Discussion: "Reinforcing RCTs with Multiple Priors While Learning about External Validity"

by F. Finan & D. Pouzo

Toru Kitagawa (Brown/UCL)

This paper

Setting:

- PM wants to perform a policy choice for a certain population
- PM has access to prior evidence, but not sure if they are useful
- PM can implement the policy in a sequential and adaptive way
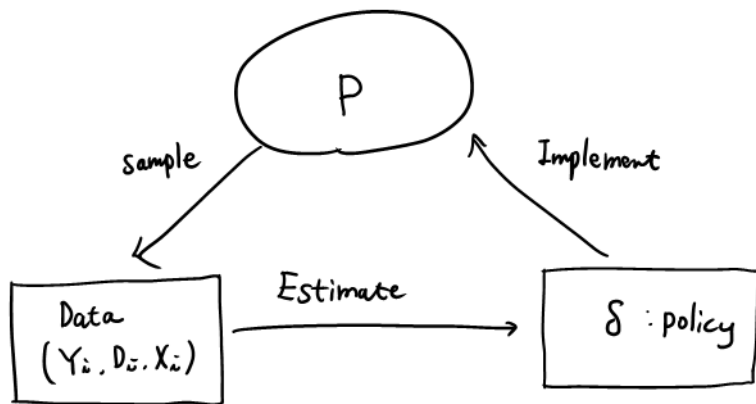
This paper

Setting:

- PM wants to perform a policy choice for a certain population
- PM has access to prior evidence, but not sure if they are useful
- PM can implement the policy in a sequential and adaptive way
- **Question**: How should PM make use of the prior evidence and adaptively implement the policy?
- **Approach**: Multiple priors + multi-arm bandit
- **Analysis**: Show statistical properties and performance guarantees
- **Application**: Charitable fundraising

Outline

1. Flash review: Adaptive experimentation and (Bayesian) bandit
2. Flash review: Multiple priors
3. Overview of the contributions
4. Comments and questions

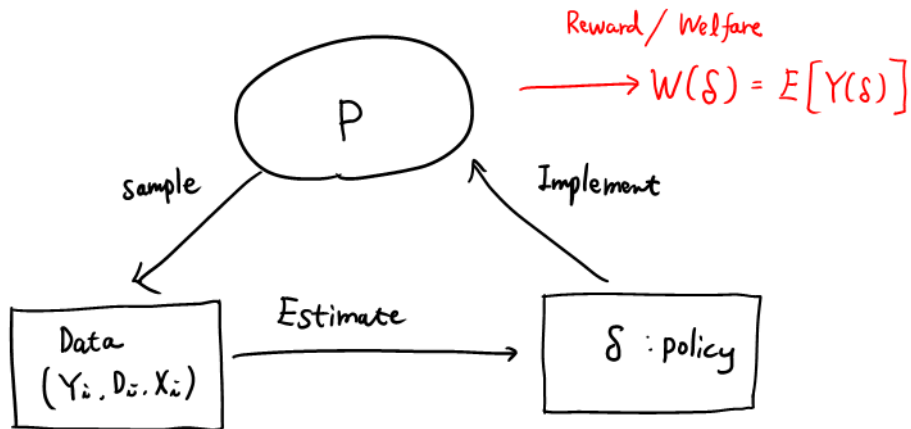# Static treatment choice

- Using the data already collected, how to learn a policy rule δ that optimizes the population welfare (Manski 2004).
- Supervised learning problem.
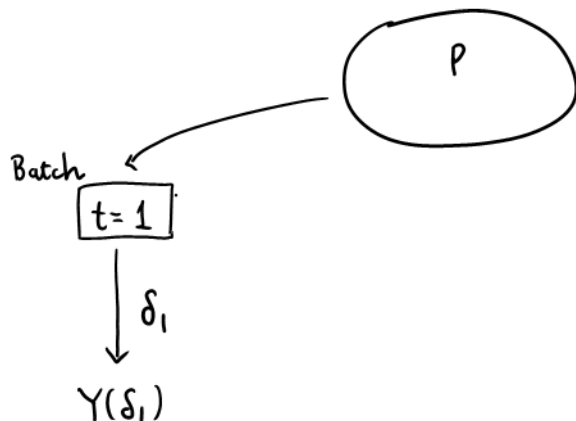
Static treatment choice

- Using the data already collected, how to learn a policy rule δ that optimizes the population welfare (Manski 2004).
- Supervised learning problem.

Reward / Welfare

$$\longrightarrow W(\delta) = E\left[Y(\delta)\right]$$

P

sample

Implement

Data
(Y_i, D_i, X_i)
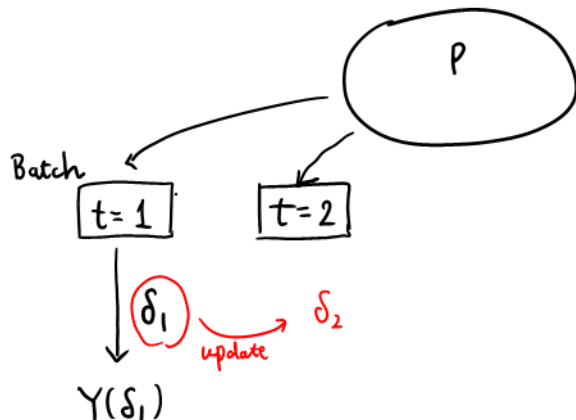
Estimate

δ : policy

Adaptive experimentation and Bandit

- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Adaptive experimentation and Bandit

- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Adaptive experimentation and Bandit

- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Adaptive experimentation and Bandit

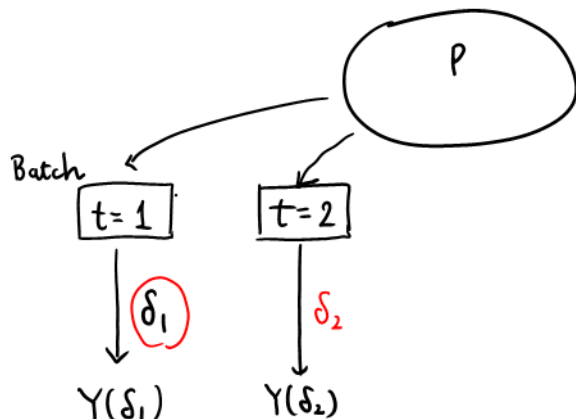- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Adaptive experimentation and Bandit

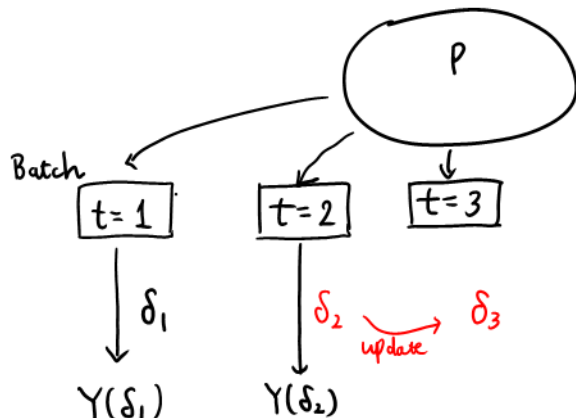- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Adaptive experimentation and Bandit

- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Adaptive experimentation and Bandit

- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Cumulative rewards

Adaptive experimentation and Bandit

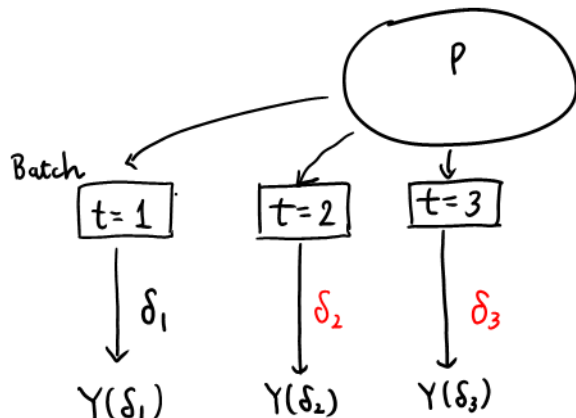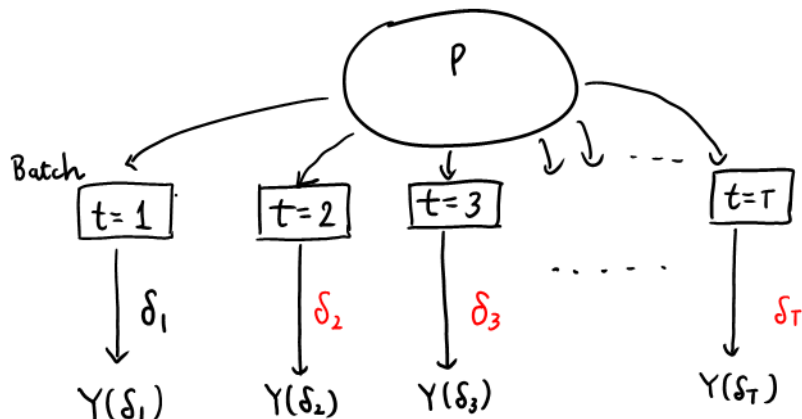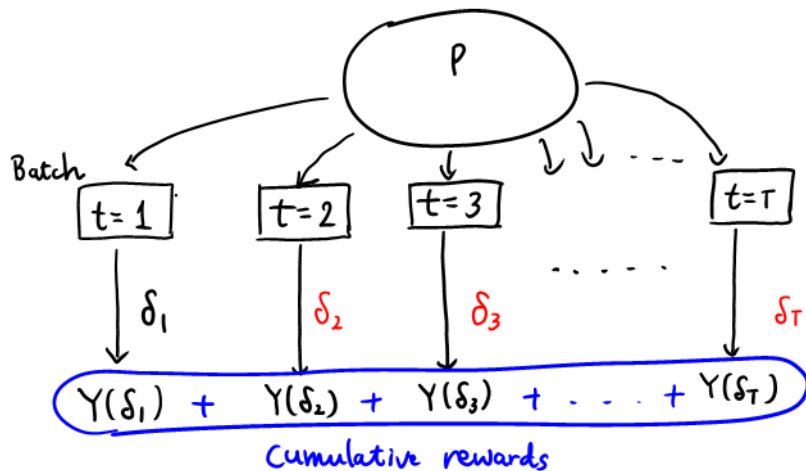- Huge literature in statistics and machine learning. Active research area recently in economics/econometrics.

Performance criteria

- Cumulative rewards (CR): $\sum_{s=1}^{t} Y_s(\delta_s)$
- Best-arm identification (BAI): $\Pr(\delta_t \neq \text{best-arm})$
- (Bayesian) Average reward (Bayes-AR): $\int_{\theta} E_{\theta}[\sum_{s=1}^{t} Y_s(\delta_s)] d\mu_0(\theta)$

Methods	Criteria	Prior
ϵ-Greedy, UCB, etc.	CR & BAI	flat
Thompson sampling	CR & BAI	diffuse
DP (Gittins index)	Bayes-AR	arbitrary

Performance criteria

- Cumulative rewards (CR): $\sum_{s=1}^{t} Y_s(\delta_s)$
- Best-arm identification (BAI): $\Pr(\delta_t \neq \text{best-arm})$
- (Bayesian) Average reward (Bayes-AR): $\int_{\theta} E_{\theta}[\sum_{s=1}^{t} Y_s(\delta_s)] d\mu_0(\theta)$

Methods	Criteria	Prior
ϵ-Greedy, UCB, etc.	CR & BAI	flat
Thompson sampling	CR & BAI	diffuse
DP (Gittins index)	Bayes-AR	arbitrary

- This paper considers informative priors and studies CR and BAI performances
- In the literature, influences of misspecified prior are not well studied

Multiple priors

Multiple priors, $\{\mu_0^0, \mu_0^1, \ldots, \mu_0^L\}$: uncertainty over prior beliefs (Good 1965).
In the current paper, each prior comes from existing evidence

1. **Hierarchical Bayes**: Prior over priors and apply the Bayes rule.
   Bayesian model averaging is a special case
2. **Empirical Bayes**: use data to select a prior and apply the Bayes rule
3. **Gamma minimax**: apply the Bayes rule prior-by-prior and do minimax

Multiple priors

Multiple priors, $\{\mu_0^0, \mu_0^1, \ldots, \mu_0^L\}$: uncertainty over prior beliefs (Good 1965).
In the current paper, each prior comes from existing evidence

1. **Hierarchical Bayes**: Prior over priors and apply the Bayes rule. Bayesian model averaging is a special case
2. **Empirical Bayes**: use data to select a prior and apply the Bayes rule
3. **Gamma minimax**: apply the Bayes rule prior-by-prior and do minimax

- **Paper's proposal**: Obtain the posterior for mean rewards in the hierarchical Bayesian way (Bayesian model averaging)

$$\mu_t^\alpha(\theta) = \sum_{o=0}^{L} \alpha_t^o \mu_t^o(\theta) \tag{1}$$

- Feed the posterior into some heuristic bandit algorithms ϵ-Greedy, Thompson sampling, etc, with a stopping option

Contributions

Conceptual

- Use available evidence as priors for bandit algorithms

Technical

- Model selection consistency: $\lim_{t \to \infty} \alpha_t^\circ = ?$
- Concentration of the posterior means of μ_t^α around the truth
- Uniform convergence rates of cumulative rewards and BAI probability for a wide class of algorithms

Nontrivial, we have to handle dependence of observations over t!

Comments/discussions

1. The Hierarchical Bayes has desirable (frequenstist) performance and robustness with informative priors. How about empirical Bayes or Gamma minimax is used instead?

Comments/discussions

1. The Hierarchical Bayes has desirable (frequenstist) performance and robustness with informative priors. How about empirical Bayes or Gamma minimax is used instead?

2. Bayesian posterior + heuristic algorithms. Since the hierarchical prior input can be viewed as a single prior, might make sense to solve the pure Bayesian DP, i.e., Gittins index policy?

Comments/discussions

1. The Hierarchical Bayes has desirable (frequenstist) performance and robustness with informative priors. How about empirical Bayes or Gamma minimax is used instead?

2. Bayesian posterior + heuristic algorithms. Since the hierarchical prior input can be viewed as a single prior, might make sense to solve the pure Bayesian DP, i.e., Gittins index policy?

3. Compared with a flat prior (i.e., ignoring prior evidence), can we quantify performance gains or losses of introducing an informative prior?

Comments/discussions

1. The Hierarchical Bayes has desirable (frequenstist) performance and robustness with informative priors. How about empirical Bayes or Gamma minimax is used instead?

2. Bayesian posterior + heuristic algorithms. Since the hierarchical prior input can be viewed as a single prior, might make sense to solve the pure Bayesian DP, i.e., Gittins index policy?

3. Compared with a flat prior (i.e., ignoring prior evidence), can we quantify performance gains or losses of introducing an informative prior?

4. For BAI, how much can we gain relative to the two-step sampling design of Hahn, Hirano, & Karlan (2011, JBES)?

Comments/discussions

1. The Hierarchical Bayes has desirable (frequenstist) performance and robustness with informative priors. How about empirical Bayes or Gamma minimax is used instead?

2. Bayesian posterior + heuristic algorithms. Since the hierarchical prior input can be viewed as a single prior, might make sense to solve the pure Bayesian DP, i.e., Gittins index policy?

3. Compared with a flat prior (i.e., ignoring prior evidence), can we quantify performance gains or losses of introducing an informative prior?

4. For BAI, how much can we gain relative to the two-step sampling design of Hahn, Hirano, & Karlan (2011, JBES)?

5. In many social programs, adaptive experimentation can be hindered by the time lag for observe welfare-relevant outcomes. Are surrogate outcomes useful?