THE UNIVERSITY of EDINBURGH
informatics

Bayesian Optimal Experimental Design for Simulator Models of Cognition

(AKA Active learning for science)

Slides adapted from **Simon Valentin's**;
work by Valentin, Kleinegesse, Bramley, Gutmann, Lucas
Computational Cognitive Science
15 November, 2022

Active learning

Previously, we discussed models of active learning based on choosing actions that maximize the information we gain.

If these models work, can they help maximize the information we gain as **scientists**?

Running an experiment is an action that can be more or less informative, depending on the experimental design.

Designing Experiments (1)

What kind of information should scientific experiments provide?

- They should allow us to distinguish between competing theories or models.
- If we have psychologically meaningful parameters $\boldsymbol{\theta}$, our data should tell us something about $\boldsymbol{\theta}$.

Designing Experiments (2)

Traditionally, people design experiments by hand, relying on prior experience and intuition. This *may* work, but:

- It can be **time-consuming**.
- It can lead to **uninformative** or even **misleading data**.
- As theories of human behaviour become richer and more complex, these problems are exacerbated.

Designing Experiments (3)

How can we automate experimental design?

- Find designs d that optimize the information we get from our results $\mathbf{y}$ for our variable of interest v, e.g., $H(v) - H(v|y, d)$.
- We can take a Bayesian approach, in which we have priors over models and v and maximize our *expected* information gain.
- In other words: Bayesian optimal experimental design (BOED).

Designing Experiments (4)

BOED is straightforward conceptually, but difficult in practice.

- Estimating expected information gain (or mutual information) is hard in general.
- Exact Bayesian inference is usually intractable.
- For realistic cognitive models, even the likelihood $p(\mathbf{y} \mid \boldsymbol{\theta}, m)$ may even be intractable!
  - E.g.,**simulator** models including our first random-walk model.
- Previous work on BOED was very limited in the class of models that could be considered.

Today: BOED for Simulator Models of Cognition.

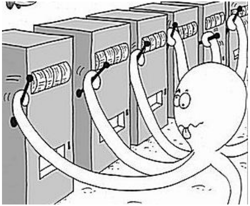
This lecture will introduce:

1. New methods for making BOED work with simulator models, leveraging recent machine learning advances.
2. A proof of concept with new empirical results from a multi-armed bandit task.

Based on work by Simon Valentin, Steven Kleinegesse, and others[1].

[1] Valentin, S., Kleinegesse, S., Bramley, N. R., Gutmann, M. U. and Lucas, C. G. Bayesian Optimal Experimental Design for Simulator Models of Cognition. In NeurIPS 2021 Workshop" AI for Science: Mind the Gaps", 2021. https://arxiv.org/abs/2110.15632

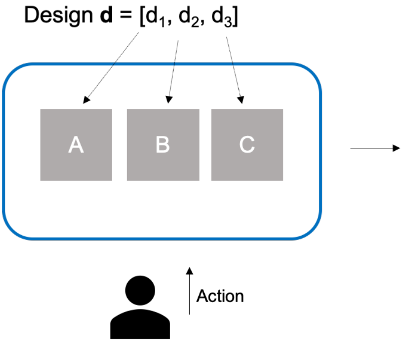
Case Study: Multi-Armed Bandit Tasks



- Here: Multi-armed (Bernoulli) bandits with k arms and fixed reward probabilities.
- Experimental design problem: How do we set the reward probabilities associated with each the k arms, possibly across multiple blocks?

Image credit: http://research.microsoft.com/en-us/projects/bandits/

Bandit Task



Observed Data

Trial	**1**	**2**	**3**	**...**
Action	A	B	B	...
Reward	0	1	0	...

Formalising the Experimental Design Problem

Notation

Notation

- Observed data: $\mathbf{y}$; actions $\mathbf{a}$ and rewards $\mathbf{r}$, so $\mathbf{y} = [\mathbf{a}, \mathbf{r}]$.
- Model parameters: $\boldsymbol{\theta}$; e.g., probability of selecting previous arm.
- Experimental design: $\mathbf{d}$; reward probabilities, e.g., $[0.5, 0.2, 0.6]$.

Recap: Bayesian Inference

Update beliefs about variable of interest $\mathbf{v}$ upon observing data $\mathbf{y}$:

$$p(\mathbf{v} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{v})p(\mathbf{v})}{p(\mathbf{y})},$$

where

$$p(\mathbf{y}) = \int p(\mathbf{y} \mid \mathbf{v})p(\mathbf{v})\ d\mathbf{v}.$$

- Usually, this integral is intractable.
- However, you have seen (and used!) methods for doing approximate Bayesian inference, such as importance sampling.

Intractable Likelihoods

With more expressive models, even the likelihood function $p(\mathbf{y} \mid \mathbf{v})$ may be intractable.

For example, due to hidden variables $\mathbf{h}$.

$$p(\mathbf{y} \mid \mathbf{v}) = \int_{\mathbf{h}} p(\mathbf{y} \mid \mathbf{h}, \mathbf{v}) p(\mathbf{h} \mid \mathbf{v}) \; d\mathbf{h}.$$

Intractable Likelihoods

- If $|\mathbf{h}|$ is large, we cannot apply standard methods here and need to resort to likelihood-free inference methods.
- We need to be able to *simulate* data from the model given prior samples and an experimental design $\mathbf{y} \sim p(\mathbf{y} \mid \mathbf{v}, \mathbf{d})$.
- These models are very common in cognitive science and can capture rich psychological processes! Example: Random walk model (Tutorial 1).

Bayesian Optimal Experimental Design

- Construct a utility function $U(\mathbf{d})$ that describes the worth of an experimental design $\mathbf{d}$.
- Task: Find $\mathbf{d}^* = \arg\max_{\mathbf{d}} U(\mathbf{d})$.
- Principled utility function: *mutual information* (MI), which is equivalent to the expected information gain.
- For $\boldsymbol{v}$, a *variable of interest* that we wish to estimate, we have

$$U(\mathbf{d}) = \mathrm{MI}(\boldsymbol{v}; \mathbf{y}|\mathbf{d}) := \mathbb{E}_{p(\mathbf{y}|\boldsymbol{v},\mathbf{d})p(\boldsymbol{v})}\left[\log \frac{p(\boldsymbol{v}|\mathbf{y},\mathbf{d})}{p(\boldsymbol{v})}\right]. \qquad (1)$$

But computing MI exactly is hard, especially with intractable models.

Our Method

BOED Method

- Maximize the MI with respect to the designs $\mathbf{d}$ using machine learning.
- MINEBED (Kleinegesse & Gutmann, 2020) idea: Train a neural network on simulated data $\mathbf{y}$ at design $\mathbf{d}$ with samples from the prior $p(\boldsymbol{v})$ to tighten a lower bound on the MI.
- We use a bespoke network architecture for behavioral experiments, based on the idea of learning approximate sufficient statistics (Chen et al., 2020).

If you're keen, you can read more about our method here: https://arxiv.org/abs/2110.15632.

BOED Framework

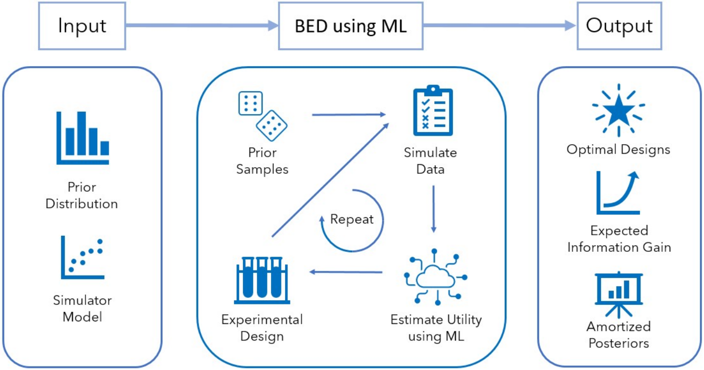


Figure 1: High-level schematic for BOED method.

Towards More Realistic Models of Human Behaviour in Bandit Tasks

Towards More Realistic Models of Human Behaviour in Bandit Tasks

Generalise, combine and modify existing models to capture more realistic behaviours.

Win-Stay Lose-Thompson-Sample (WSLTS)

- Recap: WSLS captures some people's strategy reasonable well.
- But aren't people smarter than switching to another arm selected uniformly at random?
- For example: Why switch to another arm that has failed to provide any rewards in the past?
- WSLTS: WSLS with Thompson Sampling from a reshaped posterior instead of shifting to another arm uniformly at random.

Autoregressive ε-Greedy (AEG)

- ... Or are people (noisy) greedy, but just have a tendency to "stick" to the previously selected arm (not matter whether it produced a reward or not)?
- Or are some people "anti-sticky"?
- AEG: ε-Greedy where the probability of selecting the previous arm is controlled by a separate parameter.

Generalized Latent State (GLS)

- Do people have an internal *latent* exploration-exploitation state that decides between which option to pick whenever the agent encounters the trade-off?
- GLS: Unifies and extends latent-state and latent-switching models (Lee et al., 2011) to allow for more realistic behavior.

Experiments

Simulation Study Setup

- 3 armed bandits, 30 trials per block.
- Model discrimination (MD): 2 blocks, $\boldsymbol{v} \leftarrow m$.
- Parameter estimation (PE): 3 blocks, $\boldsymbol{v} \leftarrow \boldsymbol{\theta}_m$.
- Baseline design, sample rewards from Beta$(2, 2)$ (Steyvers et al., 2009).

Simulation Study Results

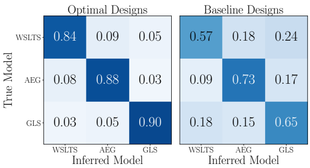


Figure 2: Model recovery confusion matrices.

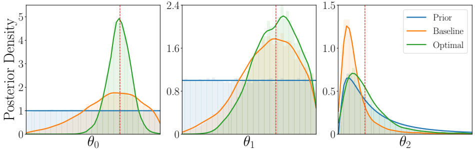


Figure 3: WSLTS model parameter recovery, marginal posterior distributions for three participants.

Real-World Experiments

We ran these two-stage experiments with 317 participants. What did we learn?

- Our designs allow us to reveal participants' strategies at an individual level.
- optimal reward distributions for our tasks included extreme values, e.g, 1 and 0
  - initially counter-intuitive.
  - very different from most previous studies, e.g., beta(2,2).
- Like Steyvers et al., we see substantial individual variability.
- Basic WSLS does not give the best account of most participants' strategies.

Real-World Experiments: A taste

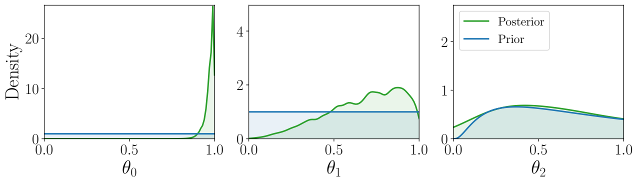
θ_0 : probability of staying after winning

θ_1 : probability of switching after a loss

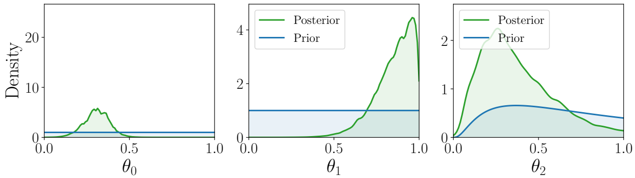
θ_2 : "temperature"

$$\theta_2 \to \infty \Rightarrow \text{WSLS};\ \theta_2 \to 0 \Rightarrow \text{greedy};\ \theta_2 \to 1 \Rightarrow \text{TS}$$

WSLTS Example Participant 1



WSLTS Example Participant 2



Conclusions

Summary

- We present a method for finding optimal experimental designs for simulator models of cognition.
- Our method can optimize experimental designs with realistic numbers of trails and dimensions of the design variable.
- Our evaluation shows that the optimal designs we find outperform designs commonly used in the literature, both with simulated and real-world data.
- Empirical data support complexity of proposed models of human behaviour in bandit tasks.
- Note: There is a lot of scope for future work in this area of research!

Key things to remember

- Bayesian Optimal Experimental Design (BOED) treats the design of experiments as an optimization problem, finding settings that maximise the expected information gain.
- Simulator models can encode realistic and complex psychological processes, and allow for sampling observed data from the model, even if the likelihood is intractable.
- BOED is challenging, but recent machine-learning-based advances make it feasible to optimize experiments used in cognitive science.

References

- Y. Chen, D. Zhang, M. U. Gutmann, A. Courville, and Z. Zhu. Neural approximate sufficient statistics for implicit models. In International Conference on Learning Representations (ICLR), 2021.
- S. Kleinegesse and M. U. Gutmann. Bayesian Experimental Design for Implicit Models by Mutual Information Neural Estimation. In Proceedings of the 37th International Conference on Machine Learning, Proceedings of Machine Learning Research. PMLR, 2020.
- M. D. Lee, S. Zhang, M. Munro, and M. Steyvers. Psychological models of human and optimal performance in bandit problems. Cog. Sys. Research, 12(2), 2011.
- W. R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika, 25(3/4):285–294, 1933.
- R. S. Sutton and A. G. Barto. Reinforcement learning: an introduction. Adaptive computation and machine learning series. The MIT Press, 2nd edition, 2018.
- Valentin, S., Kleinegesse, S., Bramley, N. R., Gutmann, M. U. and Lucas, C. G. Bayesian Optimal Experimental Design for Simulator Models of Cognition. In NeurIPS 2021 Workshop" AI for Science: Mind the Gaps", 2021.