The Speed of Sequential Asymptotic Learning

Wade Hann-Caruthers, Vadim V. Martynov and Omer Tamuz
Caltech

July 22, 2017

Sequential learning

- Banerjee 1992, Bikhchandani, Hirshleifer & Welch 1992.
- **State of the nature** $\theta \in \{h, l\}$.
- Set of **Bayesian agents** $N = \{1, 2, \ldots\}$.
- Agent i receives a **private signal** s_i which depends on θ.
- **Conditioned** on S, private signals are **i.i.d.**
  - **Example.** Gaussian private signals - either $\mathrm{N}(-1, 1)$ or $\mathrm{N}(1, 1)$, depending on θ.
  - **Example.** Bernoulli private signals - either $B(0.51)$ or $B(0.49)$, depending on θ.

Dynamics

- Actions set $A = \{h, l\}$.
- Each state associated with a preferable **action**.
- Each agent has to choose an action in turn.
- Agents observe the **actions** of their **predecessors**.
- **Time periods.** $t \in \{1, 2, \dots\}$.
- At time t, agent t
  - Observes $a_1, \dots, a_{t-1}$ and s_t.
  - Takes action $a_t \in \{h, l\}$.
- The **utility** of agent t is 1 if $a_t = \theta$ and zero otherwise.
- $a_t \in \operatorname{argmax}_{a \in \{h,l\}} \mathbb{P}\left[\theta = a \middle| a_1, \dots, a_{t-1}, s_t\right]$.

Benchmark model

- Benchmark: agents observe **signals**.
- At time t, agent t
  - Observes $s_1, \ldots, s_t$.
  - Takes action a_t.
- The error probability is $e_t = \mathbb{P}[a_t \neq \theta]$.
- **Asymptotic learning.** e_t tends to zero.
- In fact, exponentially quickly.
- The time to learn is

$$T_L = \min\{t \,:\, a_\tau = \theta \text{ for all } \tau \geq t\}.$$

- **Strong asymptotic learning.** $T_L < \infty$ almost surely.
- The public belief is $p_t = \mathbb{P}[\theta = h | a_1, \ldots, a_t]$.
- **Public learning.** $\lim p_t \in \{0, 1\}$.
- Private signals suffice to estimate θ correctly.

Bounded and unbounded signals

- Back to observing actions.
- **Does the private signal information get disseminated?**
- Answers depend on whether signals are bounded.
- Define the private log-likelihood ratio

$$L_t = \log \frac{\mathbb{P}\left[s_t | \theta = h\right]}{\mathbb{P}\left[s_t | \theta = l\right]}.$$

- **Bounded.** The support of L_t is contained in some $[-M, M]$
- There's a limit to how convincing a private signals can be.
- **Unbounded.** There's no such limit (in both directions).

Asymptotic learning

- What is the error probability $e_t = \mathbb{P}[a_t \neq \theta]$?
- What is the distribution of the public belief $p_t = \mathbb{P}[\theta = h | a_1, \dots, a_t]$?
- What about the time to learn T_L?
- Banerjee 1992, BHW 1992, Smith and Sørensen 1992.
- **Bounded signals: information cascade.**
  - No asymptotic learning: $\lim e_t > 0$.
  - No strong asymptotic learning: $\mathbb{P}[T_L = \infty] > 0$
  - No public learning: almost surely $\lim p_t \in (0, 1)$.
- **Unbounded signals: no information cascade.**
  - Asymptotic learning: $\lim e_t = 0$.
  - Strong asymptotic learning: almost surely $T_L < \infty$.
  - Public learning: $\lim p_t \in \{0, 1\}$.

Questions: unbounded signals

- Assume henceforth **unbounded signals**.
- Recall

$$\begin{aligned} p_t &= \mathbb{P}\left[\theta = h | a_1, \ldots, a_t\right] \\ e_t &= \mathbb{P}\left[a_t \neq \theta\right] \\ T_L &= \min\{t \,:\, a_\tau = \theta \text{ for all } \tau \geq t\}. \end{aligned}$$

- Condition on $\theta = h$.
- Asymptotic learning: $\lim e_t = 0$.
  How quickly does this happen?
- Strong asymptotic learning: $T_L < \infty$.
  What is the expectation of T_L?
- $\lim p_t = 1$.
  How quickly does this happen?

Related literature

- **Vives** 1993. Observing beliefs + noise.
- **Sørensen** 1996.
  - "Proof to be finished".
  - Conjecture: expected time to learn is **infinite**.
- **Chamley** 2004. Some estimate.
- **Lobel, Acemoglu, Dahleh and Ozdaglar** 2009. Social networks.
- Concurrent paper by **Rosenberg and Vieille**.

PART II: UNBOUNDED BELIEFS. In state H, $\ell_n \to 0$ and eventually everyone chooses action a_M, and so

$$\ell_{n+1} = \varphi(M, \ell_n) = \ell_n \frac{1 - F^L(\bar{p}_{M-1}(\ell_n))}{1 - F^H(\bar{p}_{M-1}(\ell_n))}$$

For now we shall assume that *extreme signals are not very rare at* 0: i.e. there exists $K > 0$ so that for all $p \in (0,1)$, $F^L(p)/p < K$, and similarly, *extreme signals are very rare at* 1 if there exist $K > 0$ so that for all $p \in (0,1)$, $(1 - F^H(p))/(1-p) < K$. By Lemma A-1.2, with unbounded beliefs and very rare extreme signals, there exist $\varepsilon > 0$ such that $(1-F^L(p))/(1-F^H(p)) > 1 - \varepsilon p$, once p is close enough to zero. Since $\bar{p}_{M-1}(\ell) = \ell/(u+\ell)$ for some $u > 0$, we can conclude that $\langle \ell_n \rangle$ eventually satisfies $\ell_{n+1} > \ell_n - \delta \ell_n^2$ for some $\delta > 0$. Fix n so large that $\ell_{n+1} > \ell_n - \delta \ell_n^2$ and $1 - 2\delta \ell_n > 0$, and take then N so large that $\ell_n \geq 1/((N+n)\log(N+n))$. Then we get $\ell_{n+1} \geq 1/((N+n+1)\log(N+n+1))$. Thus $\sum_n 1/(n \log(n)) = \infty$. (proof to be finished) □

Benchmark model again

- Benchmark: observing signals.
- Let the **public log-likelihood ratio** be

$$\ell_t = \log \frac{p_t}{1 - p_t}.$$

- When observing signals

$$\ell_t = \sum_{\tau=1}^{t} L_t,$$

where L_t is the private log-likelihood ratio.

- So conditioned on θ, ℓ_t grows linearly, by LLN:

$$\lim_t \frac{\ell_t}{t} = \mathbb{E}\left[L_t | \theta\right].$$

- $\ell_t \to \infty$ when $\theta = h$.
- $\ell_t \to -\infty$ when $\theta = l$.

Evolution of public beliefs

Theorem

When observing actions

$$\lim_t \frac{\ell_t}{t} = 0$$

almost surely.

Theorem

For every $r\colon \mathbb{N} \to \mathbb{R}^+$ *with* $\lim_t r(t)/t = 0$ *there exist private signal distributions with*

$$\lim_t \frac{|\ell_t|}{r_t} = 1$$

almost surely.

Evolution of public beliefs

- Let G_θ be the distribution of L_t, conditioned on θ.
- When $a_t = h$,

$$\ell_{t+1} \approx \ell_t + G_l(\ell_t).$$

- Condition on $\theta = h$. Then $a_t = h$ from T_L on.
- When G_l has a thick tail, ℓ_t will grow quickly. When the tail is thin, ℓ_t will grow slowely.

Evolution of public beliefs

- $\ell_{t+1} \approx \ell_t + G_l(\ell_t)$.
- Mild technical assumptions (continuity, convexity etc) on G_θ.

Theorem

Let f be a solution of the differential equation

$$\frac{\mathrm{d}f}{\mathrm{d}t}(t) = G_l(-f(t)).$$

Then, conditioned on $\theta = h$,

$$\lim_{t\to\infty} \frac{\ell_t}{f(t)} = 1$$

almost surely.

The time to learn

- The time to learn is

$$T_L = \min\{t \,:\, a_\tau = \theta \text{ for all } \tau \geq t\}.$$

Theorem (Sørensen 1996)

$\mathbb{E}\left[T_L\right] = \infty$ *for* **some** *signal distributions.*

- Intuition: when all predecessors have taken the same action, an agent is very likely to do the same.
- Enough to show that expected time until first correct action is infinite.

Conjecture (Sørensen 1996)

$\mathbb{E}\left[T_L\right] = \infty$ *for* **every** *signal distributions.*

The time to learn

Theorem

When G_l *and* G_h *have* **polynomial tails** *then* $\mathbb{E}\left[T_L\right] < \infty$.

Upsets and runs

- An **upset** is a time t such that $a_t \neq a_{t-1}$.
- A **run** is a sequence of times between upsets.
- In a run all agents take the same action.
- T_L is the time of the last upset (or 1).
- To show that T_L is small, we show that
  - The number of upsets is small.
  - The length of each run is small.

The number of upsets is small

- What is the probability that there are many upsets?

Theorem

The distribution of the number of upsets has an exponential tail for **every** *signal distribution.*

- Intuition.
  - Whenever agents take the correct action they have a positive probabilty of never making a mistake again.
  - This probability can be bounded away from zero.

The length of each run is small

- Assume private signals have polynomial tails.
- The length of each run has finite expectation.
- But expectation is bigger when run starts after long opposite run.
- But not by too much.
- Easy to see when ℓ_t is uniformly bounded after every upset.
- This is not the case, but we can bound it sufficiently.

Asymptotics of e_t

- The error probability is

$$e_t = \mathbb{P}\left[a_t \neq \theta\right].$$

- We have some upper and lower bounds.
- Large gap between bounds.

Question

How does the asymptotic behavior of e_t depend on the signal distributions?

Thanks!