EXPLORATION AND CORRELATION

Evan Piermont (RHUL)
&
Roee Teper (PITT)

3 Oct, 2017

Exploration Problems

Each period a project manager:

- Must choose to invest in project a or b, but not both.
- Observes if the chosen project succeeds or fails.
- Receives utility from the outcome.

Learning via Exploration

- The true state of affairs is a joint distribution over

$$S = S_a \times S_b = \{s_a s_b, s_a f_b, f_a s_b, f_a f_b\}.$$

- The optimal strategy depends on the manager's beliefs regarding the true generating process.
  - Learning: observes outcomes and updates her belief.
  - Tradeoff: immediate "consumption" value / future informational value.

Beliefs

- In applications we specify these beliefs over the outcome space of all projects.
- Bayes rule determines the dynamic of beliefs.
- We then (try to) solve for the optimal strategy.

This Paper

We ask different questions.

1. What can we learn from the manager's preferences over the different investment strategies?

2. Can we identify the beliefs underlying the exploration/exploitation trade off faced by the manager?

Belief Identification

- We are interested in a manger with exchangeable beliefs.
- Since only one action can be taken in each period, the agent's choices can reveal only the margins of her beliefs.
- We introduce a restriction on marginals, across-arm symmetry, ensuring they arise from an exchangeable process.

Uniqueness?

- We provide a simple example in finite horizon where the marginals determine the process.
- In the infinite horizon, even though across arm symmetry implies more restrictions, marginals do not uniquely determine the process.
- Can always find a (unique) representative for which projects are independent conditional on the true parameter.

A Single Project: Exchangeability

- State space: S, Time Periods: N (finite or countable).
- A process ζ, over sequences of realizations, S^N, is **exchangeable** if its distribution is invariant to finite permutations:
  - ζ is exch if, for every $n \in \mathbb{N}$, history $h \in S^N$ and permutation $\pi : n \to n$,
$$\zeta(h) = \zeta(\pi(h)).$$
- It is a mixture of hypergeometric processes in finite horizon, and of i.i.d. in the infinite. (more later)

Updating and Exchangeability

- To discuss Bayesian updating, one needs to observe the evolution of the joint distribution.
- In exploration models, only a single action can be taken in every period; only the margins of the process can be identified.
  - Beliefs about each individual project conditional on the observed history.

Multi-Dimensional Experiments and Limited Observability

- Two actions "project a" and "project b":

$$S = S_a \times S_b$$

- Let $\mathcal{T}$ be the collection of all sequences of the form $T_1, T_2, T_3, \ldots$, where $T_i \in \{S_a, S_b\}$ for every $i \in N$

Belief Structures

- With every $\mathbf{T} = T_1, T_2, \ldots$ we associate a process $\eta_{\mathbf{T}}$ over $\prod_{i \in N} T_i$.
- $\eta_{\mathbf{T}}$ conveys the distribution of outcomes from taking action T_{n+1} following every history of outcomes $h_i \in T_i$.
  - For a permutation $\pi : n \to n$,
$$\pi \mathbf{T} = (T_{\pi(1)}, T_{\pi(2)}, .., T_{\pi(n)}, T_{n+1}, \ldots)$$
  - Similarly, for a finite history $h = (h_1, \ldots, h_n) \in (T_1, \ldots, T_n)$,
$$\pi h = (h_{\pi(1)}, h_{\pi(2)}, .., h_{\pi(n)})$$

Example 1A

Let $N = 2$. The agent believes that each project will have exactly one success, equally likely to be in either period, and, moreover, believes the two projects will succeed and fail jointly.

		$n = 1$			
		s_a, s_b	s_a, f_b	f_a, s_b	f_a, f_b
$n = 0$	s_a, s_b	0	0	0	$\frac{1}{2}$
	s_a, f_b	0	0	0	0
	f_a, s_b	0	0	0	0
	f_a, f_b	$\frac{1}{2}$	0	0	0

Example 1A

The family of marginal beliefs associated with this joint:

$$\eta_{x,y}(s_x, s_y) = \eta_{x,y}(f_x, f_y) = 0$$
$$\eta_{x,y}(s_x, f_y) = \eta_{x,y}(f_x, s_y) = \frac{1}{2}.$$

where $(x, y) \in \{a, b\} \times \{a, b\}$.

- The joint distribution above was the unique joint consistent with these marginals.

Example 1B

What if the manager believed instead the two projects will succeed and fail independently?

$n = 1$ (columns); $n = 0$ (rows)

	s_a, s_b	s_a, f_b	f_a, s_b	f_a, f_b
s_a, s_b	0	0	0	$\frac{1}{4}$
s_a, f_b	0	0	$\frac{1}{4}$	0
f_a, s_b	0	$\frac{1}{4}$	0	0
f_a, f_b	$\frac{1}{4}$	0	0	0

Example 1B

The family of marginal beliefs associated with this joint:

$$
\begin{aligned}
\eta_{x,x}(s_x, s_x) &= \eta_{x,x}(f_x, f_x) = 0 \\
\eta_{x,x}(s_x, f_x) &= \eta_{x,x}(f_x, s_x) = \frac{1}{2} \\
\eta_{x,y}(s_x, f_y) &= \eta_{x,y}(f_x, s_y) = \frac{1}{4} && \text{if } x \neq y.
\end{aligned}
$$

where $(x, y) \in \{a, b\} \times \{a, b\}$.

Example 2

What if the manager believes the projects' intertemporal performance is i.i.d.

$n = 1$

$n = 0$		s_a, s_b	s_a, f_b	f_a, s_b	f_a, f_b	s_a, s_b	s_a, f_b	f_a, s_b	f_a, f_b
	s_a, s_b	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
	s_a, f_b	0	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
	f_a, s_b	0	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
	f_a, f_b	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

Example 2

Both joint distributions impart the exact same restrictions on marginal beliefs:

$$\eta_{x,y}(s_x, s_y) = \eta_{x,y}(f_x, f_y) = \eta_{x,y}(s_x, f_y) = \eta_{x,y}(f_x, s_y) = \frac{1}{4}$$

- In both examples, all joint distributions were exchangeable.

- Only in Example 1 did the marginals determine the joint.

AA-Symmetry

Definition.

$\{\eta_{\mathbf{T}}\}_{\mathbf{T}\in\mathcal{T}}$ satisfies across arm symmetry if

1. If $h \in \mathbf{T} \cap \mathbf{T}'$, then $\eta_{\mathbf{T}}(h) = \eta_{\mathbf{T}'}(h)$.
2. For every $\mathbf{T} \in \mathcal{T}$, $h \in \mathbf{T}$, and finite permutation π,

$$\eta_{\mathbf{T}}(h) = \eta_{\pi\mathbf{T}}(\pi h).$$

A Non-Symmetric $\{\eta_{\mathbf{T}}\}_{\mathbf{T}\in\mathcal{T}}$

- In Examples 1A/B and 2, the marginals satisfy AASym.
- When does it fail? Consider the following example:
    - As long as project a is chosen, belief regarding both projects is $\frac{1}{2}$
    - Once project b is chosen, the realized outcome occurs with probability 1 for both

$$\eta_{ab}(s_a, f_b) = \frac{1}{4}$$
$$\eta_{ba}(f_b, s_a) = 0$$

Symmetry and Consistency

Theorem.

$\{\eta_{\mathbf{T}}\}_{\mathbf{T}\in\mathcal{T}}$ satisfies across arm symmetry if and only if there exists an exchangeable distribution ζ over S^N such that

$$\mathrm{marg}_{\mathbf{T}}\zeta = \eta_{\mathbf{T}}$$

for every $\mathbf{T} \in \mathcal{T}$

AA-Symmetry and Strongly Exchangeability

AA-symmetry of $\{\eta_{\mathbf{T}}\}_{\mathbf{T}\in\mathcal{T}}$ does not uniquely determine a consistent exch process. From Example 2:

- When the projects are i.i.d. between periods, their contemporary correlation was not pinned down.
- The marginals were consistent with the projects' being contemporaneously independent.

Strong Exchangeability

Definition.

An exch distribution ζ is strongly exchangeable if for every history $h = \prod_{i=1}^{n}(h_{a_i}, h_{b_i})$ and permutations $\pi_a, \pi_b : n \to n$,

$$\zeta(h) = \zeta(\prod_{i=1}^{n}(h_{a_{\pi_a(i)}}, h_{b_{\pi_b(i)}}))$$

- Each dimension can be permuted independently.

de Finetti's Representation

- Let $N = \mathbb{N}$: ζ is exch if and only if there exists a prior distribution $\lambda \in \Delta(\Delta(S))$ such that

$$\zeta = \int_{\Delta(S)} \boldsymbol{\mu} \, \mathrm{d}\lambda(\mu)$$

- As if:
  - A parameter in $\Delta(S)$ is chosen according to λ.
  - The agent does not know the chosen parameter, but knows (or believes) λ.
  - Each period, updates her prior according to the outcome of the experiment.
- Such a representation is unique

A de Finetti like Representation of Strong Exchangeability

Theorem.

ζ over $\prod_{\mathbb{N}} S$ is strongly exchangeable if and only if if the support of λ is in $\Delta(S_a) \times \Delta(S_b)$.

- An exch distribution ζ over $S^{\mathbb{N}}$ is a λ-mixture of parameters in $\Delta(S)$.
  - In an exch process, the joint distribution of experiments' outcomes is (inter-temporally) independent conditionally on the true parameter.
- $S = S_a \times S_b$.
  - In a strongly exch process, experiments are also conditionally contemporaneously independent.

AA-Symmetry and Strongly Exchangeability

Theorem.

Assume $N = \mathbb{N}$ and $\{\eta_{\mathbf{T}}\}_{\mathbf{T} \in \mathcal{T}}$ satisfies AA-symmetry. There exists a unique strongly exchangeable distribution ζ over $\mathcal{S}^{\mathbb{N}}$ such that

$$\mathrm{marg}_{\mathbf{T}}\zeta = \eta_{\mathbf{T}}$$

for every $\mathbf{T} \in \mathcal{T}$

Eliciting $\eta_{\mathbf{T}}$

- The model above assumes the marginal—but not the joint—distributions are observable.
- We turn to a decision theoretic exercise to understand when and if this is reasonable.

To be shown:

- Assume we have access to the preferences over exploration strategies from a bandit problem.
- Axiomatization of the representation.
- Only $\{\eta_{\mathbf{T}}\}_{\mathbf{T}\in\mathcal{T}}$ can be (uniquely) elicited from the axioms.

Examples, revisited

- Recall: $N = 2$, $\mathcal{A} = \{a, b\}$, $X = \{s_a, f_a, s_b, f_b\}$.
- Let $u(s_a) = 9$, $u(f_a) = -9$, $u(s_b) = 18$, and $u(f_b) = -18$.
- The DM is an EU maximizer
- Total utility is the sum across the two periods.

Examples, revisited

- For $x, y, z \in \{a, b\}$, let $(x, (y, z))$ denote the strategy:
  - x in the first period.
  - y in the second, conditional on x's success, and z on x's failure.
- For example, $(a, (a, b))$ is the strategy dictating taking action a in the first period, and
  - action a in the second period, if it was a success in the first.
  - and action b in the second in case a failed in the first.

Example 1A, revisited

- The agent believes that each project will have exactly one success, equally likely to be in either period, and, moreover, believes the two projects will succeed and fail jointly.
- The agent's valuations for investment plans are given as follows: $V(x,(y,z)) = 0$ if $y = z$, and

$$
\begin{aligned}
V(a,(a,b)) = V(b,(a,b)) &= \frac{9}{2} \\
V(a,(b,a)) = V(b,(b,a)) &= -\frac{9}{2}.
\end{aligned}
$$

Example 1B, revisited

If on the other hand, the 2 projects were uncorrelated:
$V(x,(y,z)) = 0$ if $y = z$, and

$$
\begin{aligned}
V(a,(a,b)) &= -\frac{9}{2} & \qquad V(b,(a,b)) &= 9 \\
V(a,(b,a)) &= \frac{9}{2} & \qquad V(b,(b,a)) &= -9.
\end{aligned}
$$

Examples, revisited

- In Example 2: all strategies have value of 0.
- Marginals dictate behavior!
- Preference for strategies in bandit problems can identify:
  - Marginals, $\{\eta_{\mathbf{T}}\}_{\mathbf{T}\in\mathcal{T}}$—always.
  - Joint, ζ—only insofar as given by previous discussion (when $N = \mathbb{N}$, upto strong exch).

Framework

- Let X denote a set of outcomes.
- Let $\mathcal{A}$ denote a set of actions; think, the arms of a bandit problem.
- Each action, a, is associated with a set of possible outcomes, $S_a \subseteq X$.

Histories.

A **history of length n** is a sequence of action/outcome realizations.

- That is, let $h = (a_1, x_1) \dots (a_n, x_n)$.
- Let $\mathcal{H}$ and $\mathcal{H}^\infty$ denote all finite and infinite histories, respectively.

Strategies.

A (mixed) strategy is a mapping from finite histories into randomizations (lotteries) of actions:

$$p : \mathcal{H} \to \Delta(\mathcal{A})$$

- Specifies the action to be taken after each history (including the trivial $\varnothing$).
- Let p_h denote the lottery taken after h with $p_h(a)$ the probably of choosing a.
- Our decision theoretic primitive is a preference relation over all strategies.

Evaluations of Histories

If the manager has a utility index $u : X \to \mathbb{R}$ and discount factor δ, assume she values $h \in \mathcal{H}^\infty$ as

$$U(h) = \sum_{n \in \mathbb{N}} \delta^n u(x_n)$$

Subjective Expected Experimentation

- Let $\mu_{h,a} \in \Delta(S_a)$ denote the manager's belief about action a after having observed history h.
- $\{\mu_{h,a}\}_{h\in\mathcal{H},a\in\mathcal{A}}$ and p induce a unique measure over $\mathcal{H}$:

$$\text{pr}(h,(a,x)) = \text{pr}(h) \cdot p_h(a) \cdot \mu_{h,a}(x)$$

- Assume $U(p) = \mathbb{E}\, U(h)$.

Subjective Expected Experimentation

Equivalently:

$$U_h(p) = \mathbb{E}_{p_h}\left[\mathbb{E}_{\mu_{h,a}}\left[u(x) + \delta\, U_{h,(a,x)}(p)\right]\right] \qquad \text{(SEE)}$$

- We show $\left\langle u, \{\mu_{h,a}\}_{h\in\mathcal{H}, a\in\mathcal{A}}, \delta\right\rangle$ can be uniquely identified from preferences.

Belief Structures

The family $\{\mu_{h,a}\}_{h\in\mathcal{H},a\in\mathcal{A}}$ is identified with $\{\eta_{\mathbf{T}}\}_{\mathbf{T}\in\mathcal{T}}$

- Consider $\mathbf{T} = S_{a_1}, S_{a_2}, \ldots$ and $h \in \mathbf{T}$.
- Given $\{\mu_{h,a}\}_{h\in\mathcal{H},a\in\mathcal{A}}$

$$\eta_{\mathbf{T}}(x_1 \ldots x_{n+1}) = \prod_{i\leq n} \mu_{h_{i-1},a_i}(x_i)$$

- There exists a unique (σ-addative) extension.
- This mapping is bijective with the set of processes that satisfy (1) of AA-sym.