Sequential Optimal Inference for Experiments with Bayesian Particle Filters

Remi Daviet
Wharton Marketing Department, University of Pennsylvania
Behavioral experiments are bounded by time and resources considerations. Researchers need to optimize the amount of relevant information with each question.

Questions:

- What is "relevant information"?
- How to optimize the question?
- Can it be done adaptively?
Importance

- Topic emerged in the 70s
  - see Chaloner and Verdinelli [1995] for a review of the Bayesian approach
- A whole field is dedicated to it (Experimental Design)
- Problem well defined, solution is not
- Increase in model complexity has lead to a need to create methods for adaptive designs:
  - DOSE: Imai and Camerer [2019]
  - DEEP: Toubia et al. [2013]
  - ADO: Cavagnaro et al. [2010]
Current Adaptive Methods

<table>
<thead>
<tr>
<th>Adaptive method</th>
<th>DOSE</th>
<th>DEEP</th>
<th>ADO</th>
<th>SOI (this paper)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation in continuous space</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Model Selection</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Exact optimization</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>General inference method</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Table: Comparison of various adaptive methods available in the literature.

- Our method (SOI) is general and has several advantages:
  - Compatible with complex models
  - Multiple objectives (estimation, prediction, model selection, ...)
  - Fast computation allowing for real time estimation
Optimal design?

You are a researcher, we can define a utility for the observations in an experiment (e.g. relevance information): $u(\text{answer}|\text{question})$

e.g.: chose between the following lotteries:

- 50% chance of getting 20USD
- 20% chance of getting 10USD

Is this question useful? How to define useful?
Bayesian Information

We can use the Kullback–Leibler divergence between prior beliefs and posterior beliefs.
Bayesian Information

We can use the Kullback–Leibler divergence between prior beliefs and posterior beliefs

- **Inference**: Between the prior and the posterior on the parameters
  
  \[ p(\theta) \rightarrow p(\theta|\text{obs, question}) \]

- **Prediction**: Between the prior and the posterior on the answer \( y^* \) to a particular question
  
  \[ p(y^*) \rightarrow p(y^*|\text{obs, question}) \]

- **Model selection**: Between the prior and the posterior on models probabilities
  
  \[ p(\text{model}) \rightarrow p(\text{model}|\text{obs, question}) \]
Expected utility

Since we do not know the answer when designing the question, we use expected utility

\[
EU(question) = \sum_{answers} u(\text{answer}|\text{question})p(\text{answer}|\text{question})
\]

Or in continuous answer space:

\[
EU(question) = \int_{\text{answers}} u(\text{answer}|\text{question})p(\text{answer}|\text{question})d_{\text{answer}}
\]

Issue:

- Generally requires a complicated integral over the parameter space \( \Theta \)
Problem: Generally requires a complicated integral over the often high dimensional parameter space $\Theta$. Example for parameter estimation:

$$\max_{\eta} EU(\eta) = \max_{\eta} \int \int \log \left( \frac{p(y|\theta, \eta)}{p(y|\eta)} \right) p(y|\theta, \eta)p(\theta)d\theta dy.$$  

\(\eta\): question (design), \(\theta\): model’s parameter, \(y\): answer

How to solve this computational problem in between questions?
Introducing Sequential Monte Carlo (SMC):
Provides at any time a set of $P$ draws $\theta^{(p)}$ called particles from the prior/posterior distributions.

Benefits:

- Can be used to approximate the integral in the optimization problem
\[
\max_{\eta} \frac{1}{P} \sum_{p=1}^{P} \sum_{y \in \mathcal{Y}} \log \left( \frac{p(y|\theta^{(p)}, \eta)}{p(y|\eta)} \right) p(y|\eta, \theta^{(p)})
\]

- Handles multimodality well
- Computations are parallelizable
The Sequential Optimal Inference (SOI) method:

- Draw $P$ particles from prior
- Repeat:
  - Find optimal next question using particles
  - Observe answer
  - Update particles to reflect posterior (SMC update)
Implementation

Current applications:

- Purchase prediction (Prediction): Daviet (Original paper with theory)
- Choice with context effects (Parameter inference): Bergmann, Daviet, Fehr
- Neural normalization (Model selection): Daviet, Webb
- Social preferences (Model selection): Imai, Bose, Daviet, Nave, Camerer

Note: nobody in New York yet :(
Results

Application:

Uli gave me 30 questions (after harsh negotiations) to identify the indifference set of a given subject (2 options: red/green).

He then proceeded to ask preferences (ranking) between the 2 "indifference" options and a 3rd option (blue). We can thus "see" the indifference curve.
Results: convergence speed (simulation)

Convergence speed: SOI (red) vs. D-Optimal (green) vs. random (blue)
Challenges

- How to facilitate adoption?
  - Currently Matlab and Python algorithm are provided.

- Maximizing over multiple questions in advance?
  - Some approximate approaches are proposed (see paper).

- Possible strategic manipulation?
  - Some different incentive scheme can be used (see paper).
References:


