Limit Points of Endogenous Misspecified Learning

Drew Fudenberg, Giacomo Lanzani, and Philipp Strack

ES-NAWM

January 2021

Motivation

- People often have incorrect views of the world despite abundant data.
- Examples:
  - Belief that taxes are linear in income when they are not;
  - Belief in the "law of small numbers" and the gambler's fallacy;
  - "Causation neglect" about the impact of actions on outcomes;
  - Ignoring informative signals in the belief that they don't matter.
- It is important to understand how such agents learn from data, and how they will behave.

Introduction

- We analyze learning from endogenous data when the agent is a misspecified Bayesian: Their prior assigns probability 0 to a neighborhood of the true map from actions to outcome distributions.

- We provide:

  - A new and sharper necessary condition for an action to be a limit point of the learning process.

  - A characterization of the actions that are limit points for all "nearby" beliefs.

  - Sufficient conditions for an action to have positive probability of being the limit outcome from any initial beliefs.

- A Berk-Nash equilibrium (Esponda and Pouzo, 2016) is an action a that is myopically optimal against *some* beliefs supported on the models that are closest (wrt the KL divergence) to the true data generating process given that a is played. (formal definitions later)

- We relate limit outcomes to two refinements of this concept.

- A uniform Berk-Nash equilibrium is a best reply to *any* mixture over KL minimizers.

- A uniformly strict Berk-Nash equilibrium is an action that is a *strict* best reply to *every* mixture over KL minimizers.

General Results

- Any limit point must be a uniform Berk-Nash equilibrium.

- Uniformly strict Berk-Nash equilibria are uniformly stable: behavior converges to them with high probability from all nearby beliefs.

- Conversely uniformly stable equilibria must be uniformly strict.

- Thus

  Uniformly Strict B-N $=$ Uniformly Stable $\subseteq$ Stable $\subseteq$ Uniform B-N.

Positive Attractiveness

- Equilibria are positively attractive if they have positive probability from any starting beliefs.
- We show that uniformly strict Berk-Nash equilibria are positively attractive under various types of misspecification:
  - Causation Neglect, where the agent mistakenly believes that their action does not affect the outcome distribution,
  - Subjective Bandits, where the agent thinks that the outcomes observed when playing one action are uninformative about the distribution induced by the others,
- In supermodular environments, extremal equilibria are positively attractive.
- Some of the results extend to the case in which the agent observes a signal before taking their action.

Most Closely Related Work

- Esponda and Pouzo (2016) introduces Berk-Nash equilibrium, and proves convergence to it when the payoffs of the agent face iid shocks.
- Esponda, Pouzo, and Yamamoto (2019) focuses on convergence of beliefs and of action *frequencies* as opposed to actions.
- Frick, Iijima, and Ishii (2020) studies convergence of beliefs without explicitly modelling actions. Assumes a finite-support prior, and proves convergence to Berk-Nash equilibrium under myopia. It also introduces a measure of distance between models that we use in a proof.
- Mention other related work at the end time permitting.

Actions, Utilities and Objective Outcome Distributions

- Every period $t \in \mathbb{N}$, the agent chooses an action a from the finite set A.
- Finite set of outcomes Y.
- Action a has two consequences:
  - Induces objective probability distribution over outcomes $p_a^* \in \Delta(Y)$;
  - Directly influences the agent's payoff through $u : A \times Y \to \mathbb{R}$.

Subjective Beliefs of the Agent

- Let $P := \times_{a \in A} \Delta(Y)$ be the space of all action-dependent outcome distributions.
- Elements $p \in P$, components p_a.
- The agent is Bayesian.
- They have a prior $\mu_0 \in \Delta(P)$.
- $\mathcal{P} := \text{supp}\mu_0$ is the set of conceivable outcome distributions.
- The agent may be misspecified, i.e. we allow $p^* \notin \mathcal{P}$.

Behavior of the Agent

- A (pure) policy $\pi : \bigcup_{t=0}^{\infty} A^t \times Y^t \to A$ specifies an action for every history $(a_\tau, y_\tau)_{\tau=0}^{t} = (a^t, y^t) \in A^t \times Y^t$.

- We assume that the agent wants to maximize expected discounted utility with discount factor $\beta \in [0, 1)$.

- $A^m(\mu) = \arg\max_{a \in A} \int_P \mathbb{E}_{p_a}[u(a, y)]\, d\mu(p)$ is the set of myopic best replies to belief μ.

Berk-Nash Equilibrium

- Given two distributions over outcomes $q, q' \in \Delta(Y)$ define

$$H\left(q, q'\right) = -\sum_{y \in Y} q(y) \log q'(y).$$

- For each action a, let

$$\hat{\mathcal{P}}\left(a\right) = \underset{p \in \mathcal{P}}{\operatorname{argmin}}\, H\left(p_a^*, p_a\right) = \underset{p \in \mathcal{P}}{\operatorname{argmin}}\, H\left(p_a^*, p_a\right) - H(p_a^*, p_a^*)$$

  denote the set of conceivable action-contingent outcome distributions that minimize the KL divergence relative to p_a^* when the agent plays a.

- Action a is a Berk-Nash equilibrium (Esponda and Pouzo [2016] if there is a belief $\nu \in \Delta(\hat{\mathcal{P}}\left(a\right))$ such that *a* is myopically optimal given ν.

- Two outcome distributions $p, p' \in \mathcal{P}$ are observationally equivalent under action a, written $p \sim_a p'$, if $p_a(y) = p'_a(y)$.

- Let $\mathcal{E}_a(p) \subseteq \mathcal{P}$ denote the outcome distributions in $\mathcal{P}$ that are observationally equivalent to p under a.

- We do not assume that agents are arbitrarily patient, so no reason to expect them to have have much data about the consequences of every action.

Refinements of Berk-Nash Equilibrium

Definition (Uniform and Uniformly Strict Berk-Nash Equilibria)

Action a is a

(i) uniform Berk-Nash equilibrium if for every KL minimizing outcome distribution $p \in \hat{\mathcal{P}}(a)$, there is a belief over the observationally equivalent distributions $\nu \in \Delta\left(\mathcal{E}_a(p)\right)$ such that $a \in A^m(\nu)$.

(ii) uniformly strict Berk-Nash equilibrium if $\{a\} = A^m(\nu)$ for every observationally equivalent belief in $\nu \in \Delta(\hat{\mathcal{P}}\,(a))$.

When the agent is correctly specified (i.e. $p^* \in \mathcal{P}$),

$$\text{Uniform B-N} = \text{B-N} = \text{Self-Confirming},$$

as p_a^* is the unique KL minimizer for a.

Technical Assumptions

- **Simplyfing assumption for the talk:** For all $p \in \mathcal{P}$, p and p^* are mutually absolutely continuous. This guarantees that no conceivable distribution is ruled out after a finite number of observations.
- Also assume that the prior μ_0 has subexponential decay: there is $\Phi : \mathbb{R}_+ \to \mathbb{R}$ such that for every $p \in \mathcal{P}$ and $\varepsilon > 0$ we have

$$\mu_0(B_\varepsilon(p)) \geq \Phi(\varepsilon)$$

  with

$$\lim \Phi(K/n) \exp(n) = \infty \qquad \forall K > 0.$$

- Priors with a density that is bounded away from 0 on their support, priors with finite support, and Dirichlet priors all have subexponential decay. Fudenberg, He, and Imhof [2017] show that Bayesian updating can behave oddly on priors w/o subexponential decay.

Only Uniform-Berk Nash Equilibria are Limit Actions

Theorem (Limit Actions are Uniform Berk-Nash Equilibria)

If actions converge to a $\in A$ with positive probability, a is a uniform Berk-Nash equilibrium.

- Previous results on convergence to B-N equilibria require myopia and either i.i.d. payoff shocks or a finite-support prior (Esponda and Pouzo, 2016, Frick, Iijima, and Ishii 2020).
- Sharper conclusion: a limit action must be a best reply to all of the KL minimizers it induces.
- Key for this and a few of our other results is a lemma that say the beliefs of misspecified agents converge to the K-L minimizers at a uniform rate.
- This extends the uniform concentration result of Diaconis and Friedman [1990] to misspecified agents.

Proof Sketch for Theorem 1

- Our uniform concentration result shows that the agent's belief concentrates around the distributions that minimize the KL divergence from the empirical frequency at an exponential rate e^{Kt} that is uniform over the sample realizations.

- While playing a, the empirical frequency converges to p_a^*

- The difference between the empirical frequency and p_a^* is a random walk, and it oscillates in the direction of the different minimizers.

- By the Central Limit Theorem these oscillations die out at rate $\frac{1}{\sqrt{t}}$, which is slower than the exponential contraction of beliefs.

Proof Sketch for Theorem 1

- So we can use an extension of the second Borel-Cantelli lemma for events that are not "too correlated" to show that infinitely often the beliefs concentrate around every minimizer.

- If a is not uniform B-N, this induces the agent to switch to another action.

When are there Multiple KL Miminimizers?

- In space of *all* probability distributions there is generically a unique KL minimizer.
- But frameworks with symmetry or parametric restrictions are not generic, and there multiple KL minimizers can arise naturally.
- Example: suppose that y is the color of the ball drawn from an urn which is known to contain 6 balls.
- The agent correctly believes their action doesn't affect y.
- Outcome distributions correspond to the urn composition.
- The agent is certain that at most half of the balls have the same color, i.e., that $p(y) \leq 1/2$ for every y.
- In reality the urn has 4 white balls, 1 red, and 1 blue.
- So the two KL minimizers are (3 white, 2 blue, 1 red) and (3 white, 1 blue, 2 red).

Possible Non-convergence

- Nyarko (1991) shows by example that misspecified learning may not converge.
- However, Esponda and Pouzo (2016) show there always exists a B-N equilibrium.
- Their existence proof relies on possibly nonuniform Berk-Nash equilibria featuring multiple minimizers.
- Our theorem 1 shows that if no equilibrium is uniform, actions cannot converge; this may be easier to check than directly verifying non-convergence.
- We show by example that uniform B-N equilibria need not exist.
- One case where they do exist is if the agent is correctly specified.

Two Stability Notions

Definition (Stability)

(i) A Berk-Nash equilibrium a is stable if for every $\kappa \in (0,1)$, there is an $\epsilon > 0$ and a belief $\nu \in \Delta(\mathcal{P})$ such that for all initial beliefs in $B_\epsilon(\nu)$, the action prescribed by *some* optimal policy converges to a with probability larger than $1-\kappa$.

(ii) A Berk-Nash equilibrium a is uniformly stable if for every $\kappa \in (0,1)$, there is an $\epsilon > 0$ such that for all initial beliefs $\nu \in \Delta(\mathcal{P})$ such that $\nu(\hat{\mathcal{P}}(a)) > 1-\epsilon$, the action prescribed by *any* optimal policy converges to $a \in A$ with probability greater than $1-\kappa$.

For Nash equilibria (where the agent has correct beliefs about the consequences of every action), these two stability notions coincide if for every pair of actions a, a' there is a $\mathcal{P} \in \mathcal{P}$ such that $U(a, \delta_{\mathcal{P}}) \neq U(a', \delta_{\mathcal{P}})$

Characterization of Uniform Stability

Theorem (Characterization theorem)

Action $a \in A$ is uniformly stable if and only if it is a uniformly strict Berk-Nash equilibrium.

- This is the first if and only characterization of stability under misspecified learning.
- Differs from past work in covering the case where the agent perceives an information value from experimentation.

Proof Sketch for Uniformly Strict Implies Uniformly Stable

- Since a is a uniformly strict B-N equilibrium, a is the unique myopic best reply to every action- contingent outcome distribution p in a ball around the KL minimizers $\hat{\mathcal{P}}(a)$.
- The agent needn't be myopic, and non-equilibrium actions can convey information.
- However, this information is useless, since uniform strictness implies the agent would want to play a regardless of what they learn about $p_{a'}$ for other actions a'.
- Then we use the fact that a transformation of the odds-ratio between the non-KL minimizers and KL minimizers is a positive supermartingale (as in Frick, Iijima, and Ishii, 2020) to generalize the "active supermartingale" result of Fudenberg and Levine (1993) to misspecification.
- Use the Dubins upcrossing inequality to show that if this odd ratio starts sufficiently low, with an arbitrarily large probability it never crosses the threshold needed to switch action.

Proof Sketch for uniformly stable implies uniformly strict

- If a is not a uniformly strict B-N there is *some* belief over minimizers such that a is not strictly optimal.
- So it is not the limit outcome under *some* optimal policy.
- Theorem 1 and Theorem 2 combined give

$$\text{Unif. Strict B-N} = \text{Unif. Stable} \subseteq \text{Stable} \subseteq \text{Unif. B-N.}$$

- In a rich environment, for every KL minimizer for every action, there is a nearby model in $\mathcal{P}$ where the action's utility is relatively lower.This seems like a relatively weak condition, but it rules out the common assumption of finite-support priors.
- Theorem 3 shows that in rich environments uniformly strict B-N $\Leftrightarrow$ stability so

$$\text{Unif. Strict B-N} = \text{Unif. Stable} \overset{rich}{=} \text{Stable} \subseteq \text{Unif. B-N.}$$

Positive Attractiveness

- Another natural notion of a being a long-run outcome is that for every prior belief there is a strictly positive probability that the agent's action converges to a.

Definition (Positively Attracting)

Action $a \in A$ is positively attracting if for every optimal policy π

$$\mathbb{P}_\pi \left[\lim_{t \to \infty} a_t = a \right] > 0 \,.$$

Causation Neglect

- When the agent has causation neglect they believe that the distribution over outcomes is the same for all actions:

$$p_a = p_b \quad \forall a, b \in A, p \in \mathcal{P}.$$

Theorem

Suppose that the agent has causation neglect. If a is a uniformly strict Berk-Nash equilibrium then it is positively attracting.

- *Example:* The agent is randomly matched with an opponent and believes they are playing a simultaneous game, and they are uncertain about the distribution over strategies p in the opponents' population.
- In reality the opponents observe a noisy signal about the action taken by the agent before moving, so $p^*_a \neq p^*_b$ if $a = b$.

Sketch of the Proof of Positive Attractiveness

- Our uniform consistency result guarantees that on every path of outcome realizations, beliefs concentrate around the empirical frequency.

- We use this concentration to show that if the empirical frequency is close to p_a^*, the beliefs concentrate around $\mathcal{P}(a)$.

- Causation neglect guarantees that the empirical frequency is a sufficient statistic.

- We combine this with our stability result to guarantee that once the beliefs get sufficiently close to the KL minimizers, the agent never switches to another action.

Subjective Bandit Problems

- In a subjective bandit problem, the agent's prior μ_0 is a product measure $\mu_0 = \times_{a \in A} \mu_a$. (so the actions are independent arms.)

- In these problems, uniformly strict B-N typically don't exist, even in the correctly specified case, because optimistic off-path beliefs can make other actions better replies.

- But here we can replace uniformity requirement with the requirement that the equilibrium is *weakly identified* (Esponda and Pouzo 2016), meaning that there is a unique conceivable outcome distribution q_a that best matches p_a^*.

Definition (Weak Identification)

A Berk-Nash equilibrium a is weakly identified if for all $p, p' \in \hat{\mathcal{P}}(a)$ we have $p_a = p'_a$.

Theorem

For every subjective bandit problem there is a $\bar{\beta} < 1$ such that if the discount factor $\beta \geq \bar{\beta}$, then every weakly identified strict Berk-Nash equilibrium is positively attractive.

- The proof uses the fact that patient agents experiment with actions that they believe might give them a higher payoff.

- Note that here the result needs the agent to be sufficiently patient.

- In contrast, patience didn't matter for the causation neglect result because there the agent thinks there is no value to experimentation.

Extension to Signals

- We extend our setup to allow for the agent to observe an exogenous signal $s \in S$ before taking their action.
- Here the counterpart of the actions are strategy profiles $\sigma : S \to A$.
- Utility function $u : A \times Y \times S \to \mathbb{R}$.
- The conceivable models are in $\Delta(Y)^{A \times S}$.
- Adding the signals lets us to incorporate i.i.d. payoff shocks as in Esponda and Pouzo (2016).
- Also lets us incorporate another common form of misspecification: signal neglect (see, e.g., Molavi 2019).
- Convergence to uniform Berk-Nash equilibria and the positive attractiveness under causation neglect generalize to this setting once the equilibrium definitions are extended.

Conclusion

- We provide sharp characterizations of the long-run outcomes of misspecified learning, and propose uniformity as a learning refinement of Berk-Nash equilibria.

- We show that all uniformly strict Berk Nash equilibria are stable, and that only uniform Berk Nash equilibria can be stable.

- We then provide the first sufficient conditions for an action to be positively attracting under several forms of misspecification:
  - Causation Neglect;
  - Subjective Bandit Problems;
  - Supermodular Environments.

More Related Literature

- The statistics literature starting with Berk (1966) studies exogenous misspecified Bayesian learning.
- Already mentioned Esponda and Pouzo(2016), Esponda, Pouzo, and Yamamoto (2019), and Frick, Iijima,and Ishii (2020).
- Also related: Fudenberg, Romanyuk, and Strack (2018), Heidhues, Koszegi, and Strack (2018), (He 2019), Molavi, 2019).
- And models of misspecified social learning such as Frick, Iijima, and Ishi (2019) Bohren (2016), Bohren and Hauser (2018), Mailath and Samuelson (2019).
- Fudenberg-Lanzani (2020, in preparation) uses an evolutionary model to study which misperceptions are "robust to mutations."

Thank you!

Definition (Rich)

Definition (Rich)

A problem is *rich* if for every action a, minimizer $p \in \hat{\mathcal{P}}(a)$ and $\varepsilon > 0$ there exists a $p' \in \mathcal{P} \setminus \hat{\mathcal{P}}(a)$ with $||p - p'|| \leq \varepsilon$ such that

$$\mathbb{E}_{p_a}\left[u(a,y)\right] - \max_{b \in A} \mathbb{E}_{p_b}\left[u(b,y)\right] > \mathbb{E}_{p'_a}\left[u(a,y)\right] - \max_{b \in A} \mathbb{E}_{p'_b}\left[u(b,y)\right].$$

Back to slides