE)
- (with infinite horizons) MOE $\Leftrightarrow$ Markov Perfect Equilibrium (MPE)

Literature

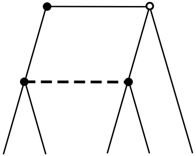
Bridging behavioral theory and standard game theory



Behavioral theory

(e.g. Spiegler, 2020, 2021)

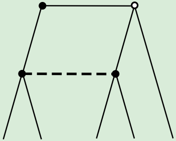
- Single-person decisions
- Directed Acyclic Graphs
- Maximum entropy
- Subjective best responses



Standard game theory

(e.g. Kreps and Wilson, 1982)

- Multiple players
- Observe terminal nodess
- Correct beliefs
- Objective best responses



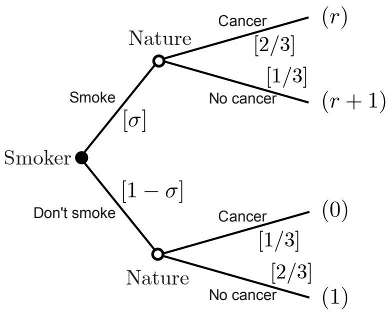
$$C = \begin{bmatrix} \quad \end{bmatrix}$$

My paper (MOE)

- Multiple players
- Observational structure (C)
- Maximum entropy
- Subjective best responses

Simplest example

- Smoker chooses to smoke ($s = 1$) or not ($s = 0$).
  - If he smokes, Nature gives him cancer with prob $\pi_1 = 2/3$.
  - If not, Nature gives him cancer with prob $\pi_0 = 1/3$.
- He gets $r < \frac{1}{3}$ if he smokes and loses 1 if he gets cancer.
- Smoker's **strategy** is the prob $\sigma \in [0, 1]$ of smoking.
- Smoker's **belief** is $\beta = (\beta_0, \beta_1)$ where β_s is the subjective probability of getting cancer given s.



Smoker's Problem

$\Rightarrow$ Under rational expectations, one shouldn't smoke because the causal effect of smoking on cancer ($\frac{2}{3} - \frac{1}{3} = \frac{1}{3}$) is larger than the reward r

Observational consistency

"Observational structure" Smoker observes only the marginal prob of cancer.

Definition

Given strategy $\sigma \in [0,1]$, a belief $\beta \in [0,1]^2$ is **observation-consistent** if

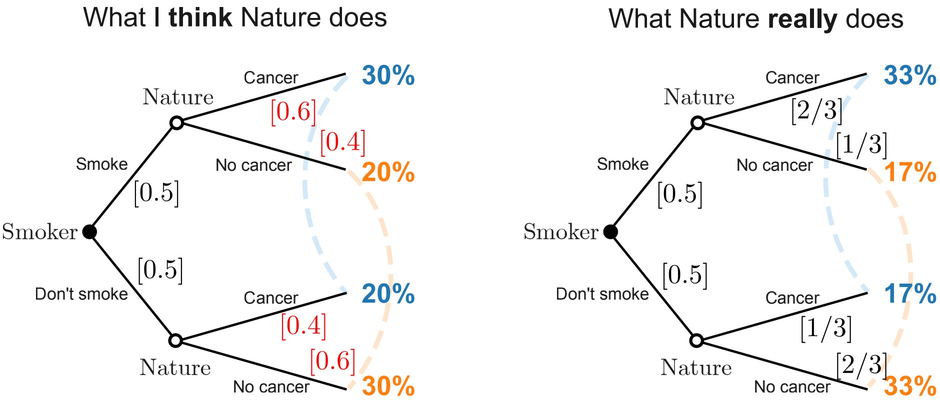
$$\underbrace{\sigma\beta_1 + (1-\sigma)\beta_0}_{\text{perceived marginal prob of cancer}} = \underbrace{\sigma \cdot \frac{2}{3} + (1-\sigma) \cdot \frac{1}{3}}_{\text{actual marginal prob of cancer}}$$

Interpretation Smoker sees a population of smokers choosing σ and sees the overall rate of cancer patients, but do not know the conditional probabilities.

Problem There are many observation-consistent beliefs.

Illustration of an observational consistency

Suppose I smoke half of the time ($\sigma = 0.5$).



Principle of Maximum Entropy

Notation

- $\mathbf{p}(\sigma, \beta)$: vector of probabilities over the 4 terminal nodes.
- $G(\cdot)$: Shannon entropy function.

Definition

Given strategy $\sigma \in (0,1)$, an observation-consistent belief $\beta^* \in [0,1]^2$ **maximizes the entropy** if

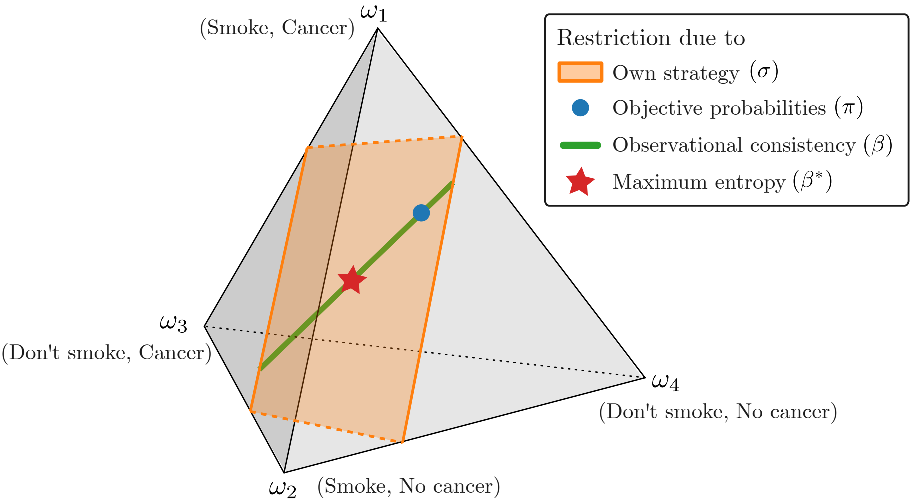
$$\beta^* \in \underset{\beta \text{ is observation-consistent}}{\operatorname{argmax}} G(\mathbf{p}(\sigma, \beta)).$$

Interpretation

- Among many worldviews consistent with observation, players believe in the one that assumes the least information

Illustration of maximum entropy

A point prediction on belief



Maximum entropy $\Rightarrow$ correlation neglect

Claim

For every $\sigma \in (0,1)$, the maximum-entropy belief β^* satisfies

$$\beta_0^* = \beta_1^* = (1-\sigma)\cdot\frac{1}{3} + \sigma\cdot\frac{2}{3}.$$

Meaning The smoker doesn't think smoking causes cancer

Intuition The smoker observes no evidence of dependence between smoking and cancer, so he believes in none.

General result (Shore and Johnson, 1980; Csiszar, 1991)
Correlation neglect $\Leftrightarrow$ maximum entropy, whenever agents observe only the marginal prob. distribution between two variables

Equilibrium

Definition

A strategy-belief pair (σ, β) is an **observation-consistent equilibrium (OE)** if

1. Given the belief β, the strategy σ is a best response, and
2. Given the strategy σ, the belief β is observation-consistent.

Interpretation

- OE is a prediction of how the smoker behaves, given his possibly wrong but observationally consistent belief

OE is too permissive

Every strategy is rationalizable by some observation-consistent belief

Claim

Every strategy σ has a belief β such that (σ, β) is an OE.

Note: Specifically, the OCE equilibria are

1. $\sigma = 0$, $\beta_0 = \frac{1}{3}$, and $\beta_1 - \beta_0 \geq r$,
2. $\sigma = 1$, $\beta_1 = \frac{2}{3}$, and $\beta_1 - \beta_0 \leq r$, and
3. $\sigma \in (0, 1)$, $\beta_0 = \sigma \cdot (\frac{2}{3} - r) + (1 - \sigma) \cdot \frac{1}{3}$, and $\beta_1 = \sigma \cdot \frac{2}{3} + (1 - \sigma)(\frac{1}{3} + r)$.

Idea Because there are many observation-consistent beliefs, there are many OEs.

Definition of MOE

Definition

An OE (σ, β) is a **maximum-entropy observation-consistent equilibrium (MOE)** if β maximizes the entropy given $\sigma \in (0, 1)$.

* For $\sigma \notin (0, 1)$, an OE is an MOE if some $\{(\sigma^k, \beta^k)\}_{k=1}^{\infty} \to (\sigma, \beta)$ and each β^k maximizes the entropy given σ^k

Interpretation

- MOE is an OE with the extra requirement that the smoker believes in the simplest explanation consistent with observation

MOE gives a sharper prediction

Claim

A strategy-belief pair (σ, β) is an MOE if and only if

$$\sigma = 1 \quad \text{and} \quad \beta_0 = \beta_1 = \frac{2}{3}.$$

Meaning

- Smoker keeps smoking while thinking that smoking doesn't cause cancer

Intuition

- Maximum-entropy belief features correlation neglect, so no other strategy is a best response.

General framework

Model (Γ, C) where

- Γ: a finite extensive-form game with perfect recall, and
- C: **observational structure**, a linear map from outcomes $(\Delta(\Omega))$ to observable outcomes $(\mathbb{R}^\ell)$

Observational consistency Given a strategy σ_i, a belief β_i is observation-consistent if

$$C\mathbf{p}(\sigma_i, \beta_i) = C\mathbf{p}(\sigma_i, (\sigma_{-i}, \pi)).$$

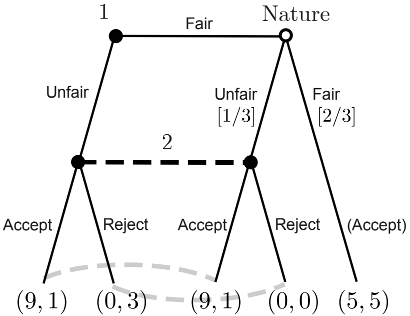
Equilibrium (MOE) A profile of strategies, beliefs, and posterior functions such that

- everyone's strategy is (subjectively) sequentially rational,
- everyone's belief is max-ent observational-consistent, and
- everyone's posterior function is Bayes-consistent

Example: An ultimatum-game-like scenario

Manager-Worker game

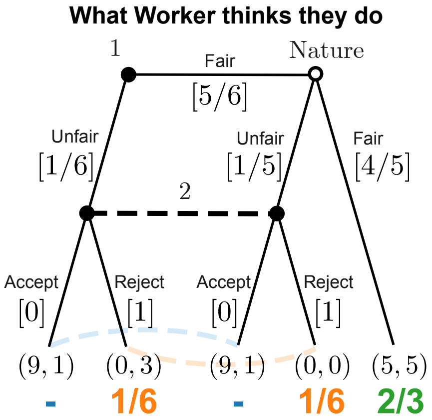
- Manager (Player 1) decides a fair or unfair bonus to Worker (Player 2)
- Even if Manager chooses a fair bonus, Nature might change it to unfair or keep it fair
- If Worker receives fair bonus, he accepts. If not, he either accepts or rejects.
  - He gets a thrill for rejecting an unfair Manager
- Worker doesn't know how likely Manager treats him unfairly in the interim or ex post (in a population)

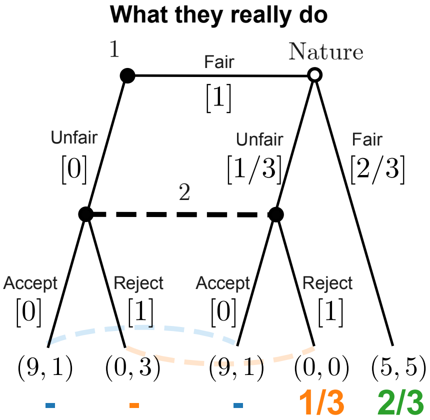


$$C = \begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix}$$

Unique MOE

Manager always tries to be fair





Lesson Limited observation of outcomes lead to more extreme (pure) strategies

Takeaway

Consider using MOE if you want to

- allow causal misperception in a dynamic model,
- let misperceptions arise endogenously from the observational structure, and
- get narrow predictions

Takeaway

Consider using MOE if you want to

- allow causal misperception in a dynamic model,
- let misperceptions arise endogenously from the observational structure, and
- get narrow predictions

Thank you!

Appendix

References I

Csiszar, Imre (1991) "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *The Annals of Statistics*, 2032–2066.

Kendall, Chad W and Constantin Charles (2022) "Causal narratives,"Technical report.

Kreps, David M and Robert Wilson (1982) "Sequential equilibria," *Econometrica: Journal of the Econometric Society*, 863–894.

Shore, John and Rodney Johnson (1980) "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Transactions on Information Theory*, 26 (1), 26–37.

Spiegler, Ran (2020) "Behavioral implications of causal misperceptions," *Annual Review of Economics*, 12, 81–106.

——— (2021) "Modeling players with random "data access"," *Journal of Economic Theory*, 198, 105374.