Experimentation and Dynamic Information Acquisition

OSub Kwon
Ohio State University

March 11, 2021

Introduction

This is a follow-up on:

- My own literature review in this lunch group 3 years ago;
- Renkun's presentation on relating papers and developing ideas.

Non-Strategic Information Source

- Information Acquisition: learning about a payoff relevant state of the world by incurring some (direct or indirect) cost.
- Unlike strategic communication, the source of information is exogenous and non-strategic.

Experimentation vs. Rational Inattention

- Experimentation: Choose something to learn about it, but it can be bad.
- Rational Inattention: Incur a cost to learn about something, and then choose.

Experimentation vs. Rational Inattention

- Experimentation: Choose something to learn about it, but it can be bad.
- Rational Inattention: Incur a cost to learn about something, and then choose.

For rational inattention, a decision ends the process;

Experimentation vs. Rational Inattention

- Experimentation: Choose something to learn about it, but it can be bad.
- Rational Inattention: Incur a cost to learn about something, and then choose.

For experimentation, a decision calls upon another decision.

Experimentation vs. Rational Inattention

- Experimentation: Choose something to learn about it, but it can be bad.
- Rational Inattention: Incur a cost to learn about something, and then choose.

They look at different phenomena.

Experimentation vs. Rational Inattention

- Experimentation: Bolton and Harris (1999), Keller, Rady and Cripps (2005), Guo (2016)
- Rational Inattention: Che and Mierendorff (2019), Zhong (2019)

Two Basic Models of Experimentation

There are two workhorse models of experimentation, Brownian bandit and exponential bandit. In both model:

- Time is continuous.
- Two options (arms) available. The quality of one option (risky arm) is unknown.
- Player(s) can learn faster about the risky arm by pouring more resource to it.

Brownian Bandit

Suppose a player devotes a proportion $\alpha \in [0, 1]$ to the risky arm at this instant (dt): then she receives

$$(1-\alpha)sdt + \sqrt{1-\alpha}\sigma dZ^0 \text{ from the safe arm,}$$

and

$$\alpha\mu dt + \sqrt{\alpha}\sigma dZ^1 \text{ from the risky arm,}$$

where s is known, $\mu \in \{b, g\}(b < s < g)$ is unknown, and both Z's independently follow standard Brownian motion.

Brownian Bandit: Information

Given the accumulation of resource $A = \int \alpha dt$:

- if the state is b, accumulated payoff $\sim N(bA, \sigma^2 A)$;
- if the state is g, accumulated payoff $\sim N(gA, \sigma^2 A)$;

Bayes rule gives the posterior. Note that you learn directly from the payoff, and there is no other signals;

Brownian Bandit: Belief

From Bayes rule, we can derive the evolution of belief

$$dp = p(1-p)\frac{(g-b)}{\sigma}\sqrt{\alpha}dZ^1$$

Belief changes fastest when $p = 1/2$ and stops changing when p is either 0 or 1.

Brownian Bandit: Public Good Problem

Bolton and Harris (1999) considers a public good problem with n players, assuming actions and payoffs are both observable:

- Because everything is observable, the current belief can serve as a state variable;
- Central Trade-off: Risk of choosing a bad arm vs. option value of information.
- Socially optimal strategy is a cutoff strategy, with a stopping belief p_n^* (decreasing in n);
- Symmetric Markov equilibrium strategy is not a cutoff strategy, its stopping belief is strictly between p_n^* and p_1^*;

Brownian Bandit: Symmetric Markov Equilibrium

- Cutoff strategy cannot be an equilibrium strategy (free-riding). Interior solution to keep other players indifferent.
- You do not fully internalize the value of information, so stopping belief cannot reach p_n^* (free-riding).
- Experimenting with risky arm may boost up the belief of other players, so stopping belief lower than p_1^* (encouragement effect).
- Big Caveat: Horner et al. (2019) says that without Markov restriction, efficiency can be achieved.

Exponential Bandit

If a player devotes a proportion $\alpha \in [0,1]$ to the risky arm at this instant (dt), then she receives

$$(1-\alpha)sdt \text{ from the safe arm,}$$

and

$$\text{a lump sum benefit of } \mu \text{ w.p. } \alpha\lambda_\mu dt \text{ from the risky arm}$$

where s is known, $\mu \in \{b, g\}(b < s < g)$ is unknown. Note that the arrival rate of lump sum payoff can depend on state.

Exponential Bandit: Special Cases

Literature focuses on two special cases

- Good news model: $\lambda_g > \lambda_b = 0$. Only good arm generates lump sum benefit.
- Bad news model: $\lambda_b > \lambda_g = 0$. Only bad arm generates lump sum benefit.

Exponential Bandit: Information

In the good news model, only good arm generates lump sum benefit.

- If the lump sum benefit comes, it fully reveals that the arm is good.
- If the lump sum benefit doesn't come, it is either because the arm is bad or the arm is good but you are unlucky.
- The lack of lump sum benefit makes you revise the belief downward.

The opposite applies to the bad news model.

Exponential Bandit: Belief

Along the path where no lump-sum benefit arrives, from Bayes rule, we can derive the evolution of belief:

$$\dot{p} = -(\lambda_g - \lambda_b)p(1-p)\alpha$$

In a good news model ($\lambda_b = 0$), the belief strictly decreases along this path.

Exponential Bandit: Public Good Problem

Keller, Rady, and Cripps (2005) considers a public good problem with n players, assuming actions and payoffs are both observable:

- Because everything is observable, the current belief can serve as a state variable;
- Central Trade-off: Risk of choosing a bad arm vs. option value of information.
- Socially optimal strategy is a cutoff strategy, with a stopping belief p_n^* (decreasing in n);
- Symmetric Markov equilibrium strategy is not a cutoff strategy, its stopping belief is equal to p_1^* (no encouragement effect);

Exponential Bandit: Dynamic Delegation (Guo, 2016)

Say a vaccine company (principal, P) needs to delegate the research of a vaccine to a scientist (agent, A):

- The vaccine generates lump sum benefit as in the good news model.
- A has private information regarding the quality of the vaccine.
- Both P and A prefers working on a good vaccine to a bad vaccine (same ordinal preference).
- A values good vaccine more than P (different cardinal preference).
- P only observes the lump-sum payoff if it arrives, but does not observe A's action.
- P designs a menu of contracts to elicit A's private information.

Exponential Bandit: Dynamic Delegation (Guo, 2016)

The optimal strategy is a simple deadline: if no lump-sum payoff is generated before a date, A is asked to stop.

- Note that private information makes some types of A initially more or less pessimistic than P.
- The existence of private information makes the posterior no longer a state variable.
- A time cutoff (i.e. deadline) is needed to stop optimistic A from experimenting too much.
- The deadline is time consistent in that the principal does not want to revise the deadline even if she has a chance to do so.

Information Acquisition

Note that in all of the above papers, the players learn from payoffs (i.e., signals are payoffs).

However, we can use also separate the signals from the payoffs to discuss dynamic information acquisition.

Information Acquisition: Components

A typical dynamic information acquisition problem has four components:

- a decision problem that requires information;
- a class of dynamic learning processes to choose from;
- a flow cost that prevents the DM from acquiring all information at once;
- a stopping time problem of when to stop acquiring information and make decision;

Information Acquisition: Che and Mierendorff (2019)

A decision maker wants to match a binary state (L or R):

- she chooses either to make decision now or to learn and incur a flow cost c;
- if she learns, she chooses to allocate $\alpha \in [0, 1]$ between L-news learning and R-news learning;
- recall that the lack of L-news learning makes you more right (i.e., L-news model is mostly rightward biased), and vice versa.

Information Acquisition: Che and Mierendorff (2019)

Optimal strategy makes extreme DM more extreme and moderate DM more moderate:

- Very extreme DM makes decision without any more information;
- Extreme DM only acquires information that mostly make them more extreme and make fast decision;
- Moderate DM only acquires information that mostly makes her more moderate and make decision only after either L-news or R-news;
- Novel trade-off between SPEED and ACCURACY.

Information Acquisition: Zhong (2019)

Zhong (2019) extends Che and Mierendorff (2019) model to allow any for learning process, any cost function, and any finite state space.

- A combination of exponential learning process and Brownian learning process is always optimal;
- Exponential learning process is almost always better than Brownian learning.

Follow-ups on Renkun's Insight

Renkun talked about how he locates papers he read and how some ideas are generated based on his observation. To his insight, I want to add:

- we want to set priorities in readings;
- we don't want read too many papers on one line of literature;
- we want to do some "reverse engineering" at high level;

Disclaimer: Renkun and I are different people with very different working style.

Setting Priorities: Not All Readings Are Equal

Because we have limited time and energy, we want to give priorities to "important papers" that help us generate ideas:

- Papers that I make a living out of: Keller, Rady, and Cripps (2005), Kamenica and Gentzkow (2011)...
- Papers related to what I work on: Guo (2016)...
- Papers that teaches me high level insight: Kagel and Roth (2000), Charness and Levin (2005)
- Garbage.

Don't Read Too Many Papers in One Literature

However, making a living out of a literature does not mean we want to always stick to this literature:

- The originality of the papers you read decreases very quickly as you get familiar with a literature.
- You want to imitate the first three papers in the literature in terms of idea generation, not the 10th or 11th.

One Example: Gender Manipulation

My idea of "Gender Manipulation" that I am working with Xiaomin is actually inspired by Bayesian Persuasion.

- Basic logic of Bayesian Persuasion: "Information structure" is usually exogenous, let's make it a choice variable.
- My logic of Gender Manipulation: "gender" is usually exogenous, let's make it a choice variable.

Another Example: Contingent Reasoning

I am thinking about a new project on contingent reasoning, somehow inspired by Guo (2016).

- Basic logic of Guo (2016): we have "experimentation", it's usually used under public good, but let's relate that to "delegation."
- My logic: we have "contingent reasoning", it's usually used under adverse selection, but let's relate that to moral hazard/ delegation/cheap talk... anything that works.

Reverse Engineering at High Level

In general, there seems to be a pattern:

- Step 1: Take an important paper, think about how the idea is generated;
- Step 2: Digest that idea generating process at a high level;
- Step 3: Transplant that idea generating process to another literature.

This process is implied when Renkun talks about application, combination, incorporation, and appropriation.

A Counterexample: YouTube

All this being said, there is no fixed template of generating ideas. Here is an idea that came to me from nowhere.

- Recently I have transitioned from a very heavy YouTube user to a very light user;
- All I did is asking YouTube to delete my preference and to not record new preference;
- Given the necessary resource, this can lead to a very interesting field experiment.

Thank You

Questions and Comments: kwon.406@osu.edu
Website: kwonosub.weebly.com