So far:

- Learning how to play well
- Learning how to play a Nash, or an optimal strategy in a one-shot game.

Today:

- There is some random "state of nature" θ, whose value agents would like to learn, *e.g.* which technology/restaurant/crop is better, etc.
- How do they learn ? by observing "signals" and/or actions of the players.
- What does this have to do with game theory ?

3 parts:

- social learning: long-lived agents.
- sequential learning: short-lived agents, acting in turn.
- merging and applications.

Social learning

- State of nature θ drawn from some (known) distribution.
- Agents receive private information on θ.
- Form a belief over θ (= distribution of θ, conditional on their information).
- Adjust their behavior/belief over time, as a function of their observations.

Questions:

- Does a consensus eventually emerge ?
- Does this consensus reflect all available information ?

A "simple" example

Setup:

- I players
- binary state of nature $\theta = 0, 1$ with equal probabilities.
- $(\theta, s^1, \ldots, s^I)$ random vector with known distribution.
- vector $s = (s^1, \ldots, s^I)$ of signals.
- each agent i observes (only) s^i.

Dynamics.

- Set $q_1^i := \mathbf{P}(\theta = 1 \mid s^i) = \mathbf{E}[\theta \mid s^i]$ initial belief of i.
- Beliefs are announced and updated. Define recursively $q_n^i := \mathbf{E}(\theta \mid s^i, q_1, \ldots, q_n)$.

Interpretation:

- assume i derives utility $-(\theta - a)^2$ from choosing $a \in [0, 1]$.
- assume i is myopic and actions are observed.
- then $a_n^i = q_n^i$ for each i and n.

A consensus is reached

- (q_n^i) random sequence – it is a function of all initial signals.
- define $\mathcal{F}_n^i := \sigma(s^i, q_1, \ldots, q_{n-1})$ – information of i at time n.
- Then $q_n^i = \mathbf{E}(\theta \mid \mathcal{F}_n^i)$ is a bounded martingale.
- Hence converges a.s., with limit $q_\infty^i = \mathbf{E}(\theta \mid \mathcal{F}_\infty^i)$.
- q_∞^i is the unique minimizer of $\mathbf{E}[(\theta - a)^2 \mid \mathcal{F}_\infty^i]$.
- q_n^j is $\mathcal{F}_{n+1}^i$-measurable, hence q_∞^j is $\mathcal{F}_\infty^i$-meas.
- on the event $q_\infty^i \neq q_\infty^j \in \mathcal{F}_\infty^i \cap \mathcal{F}_\infty^j$, one has

$$\left(\theta - \frac{q_\infty^i + q_\infty^j}{2}\right)^2 < \max\left((\theta - q_\infty^i)^2, (\theta - q_\infty^j)^2\right).$$

- "Hence" $\mathbf{P}(q_\infty^i \neq q_\infty^j) = 0$: all beliefs coincide.

The limit consensus need not reflect all available information.

Example.

- Assume:
  - $s^1 \coprod s^2 \sim \mathcal{B}(\frac{1}{2})$
  - and $\theta = s^1 + s^2 \ mod\ 2$.
- Then $q_1^i = \frac{1}{2}$ w.p. 1 hence $q_\infty = \frac{1}{2}$ w.p. 1.
- However, θ can be deduced from the pair (s^1, s^2).

Here is a positive result:

Theorem

Assume that $s^1, \ldots, s^I$ are iid conditional on θ. Then $q_\infty = \mathbf{P}(\theta = 1 \mid s^1, \ldots, s^I)$ w.p. 1.

The emergence of a consensus is a robust phenomenon.

Let's allow for:

- a coarser action set.
- a richer "monitoring" structure: a player may observe a subset of the other players, possibly random and time-dependent.
- a richer signal structure: signals may be received over time and be influenced by previous actions.
- possibly far-sighted sighted agents.

This includes e.g. multi-armed bandit problems.

Setup

- a set $(\Theta, \mathcal{A}, \mathbf{P})$ states of nature, with $\theta \sim \mathbf{P}$.
- a set I of players (finite or countably infinite).
  - with (compact metric) action set A
  - utility function $u : \Theta \times A \to \mathbf{R}$.
- at each stage n, i observes $s_n^i \in S$, then chooses $a_n^i \in A$.
- transition probability from $\Theta \times (S^I \times A^I)^{n-1} \to S^I$.
- Overall payoff $\sum_{n=1}^{+\infty} \delta^{n-1} u(\theta, a_n^i)$.

Includes:

- previous example
- or players observe only his neighbors in a given graph
- or a random sample
- multi-armed bandit problems.

- a strategy σ^i of i specifies "how to play" in each round n, as a function of $s_1^i, \ldots, s_n^i$.
- a strategy profile $\sigma = (\sigma^i)$ induces a distribution, denoted $\mathbf{P}_\sigma$, over $H_\infty = \Theta \times (S^I \times A^I)^{\mathbf{N}}$ (with the product σ-field $\mathcal{H}_\infty$).
- this defines a *game* with payoff function

$$g^i(\sigma) := \mathbf{E}_\sigma \left[\sum_1^{+\infty} \delta^{n-1} u(\theta, a_n^i) \right].$$

For the presentation, let's

- assume that
  - there is a directed graph over I, such that s_n^i includes a_{n-1}^j whenever $i \to j$: i observes his neighbors' actions.
  - Θ is finite.
- fix a Nash equilibrium σ.

Results

Denote

- p_n^i belief of i at stage n, and $p_\infty^i = \lim_n p_n^i$ (a.s. and L^1).
- A_∞^i set of limit points of (a_n^i) (r.v., $\mathcal{H}_\infty^i$-measurable).
- For $p \in \Delta(\Theta)$, $BR(p) = \text{argmax}_{a \in A} \mathbf{E}_p[u(\theta, a)]$ actions that are myopically optimal when holding the belief p.
- $u_*(p) = \max_{a \in A} u(p, a)$.

Proposition

One has $A_\infty^i \subseteq BR(p_\infty^i)$, $\mathbf{P}_\sigma$*-a.s.*

- Agents *eventually* behave in a myopically optimal way.
- can't keep experimenting forever.

Theorem

Assume that graph G over I is strongly connected. Then

P1 $\mathbf{E}_\sigma[u_*(p^i_\infty)] = \mathbf{E}_\sigma[u_*(p^j_\infty)]$ *for any* i, j.

P2 *if* $i \to j$, *then* $A^j_\infty \subseteq BR(p^i_\infty)$.

Interpretation.

- The *ex ante* expectation of the limit stage payoff is the same for all agents.
- Eventually, each agent thinks that his neighbors are acting optimally: any action that j keeps playing is optimal in the eyes of i.

Warning (see forthcoming examples)

- **P1** does not imply that $u_*(p^i_\infty) = u_*(p^j_\infty)$, $\mathbf{P}_\sigma$-a.s.
- **P2** does not extend to neighbors of neighbors.

Proof.

- one has $u(p^i_\infty, a^i_\infty) \geq u(p^i_\infty, a^j_\infty)$, $\mathbf{P}_\sigma$-a.s. whenever $i \to j$: i believes he gets a payoff at least as good as j
- taking expectations, $\mathbf{E}[u(p^i_\infty, a^i_\infty)] \geq \mathbf{E}_\sigma[u(p^i_\infty, a^j_\infty)]$.
- since a^j_∞ is $\mathcal{H}^i_\infty$-measurable, this implies

$$\mathbf{E}[u(\theta, a^i_\infty)] \geq \mathbf{E}_\sigma[u(\theta, a^j_\infty)].$$

- by strong connectedness, all these expectations are equal.
- hence $u(p^i_\infty, a^i_\infty) = u(p^i_\infty, a^j_\infty)$, $\mathbf{P}_\sigma$-a.s.

□

In the rest of the lecture, we assume – as in the introductory example, that:

- The state of nature is binary: $\Theta = \{0, 1\}$.
- agents are myopic and try to guess the state of nature: $\delta = 0$ and $u(\theta, a) = -(\theta - a)^2$.
- but have only two actions $A = \{0, 1\}$.
- Hence a_n^i is the MAP of θ, *i.e.*, matches the most likely state.
- In addition, agents receive signals only at stage 1, with $(s^i)_{i \in I}$ *iid* given θ.
- In later stages, agents observe the actions of their neighbors in an *undirected* graph G.

Example 1: the line network

[drawing]

- signals are either 0 or 1.
- signals are correct with probability $q > \frac{1}{2}$: $\mathbf{P}(s = \theta \mid \theta) = q$, a.s.
- assume an agent follows his signal whenever indifferent.

Then

- if two neighbors happen to receive the same signal, they repeat the same action forever.
- This is true irrespective of the number of agents.

Example 2: the royal family

[drawing]

We further assume:

- the distribution of $q_1^i = \mathbf{P}(\theta = 1 \mid s^i)$ is absolutely continuous (has a density w.r.t. Lebesgue measure).

Theorem

W.p. 1, the set A_∞^i is the same for each i.

Theorem

There is a sequence (ρ_n) with limit 1 such that

$$\mathbf{P}_\sigma(a_\infty^i = \theta) \geq \rho_n$$

whenever G has n vertices.

Proof sketch of the first theorem

Write:

- $q^i = \mathbf{P}(\theta = 1 \mid s^i)$: private belief (as before).
- $\pi_n^i = \mathbf{P}(\theta = 1 \mid \phi_n^i)$: social belief, where ϕ_n^i is the list of i's neighbors previous actions.
- $p_n^i = \mathbf{P}(\theta = 1 \mid s^i, \phi_n^i)$: posterior belief.
- i picks a_n^i in round n iff $p_n^i := \mathbf{P}(\theta = 1 \mid s^i, \phi_n^i) \geq \frac{1}{2}$.
- one has $\dfrac{p_n^i}{1 - p_n^i} = \dfrac{q^i}{1 - q^i} \times \dfrac{\pi_n^i}{1 - \pi_n^i}$ by Bayes rule.
- Hence i picks $a_n^i = 1$ iff $\pi_n^i + q^i \geq 1$.

We prove that $p^i_\infty = \frac{1}{2} \Leftrightarrow A^i_\infty = \{0,1\}$.

- π^i_n is a function of s^{-i} and of i's past *actions*.
- (π^i_n) converges a.s. with limit π^i_∞.
- conditional on s^{-i} and on θ,

$$\mathbf{P}(p^i_\infty = \frac{1}{2}, \vec{a}^i = \bar{a}^i \mid s^{-i}, \theta) = 0,$$

  where $\vec{a}^i$ is the (random) sequence of i's actions, and $\bar{a}^i$ is any fixed sequence in $\{0,1\}$.
- There are only countably many converging sequences $\bar{a}^i$ with values in $\{0,1\}$. Hence, when summing over all such sequences

$$\mathbf{P}(p^i_\infty = \frac{1}{2}, |A^i_\infty| = 1 \mid s^{-i}, \theta) = 0.$$

- Hence the event $\{p^i_\infty = \frac{1}{2}, |A^i_\infty| = 1\}$ has probability zero.

Proof sketch of the second theorem

Some graph preliminaries.

- $\mathcal{G} := \{(V, E)\}$ is the set of undirected and connected graphs with at most countably many vertices.
- $\mathcal{G}_r : \{(G, i), G \in \mathcal{G}, i \in V\}$ is the set of *rooted* graphs.
- $\mathcal{G}_r^d$ are those rooted graphs with degree at most d.
- For $(G, i) \in \mathcal{G}_r$, and $k \in \mathbf{N}$, $B_k(G, i)$ is the graph obtained from G when dropping vertices whose distance to i exceeds k: k-neighborhood of i in G.
- Two (rooted) graphs (G, i) and $(\tilde{G}, \tilde{i})$ are *isomorphic* if there is a bijection $h : V \to \tilde{V}$ that preserves edges and roots. We then write $(G, i) \sim (\tilde{G}, \tilde{i})$.
- We define a topology on the quotient space:
  - Given (G, i) and $(\tilde{G}, \tilde{i})$, set

$$\delta((G, i), (\tilde{G}, \tilde{i})) := \frac{1}{2^K}, \text{ where } K := \sup\{k : B_k(G, i) \sim B_k(\tilde{G}, \tilde{i})\}.$$

Proposition

$(\mathcal{G}_r^d, \delta)$ is a compact metric space.

- Define k_n^i to be the MAP of a_∞^i given $\mathcal{H}_n^i$.
- For each $G \in \mathcal{G}^d$ and $i \in V$, one has $\mathbf{P}(k_n^i = a_\infty^i) \geq 1 - \varepsilon$ for each $n \geq T$ large enough.

.... : compactness implies that T can be upper bounded, uniformly over *all* $(G, i) \in \mathcal{G}_r^d$.

- Take any two vertices i and j whose distance exceeds $2T$.
- Then k_T^i and k_T^j are conditionally independent, given θ.
- With high probability, they are equal to the same random variable, since $a_\infty^i = a_\infty^j$.
- This can only happen if this r.v. is non-random (given θ).
- Therefore, $a_\infty^i = \theta$, w.p. 1.

Sequential models

The setup

- $\theta \in \{0, 1\}$, with equal probabilities.
- *infinite* sequence of agents, who act *once* in turn.
- agents try to guess θ.
- each agent n chooses $a_n \in \{0, 1\}$ that matches the most likely state given available information.
- Available information consists of a (private) signal s_n, and of some information ϕ_n relative to past decisions.
- Signals are conditionally *iid* given θ.

Q: does (a_n) converge ? to θ ? at which speed ?

An example: assume $s_n \in \{0,1\}$ with $\mathbf{P}(s_n = \theta \mid \theta) = p > \frac{1}{2}$.

- agent 1 picks $a_1 = s_1$.
- agent 2:
  - if $s_2 = a_1$, picks $a_2 = s_2 = a_1$.
  - if $s_2 \neq a_1$, agent 2's belief is $\frac{1}{2}$. Assume $a_2 = s_2$.
- agent 3:
  - if $a_1 = a_2$, $\mathbf{P}(\theta = a_1 \mid a_1 = a_2) > p$. Then, $a_3 = a_1 = a_2$, irrespective of s_3. The same holds for later agents.
  - if $a_1 \neq a_2$, then $\mathbf{P}(\theta = 1 \mid a_1, a_2) = \frac{1}{2}$: a_1 and a_2 "cancel" each other.
- the sequence (a_n) converges a.s, and quickly;
  - there is an asymptotic consensus w.p. 1.
  - with positive probability, $a_\infty \neq \theta$ and learning fails.

- Failure due to the fact that eventually agents (find it optimal to) ignore their private information.
- What happens if:
  - agents need not have access to the entire sequence of past choices ?
  - or if there is always a positive probability that one's signal overturns available evidence, how strong this evidence is ?

A variation

We assume:

- agent n observes either the entire sequence ($I_n = 1$) or his own private signal only ($I_n = 0$).
- The r.v.'s (I_n) are independent, with $I_n \sim \mathcal{B}(p_n)$.
- The I_n are *private information*: agent n does not know exactly what information was available to previous agents.

Set $e_n := \mathbf{P}(a_n \neq \theta)$.

Theorem

There is $k_* > 0$ such that

$$\inf_{(p_n)} \limsup_{n \to +\infty} n e_n = k_*.$$

- "asymptotic" learning is possible: here, $(a_n) \to \theta$ in prob.
- the optimal rate of learning is $\frac{1}{n}$.
- k_* is known.
- an approximately optimal sequence (p_n) is $p_n = \dfrac{1+\varepsilon}{1-p}\dfrac{k_*}{n}$.

It is not known whether "random" monitoring is really needed.

Beyond this example

- Recall that signals (s_n) are iid given θ.
- define $q_n := \mathbf{P}(\theta = 1 \mid s_n)$: *private* belief of agent n.
- We assume that the distribution of q_n is absolutely continuous:
    - signals are informative.
    - no "two" signals have the same informational content.
    - w.p. 1, no player is ever indifferent between the two guesses $\Rightarrow$ no tie-breaking rule.
- Denote by F the cdf of q_n, and by f its density.
- Since $\mathbf{E}[q_n] = \frac{1}{2}$, one has $\int_0^1 F = \frac{1}{2}$.
- We assume for simplicity that the support of f is $[q_{min}, q_{max}]$.
- We will mostly assume $q_{min} = 0$ and $q_{max} = 1$, hence $0 < F < 1$ on $(0, 1)$.

Some basic facts (ctn'd).

- Denote f_θ the density of q_n conditional on θ.
- f_0 and f_1 are uniquely determined by F, with $\frac{f_1(q)}{f_0(q)} = \frac{q}{1-q}$.
- Then $F_1 \leq F_0$ on $[0,1]$, and $F_1 < F_0$ on (q_{min}, q_{max}): the belief that $\theta = 1$ tends to be higher if $\theta = 1$ than not.
- The ratio $\frac{F_1}{F_0}$ is (strictly) increasing on $(q_{min}, q_{max}]$, with limit 0 at q_{min}.
- Similarly, the ratio $\frac{1-F_1}{1-F_0}$ is strictly monotonic as well (exchange the role of the two states).

- Recall that ϕ_n is the extra information available to agent n.
- Define:
  - $\pi_n := \mathbf{P}(\theta = 1 \mid \phi_n)$: "social" belief.
  - $p_n := \mathbf{P}(\theta = 1 \mid s_n, \phi_n)$: posterior belief.
- player n picks $a_n = 1$ iff $p_n \geq \frac{1}{2}$.
- Since signals are conditionally independent,

$$\frac{p_n}{1 - p_n} = \frac{\pi_n}{1 - \pi_n} \times \frac{q_n}{1 - q_n}.$$

- hence $p_n \geq \frac{1}{2}$ iff $q_n + p_n \geq 1$.
- This has probability $1 - F_\theta(1 - \pi_n)$ in state θ.

The case of public actions

Here, $\phi_n = (a_1, \dots, a_{n-1})$.

- $\pi_n = \mathbf{P}(\theta = 1 \mid a_1, \dots, a_{n-1})$.
- Since $a_n \in \{0, 1\}$, π_{n+1} may take only two values (given $(a_1, \dots, a_{n-1})$), with

$$\frac{\pi_{n+1}}{1-\pi_{n+1}} = \frac{\pi_n}{1-\pi_n} \times \frac{1-F_1(1-\pi_n)}{1-F_0(1-\pi_n)} \text{ if } a_n = 1$$

  and

$$\frac{\pi_{n+1}}{1-\pi_{n+1}} = \frac{\pi_n}{1-\pi_n} \times \frac{F_1(1-\pi_n)}{F_0(1-\pi_n)} \text{ if } a_n = 0$$

- (π_n) is a Markov chain (homogenous, with countable state space).
- $\pi_{n+1} \geq \frac{1}{2}$ if $a_n = 1$, $\pi_{n+1} \leq \frac{1}{2}$ if $a_n = 0$.

Analysis (ctn'd).

- The sequence of likelihood ratios $l_n := \frac{\pi_n}{1-\pi_n}$ is a martingale under $\mathbf{P}_0$. Indeed:

$$
\begin{aligned}
E_0\left[l_{n+1} \mid \pi_n\right] &= \mathbf{P}_0(a_n = 1 \mid \pi_n) \times \frac{\pi_n}{1-\pi_n} \times \frac{1-F_1(1-\pi_n)}{1-F_0(1-\pi_n)} \\
&\quad + \mathbf{P}_0(a_n = 0 \mid \pi_n) \times \frac{\pi_n}{1-\pi_n} \times \frac{F_1(1-\pi_n)}{F_0(1-\pi_n)} \\
&= \frac{\pi_n}{1-\pi_n}
\end{aligned}
$$

- Therefore, (l_n) converges $\mathbf{P}_0$-a.s., with limit $l_\infty \in L^1$.

- Recall that $q_{min} = 0$ and $q_{max} = 1$.
- This implies $l_\infty \in \{0, +\infty\}$, w.p. 1.

  Indeed,

  - fix a compact set $K \subset (0, 1)$. Whenever $\pi_n \in K$, the ratio $\frac{\pi_{n+1}}{\pi_n}$ is bounded away from 1.
  - Hence $\mathbf{P}(\pi_\infty \in K) = 0$
- Since $l_\infty \in L^1$, one has $l_\infty = 0$, $\mathbf{P}_0$-a.s.
- Hence $(a_n) \to 0$, $\mathbf{P}_0$-a.s.: eventually, all agents make the correct guess.

Speed of learning.

- What is the expected number of wrong guesses ? This is a measure of welfare loss.
- Define $\tau := \inf\{n : a_n = \theta\}$, the *first* correct guess. What is $\mathbf{E}_0[\tau]$?
- One has

$$\begin{aligned} \mathbf{P}_0(\tau > n) &= \mathbf{P}_0(a_1 = \cdots = a_n = 1) \\ &= \prod_{k=1}^{n} (1 - F_0(1 - \pi_k)) \end{aligned}$$

  where (π_k)'s are social beliefs along the sequence where all guesses are one.
- Along this sequence, $(\pi_n) \to 1$ and

$$\frac{\pi_{n+1}}{1 - \pi_{n+1}} = \frac{\pi_n}{1 - \pi_n} \times \frac{1 - F_1(1 - \pi_n)}{1 - F_0(1 - \pi_n)}$$

- hence

$$\prod_{k=1}^{n}(1 - F_0(1-\pi_k)) \sim \times(1-\pi_{n+1})\prod_{k=1}^{n}(1 - F_1(1-\pi_k)),$$

  and thus

$$\mathbf{P}_0(\tau > n) \sim \times(1-\pi_{n+1})\mathbf{P}_1(a_k = 1 \text{ for all } k \leq n)$$

- By a martingale argument, $\mathbf{P}_1(a_k = 1 \text{ for all } n) > 0$.
- We thus get $\mathbf{P}_0(\tau > n) \sim \mathrm{cte} \times (1-\pi_{n+1}) \sim \mathrm{cte} \times e^{-r_{n+1}}$, where $r_n := \ln \frac{\pi_n}{1-\pi_n}$ is the LLR.
- Therefore, $\mathbf{E}_0[\tau] < +\infty \Leftrightarrow \sum_n e^{-r_n} < +\infty$.
- Along the sequence where $a_n = 1$ for all n, (r_n) is given recursively by

$$r_{n+1} - r_n = \ln \frac{1 - F_1(1-\pi_n)}{1 - F_0(1-\pi_n)} \sim 2F\left(\frac{1}{1+e^{r_n}}\right).$$

- $r_{n+1} - r_n \sim 2F\left(\frac{1}{1+e^{r_n}}\right).$

[Drawing]

- $(r_n) \to +\infty$, but $r_{n+1} - r_n \to 0$.

We have shown that

$$\mathbf{E}_0[\tau] < +\infty \Leftrightarrow \sum_n e^{-r_n} < +\infty,$$

with

$$r_{n+1} - r_n = 2F\left(\frac{1}{1+e^{r_n}}\right).$$

This suggests a continuous-time approximation, and the following theorem.

Theorem

One has $E_0[\tau] < +\infty \Leftrightarrow \int_0^{+\infty} e^{-r(t)} dt < +\infty$, *where*

$$r'(t) = 2F\left(\frac{1}{1+e^{r(t)}}\right) \text{ (and } r(0) = 0\text{).}$$

Setting $u = r(t)$, one obtains.

Corollary

$$\mathbf{E}_0[\tau] < +\infty \Leftrightarrow \int_0^1 \frac{1}{F} < +\infty.$$

This is also equivalent to $\mathbf{E}_0[N] < +\infty$, where $N := |\{n : a_n \neq \theta\}|$ is the total number of wrong guesses.

To sum up:

- $N < +\infty,\ \mathbf{P}_0 - a.s. \Leftrightarrow F > 0$ on $(0,1]$.
- $\mathbf{E}_0[N] < +\infty \Leftrightarrow \int_0^1 1/F < +\infty$.

Applications

- *if $q \sim \mathcal{U}([0,1])$, then $\mathbf{E}_0[\tau] = +\infty$.*
- *if $s \sim \mathcal{N}(\mu_\theta, \sigma^2)$ with $\mu_0 \neq \mu_1$, $\mathbf{E}_0[\tau] = +\infty$ (not obvious).*

The case of private actions

- We assume here $\phi_n = a_{n-1}$: agent n observes s_n and the most recent choice.
- for simplicity, assume symmetry: $f_0(q) = f_1(1-q)$. Equivalently,

$$F_0(q) + F_1(1-q) = 1 \text{ for each } q.$$

- Under symmetry, $e_n := \mathbf{P}(a_n \neq \theta) = \mathbf{P}_0(a_n = 1) = \mathbf{P}_1(a_n = 0)$.
- What is the evolution of (e_n) ?
- The social belief is here $\pi_{n+1} = \mathbf{P}(\theta = 1 \mid a_n)$. It is equal to:
    - $1 - e_n$ following $a_n = 1$
    - and e_n following $a_n = 0$.

- Thus:
  - $\mathbf{P}_0(a_{n+1} = 1 \mid a_n = 1) = 1 - F_0(e_n)$
  - $\mathbf{P}_0(a_{n+1} = 1 \mid a_n = 0) = 1 - F_0(1 - e_n)$.
- and $e_{n+1} = e_n\,(1 - F_0(e_n)) + (1 - e_n)\,(1 - F_0(1 - e_n))$.
- The relations $\dfrac{f_1(q)}{f_0(q)} = \dfrac{q}{1-q}$ and $f_0(q) + f_1(q) = 2f(q)$ allow one to express F_0 as a function of F. Substituting leads to

$$e_{n+1} = e_n - 2\int_0^{e_n} F.$$

- It follows that (e_n) is non-increasing (anyway obvious).
- and that $\lim_n e_n = 0$: (a_n) converges to θ, in probability.
- but not w.p. 1. Intuition.
- **This holds quite generally**

Speed of learning

- Recall that $e_{n+1} - e_n = -2 \int_0^{e_n} F$.
- And $\mathbf{E}_0[N] = \sum_n e_n$ (where N is the number of wrong guesses).
- This again suggests a continuous-time approximation.

Theorem

One has $\mathbf{E}_0[N] < +\infty \Leftrightarrow \int_0^{+\infty} e(t)dt < +\infty$*, where*

$$e'(t) = -2 \int_0^{e(t)} F.$$

Theorem

One has $\mathbf{E}_0[N] < +\infty \Leftrightarrow \int_0^{+\infty} e(t)dt < +\infty$, *where*

$$e'(t) = -2\int_0^{e(t)} F.$$

A change of variable leads to

Corollary

One has

$$\mathbf{E}_0[N] < +\infty \Leftrightarrow \int_0^1 \frac{q}{\int_0^q F} dq < +\infty.$$

- This is the *same* condition as when all actions are public.
- What happens with intermediate monitoring structures is basically unknown.

Merging and applications

Q: an observer has a belief over a data-generating process, and updates his belief over time, as data accumulates. Will he eventually be able to forecast accurately future data ?

This is a purely statistical question.

Game-theoretic motivation:

- repeated games with incomplete information.
- one player has knowledge of the "state of nature", while the other doesn't.
- To what extent will the uninformed agent be able to eventually predict the behavior of the informed player ?

Setup.

- $(\Omega, \mathcal{F})$ is a measurable space, and $(\mathcal{F}_n)$ a filtration over Ω.
- Assume:
  - $\sigma(\mathcal{F}_n, n \in \mathbf{N}) = \mathcal{F}$.
  - each $\mathcal{F}_n$ is generated by a (at most) countable partition of Ω.
- True distribution is $\mathbf{P}$.
- Let μ be an alternative distribution – the belief.
- Given $A \in \mathcal{F}$ and $n \in \mathbf{N}$:
  - $\mathbf{P}(A \mid \mathcal{F}_n)$ is the (conditional) probability of A, given information available at stage n.
  - $\mu(A \mid \mathcal{F}_n)$ is the belief of the observer.

Q: are there conditions that ensure that $\mathbf{P}(\cdot \mid \mathcal{F}_n)$ and $\mu(\cdot \mid \mathcal{F}_n)$ get closer ?

Definition

μ merges with $\mathbf{P}$ if

$$\sup_{A \in \mathcal{F}} |\mu(A \mid \mathcal{F}_n) - \mathbf{P}(A \mid \mathcal{F}_n)| \to 0$$

as $n \to +\infty$, $\mathbf{P}$-a.s..

The prediction error converges to zero, uniformly over all events.

Definition

μ weakly merges with $\mathbf{P}$ if

$$\sup_{A \in \mathcal{F}_{n+1}} |\mu(A \mid \mathcal{F}_n) - \mathbf{P}(A \mid \mathcal{F}_n)| \to 0$$

as $n \to +\infty$, $\mathbf{P}$-a.s..

The short-term error converges to zero.

When this holds, one also has for each k;

$$\sup_{A \in \mathcal{F}_{n+k}} |\mu(A \mid \mathcal{F}_n) - \mathbf{P}(A \mid \mathcal{F}_n)| \to 0$$

Examples. Let $\Omega = \{0,1\}^{\mathbf{N}}$ – repeated coin-tossing.

1. $\mathbf{P}$: the successive outcomes are *iid* with parameter $\frac{2}{3}$. μ assumes that outcomes are conditionally iid given a parameter, which is first drawn from $\mathcal{U}([1/4, 3/4])$.
   - μ weakly merges with $\mathbf{P}$.
   - but μ does not merge with $\mathbf{P}$. Let A be the event where the limit frequency of 1's is 2/3. One has $\mathbf{P}(A \mid \mathcal{F}_n) = 1$ $\mathbf{P}$-a.s. 1 for each n, while $\mu(A \mid \mathcal{F}_n) = 0$ w.p. 1.
2. $\mathbf{P}$: the parameter is equal to $\frac{2}{3}$. Under μ the parameter is either $\frac{1}{2}$ or $\frac{3}{4}$. Then μ does not weakly merge with $\mathbf{P}$.
3. $\mathbf{P}$: the parameter is equal to $\frac{2}{3}$. Under μ, the parameter is either $\frac{1}{2}$ or $\frac{2}{3}$. Then μ merges with $\mathbf{P}$.
4. under $\mathbf{P}$ and μ, the parameter is first drawn according to some densities f, g (respectively), with $f, g > 0$ on $[0,1]$. Then μ merges with $\mathbf{P}$.

Theorem

$P << \mu$ if and only if μ merges with $\mathbf{P}$ for every $(\mathcal{F}_n)$.

- We deal only with the direct implication.
- Since $P << \mu$, there exists (a Radon-Nikodym derivative) $f = \frac{d\mathbf{P}}{d\mu}$ such that $\mathbf{P}(A) = \int_A f d\mu$ for each $A \in \mathcal{F}$. One has $f \geq 0$, and $f > 0$ $\mathbf{P}$-a.s.
- We assume:
  - $\Omega = \mathcal{X}^{\mathbf{N}}$, with $\mathcal{X}$ countable
  - $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, with $\omega = (x_1, \ldots, x_n)$.

- The density of $\mathbf{P}(\cdot \mid x_1, \ldots, x_n)$ wrt $\mu(\cdot \mid x_1, \ldots, x_n)$ is

$$\alpha_n[x_1, \ldots, x_n](\omega) := \frac{f(\omega)}{\mathbf{E}_\mu[f \mid x_1, \ldots, x_n]}$$

  (with $0/0 = 1$, and those ω consistent with $(x_1, \ldots, x_n)$).
- On the other hand, $\mathbf{E}_\mu[f \mid \mathcal{F}_n] \to f$, μ-a.s. as $n \to +\infty$.
- Hence $\lim \alpha_n = 1$, $\mathbf{P}$-a.s.
- This implies $\mathbf{P}(\{\alpha_n > 1 + \varepsilon\} \mid \mathcal{F}_n) \to 0$, for each $\varepsilon > 0$.
- Now, $\sup_{A \in \mathcal{F}} |\mu(A \mid \mathcal{F}_n) - \mathbf{P}(A \mid \mathcal{F}_n)| = \int_{\alpha_n > 1} (\alpha_n - 1) d\mu(\cdot \mid X_1, \ldots, X_n)$.

- To repeat:

$$\sup_{A\in\mathcal{F}} |\mu(A \mid \mathcal{F}_n) - \mathbf{P}(A \mid \mathcal{F}_n)| = \int_{\alpha_n>1} (\alpha_n - 1) d\mu(\cdot \mid X_1, \dots, X_n)$$

$$\begin{aligned} &\leq \varepsilon + \int_{\alpha_n > 1+\varepsilon} (\alpha_n - 1) d\mu(\cdot \mid X_1, \dots, X_n) \\ &\leq \varepsilon + \mathbf{P}(\alpha_n > 1 + \varepsilon \mid X_1, \dots, X_n) \end{aligned}$$

- This implies

$$\limsup \sup_{A\in\mathcal{F}} |\mu(A \mid \mathcal{F}_n) - \mathbf{P}(A \mid \mathcal{F}_n)| \leq \varepsilon,$$

  P-a.s. for each $\varepsilon > 0$.
- Hence $\sup_{A\in\mathcal{F}} |\mu(A \mid \mathcal{F}_n) - \mathbf{P}(A \mid \mathcal{F}_n)| \to 0$, **P**-a.s.

In most game-theoretic applications:

- only short-run prediction errors matter.
- a property stronger than absolute continuity holds – *grain of truth*:

$$\mu = \rho \mathbf{P} + (1-\rho)\mathbf{Q},$$

  for some $\rho > 0$ and some alternative distribution $\mathbf{Q}$.

This allows for quantitative estimates, and not only asymptotic statements.

Motivation:

- Repeated games with incomplete, one-sided, information.
- Reputation models.

Recall: $\mu = \rho \mathbf{P} + (1-\rho)\mathbf{Q}$.

Introduce a "state of nature" θ, equal to $\mathbf{P}$ w.p. ρ, and to $\mathbf{Q}$ with probability $1-\rho$.

Write $p_n := \mu(\theta = \mathbf{P} \mid \mathcal{F}_n)$.

Define:

- $e_n(\mu, \mathbf{P}) := \sum_{x \in \mathcal{X}} |\mu(X_{n+1} = x \mid \mathcal{F}_n) - \mathbf{P}(X_{n+1} = x \mid \mathcal{F}_n)|$: short-run prediction error.
- $V_n := \mathbf{E}_\mu[|p_{n+1} - p_n| \mid \mathcal{F}_n]$: one-step ($L^1$) variation of the martingale (p_n).

Intuitively:

- either the two conditional distributions $\mu(\cdot \mid \mathcal{F}_n)$ and $\mathbf{P}(\cdot \mid \mathcal{F}_n)$ are close, and then e_n is small.
- or X_{n+1} carries information about θ, and then $p_{n+1} - p_n$ is likely not to be small.
- but this cannot happen too often.

Lemma

One has $V_n = p_n e_n(\mu, \mathbf{P})$, μ-a.s. for each n.

- Let $\omega \in \Omega = \mathcal{X}^{\mathbf{N}}$. Write $\omega = (x_n)$, and $\vec{x}_n = (x_1, \ldots, x_n)$.
- Note that $V_n(\omega) = \sum_{x \in \mathcal{X}} \mu(X_{n+1} = x_{n+1} \mid \vec{x}_n) |p_{n+1}(\vec{x}_n, x_{n+1}) - p_n(\vec{x}_n)|$
- Bayes rule gives $p_{n+1}(\vec{x}_n, x_{n+1}) = p_n(\vec{x}_n) \times \dfrac{\mathbf{P}(x_{n+1} \mid \vec{x}_n)}{\mu(x_{n+1} \mid \vec{x}_n)}$.
- Thus

$$\begin{aligned} V_n(\omega) &= \sum_x p_n(\vec{x}_n) |\mathbf{P}(X_{n+1} = x \mid \vec{x}_n) - \mu(X_{n+1} = x \mid \vec{x}_n)| \\ &= p_n(\omega) e_n(\mu, \mathbf{P})(\omega) \end{aligned}$$

Theorem

Let M, α, t be arbitrary. Then:

- P1 $\mathbf{P}(e_n \geq \frac{1}{t\rho M^{1/4}}) \leq t + \frac{1}{\rho\sqrt{M}}$, *for all stages except at most M.*
- P2 $\mathbf{P}(|\mathcal{B}| \geq M) \leq t + \frac{1}{\rho M \alpha^2}$, *where $\mathcal{B} := \{n : e_n \geq \frac{\alpha}{\rho t}\}$.*

According to **P1**, in most stages, the short-run prediction error is small with high probability.

According to **P2**, with high probability, the short-run prediction error is small in most stages.

The proofs of **P1** and **P2** are similar, and are based on two central observations.

First,

- Define $W_n := \mathbf{E}_\mu \left[(p_{n+1} - p_n)^2 \mid \mathcal{F}_n \right]$, quadratic variation of (p_n).
- Since (p_n) is a martingale in L^2,
$$\mathbf{E}_\mu \left[\sum_{n=1}^{+\infty} W_n \right] = \mathbf{E}[p_\infty^2 - p_1^2] \leq 1.$$
- By the (conditional) Cauchy-Schwarz inequality, $V_n \leq \sqrt{W_n}$.

Second,

- the inequality $p_n(\omega) \leq \rho t$ means $\rho \frac{\mathbf{P}(\vec{x}_n)}{\mu(\vec{x}_n)} \leq \rho t$ (with $\omega = (x_n)$).
- by summation, this implies $\mathbf{P}(\{p_n \leq t\rho\}) \leq t$ for each fixed n.
- the same argument shows that $\mathbf{P}(\tau < +\infty) \leq t$, where $\tau := \inf\{n : p_n \leq \rho t\}$.

Proof of **P1**

Remember **P1**: $\mathbf{P}(e_n \geq \frac{1}{t\rho M^{1/4}}) \leq t + \frac{1}{\rho\sqrt{M}}$, for all stages except at most M.

- Let $\mathcal{M} := \{n : \mathbf{E}_\mu[W_n] \leq \frac{1}{M}\}$.
- Since $\mathbf{E}_\mu\left[\sum W_n\right] \leq 1$, $|\mathcal{M}^c| \leq M$. We now let $n \in \mathcal{M}$.
- Since $\mathbf{E}_\mu[W_n] \leq \frac{1}{M}$, one has $\mu\left(W_n \geq \frac{1}{\sqrt{M}}\right) \leq \frac{1}{\sqrt{M}}$.
- Since $V_n \leq \sqrt{W_n}$, $\mu\left(V_n \geq \frac{1}{M^{1/4}}\right) \leq \frac{1}{\sqrt{M}}$.
- $\{e_n \geq \frac{1}{t\rho M^{1/4}}\} \subseteq \{p_n \leq t\rho\} \cup \{p_n > t\rho, V_n \geq \frac{1}{M^{1/4}}\}$.
- Hence $\mathbf{P}(e_n \geq \frac{1}{t\rho M^{1/4}}) \leq t + \frac{1}{\rho\sqrt{M}}$.

Proof of **P2**

Remember **P2**: $\mathbf{P}(|\mathcal{B}| \geq M) \leq t + \frac{1}{\rho M \alpha^2}$, where $\mathcal{B} := \{n : e_n \geq \frac{\alpha}{\rho t}\}$.

- for each n, $V_n \geq \alpha \Rightarrow W_n \geq \alpha^2$.
- Since $\mathbf{E}_\mu \left[\sum_n W_n\right] \leq 1$, the μ-probability of the event $|\{n : W_n \geq \alpha^2\}| \geq M$ does not exceed $\dfrac{1}{M\alpha^2}$.
- One has $\{e_n \geq \dfrac{\alpha}{\rho t}\} \subseteq \{\tau = +\infty\} \cup \{V_n \geq \alpha\} \subseteq \{\tau = +\infty\} \cup \{W_n \geq \alpha^2\}$.
- Hence $\mathbf{P}(|\mathcal{B}| \geq M) \leq t + \frac{1}{\rho M \alpha^2}$.

Some bibliography.

- Part 1:

Aumann. Agreeing to disagree, Ann Stats, 1976.

Bala and Goyal. Learning from neighbors, ReStud, 1998.

Gale and Kariv. Bayesian learning in social networks, GEB, 2003.

Geanakoplos and Polemarchakis. We can't disagree forever, JET, 1982.

Mossel and Tamuz. Opinion dynamics, Prob. Surveys, 2017.

Mossel, Sly and Tamuz. Asymptotic learning on Bayesian social networks, PTRF, 2013.

Mossel, Sly and Tamuz. Strategic learning and the topology of social networks, Ecta, 2015.

Rosenberg, Solan and Vieille. Informational externalities and emergence of consensus, GEB, 2009.

- Part 2:

Bikhandani, Hirshleifer and Welch. A theory of fads, fashion, custom and cultural change as informational cascades, JPE, 1992.

Banerjee. A simple model of herd behavior, QJE, 1992.

Smith and Sørensen. Pathological outcomes of observational learning, Ecta, 2000.

Golub and Sadler. Learning in social networks, Oxford Handbooks, 2016.

Rosenberg and Vieille. On the efficiency of social learning, 2018.

Acemoglu, Dahleh, Lobel, Ozdaglar. Bayesian learning in social networks, ReStud, 2011.

Celen and Kariv. Observational learning under imperfect observation, GEB, 2001.

Hahn-Caruthers, Martynov and Tamuz. The speed of sequential asymptotic learning, JET, 2018.

Monzon and Rapp. Observational learning with position uncertainty, JET, 2014.

Perez, Racz, Sly, Stuhl. How fragile are information cascades, ArXiv, 2018.

Smith and Sørensen. Rational social learning with random sampling, 2013.

- Part 3:

Blackwell and Dubins. Merging of opinions with increasing information, Ann Math. Stats, 1962.

Kalai and Lehrer. Rational learning leads to Nash equilibria, Ecta, 2013.

Kalai and Lehrer. Weak and strong merging of opinions, JME, 1994.

Sorin. Merging, reputation and repeated games with incomplete information, GEB, 1999.

Lehrer and Smorodinski. Merging and learning, in Papers in Honor of D. Blackwell, 1996.