A Bandit Approach to Indirect Inference

Felix Steinberger Eriksson

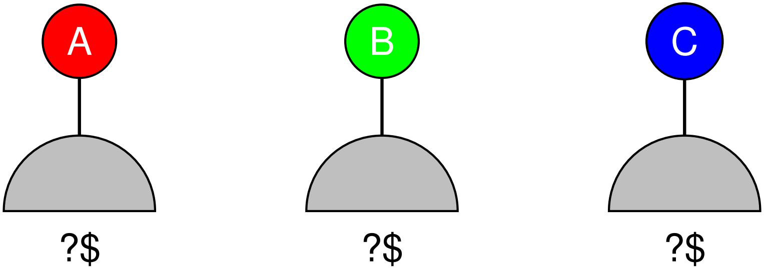
17 May 2023

How to Successfully Gamble Your Way to a Parameter Estimate

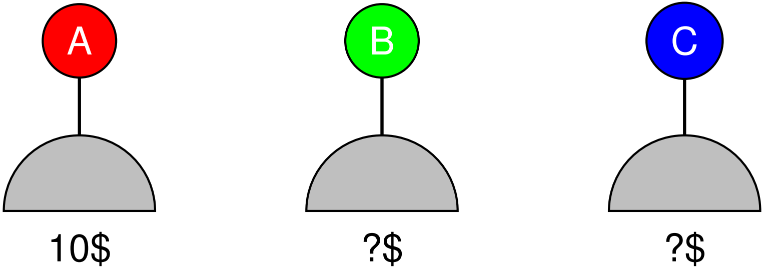
Felix Steinberger Eriksson

17 May 2023

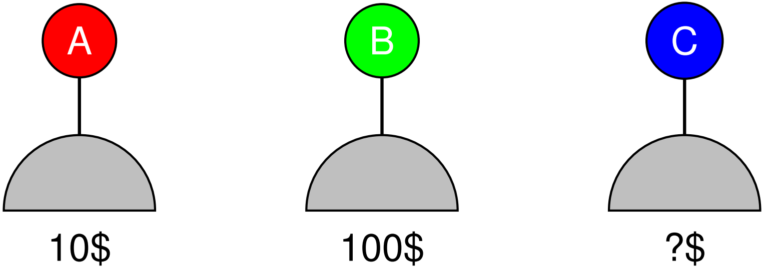
Reward Levers



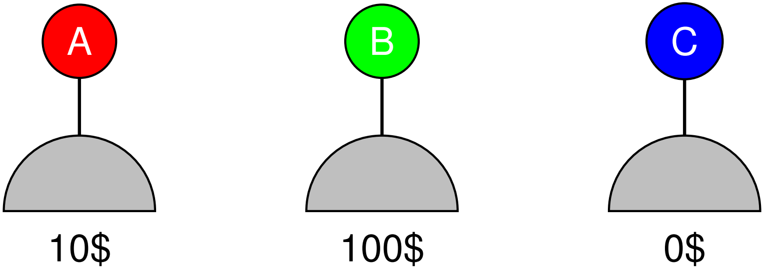
Reward Levers



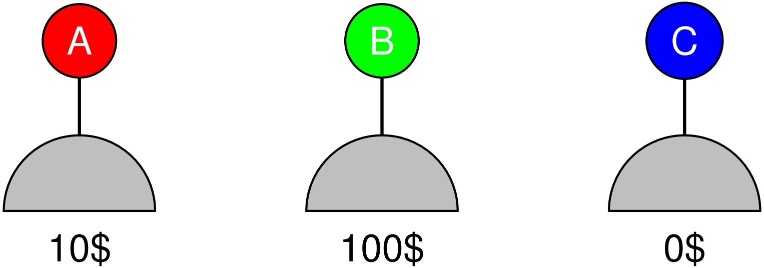
Reward Levers



Reward Levers

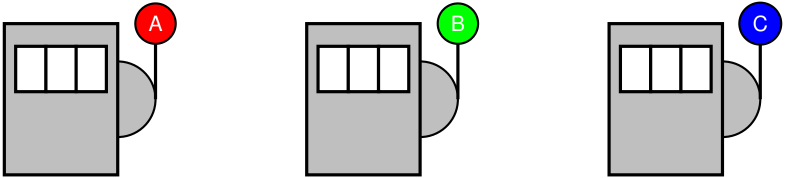


Reward Levers



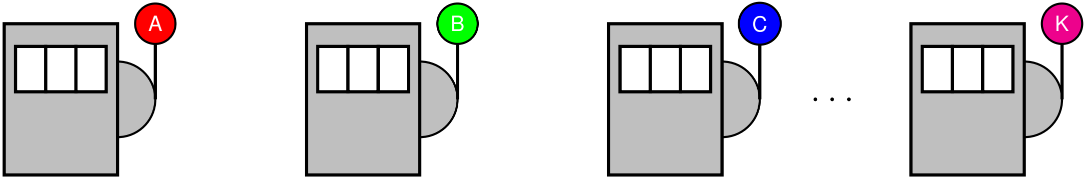
Pull green lever lots, get rich!

Stochastic Bandits



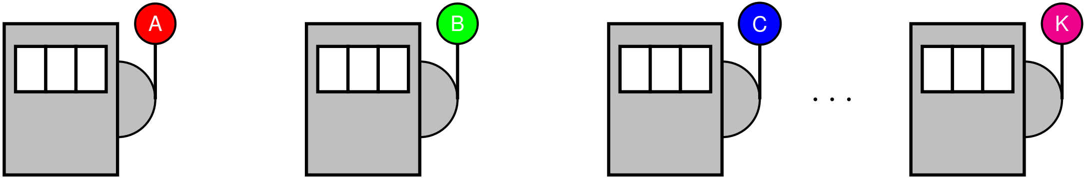
$r_A \sim F_A$ $\quad$ $r_B \sim F_B$ $\quad$ $r_C \sim F_C$

Stochastic Bandits



$r_A \sim F_A$ $\quad$ $r_B \sim F_B$ $\quad$ $r_C \sim F_C$ $\quad$ $r_K \sim F_K$

Stochastic Bandits



$r_A \sim F_A$ $r_B \sim F_B$ $r_C \sim F_C$ $r_K \sim F_K$

Random rewards, need some strategy to select arms.

Bandit algorithms

Bandit algorithms

ϵ-greedy

Keep estimates of rewards of arms. With some small probability ϵ, select an arm uniformly at random. Otherwise, select the arm with best estimated reward.

Bandit algorithms

ϵ-greedy

Keep estimates of rewards of arms. With some small probability ϵ, select an arm uniformly at random. Otherwise, select the arm with best estimated reward.

Upper Confidence Bound (UCB)

Keep estimates of rewards that incorporate growing uncertainty for arms not played in a long time.

Bandit algorithms

Takeaway

There is theory on **smart** ways to select which arm to play.

Parameter estimation

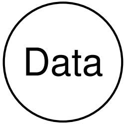
Collection $\mathcal{M} = \{M(\theta) \big| \theta \in \Theta\}$ of parametric models.

Parameter estimation

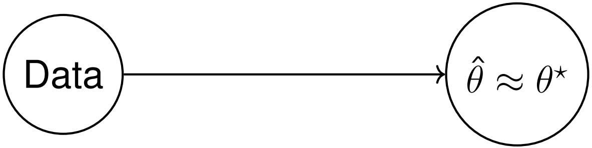
Collection $\mathcal{M} = \{M(\theta) \big| \theta \in \Theta\}$ of parametric models.

Given data sampled from some $M(\theta^\star)$, want to identify $\theta^\star$.

Parameter estimation



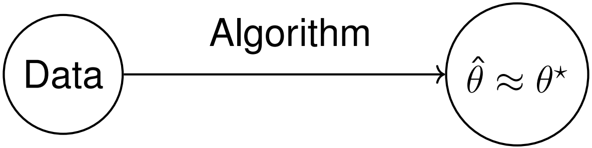
Parameter estimation



Parameter estimation

$$\text{Data} \xrightarrow{\mapsto} \hat{\theta} \approx \theta^{\star}$$

Parameter estimation



Parameter estimation

$$\text{Data} \xrightarrow{\text{Divine intuition}} \hat{\theta} \approx \theta^{\star}$$

Indirect inference

Indirect inference

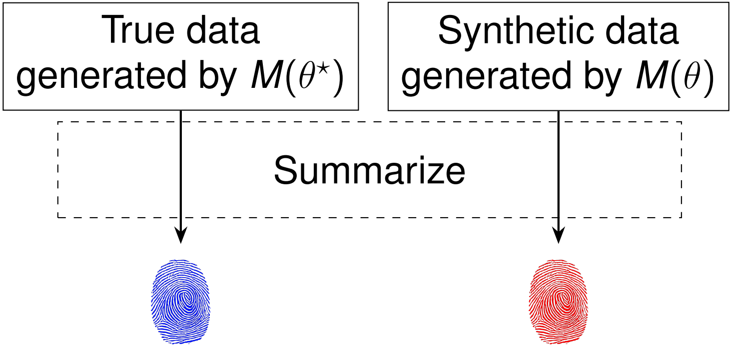
$$\boxed{\begin{array}{c}\text{True data}\\ \text{generated by } M(\theta^\star)\end{array}}$$

Indirect inference

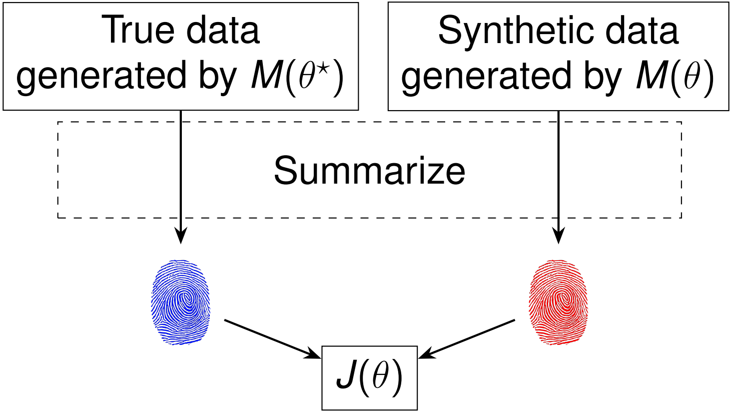
True data generated by $M(\theta^\star)$

Synthetic data generated by $M(\theta)$

Indirect inference



Indirect inference



Indirect inference

True data generated by $M(\theta^\star)$

Synthetic data generated by $M(\theta)$

Summarize

$J(\theta)$

Pick $\hat{\theta} \in \arg\min_{\theta \in \Theta} J(\theta)$!

Easy, right?

$J(\theta)$ has no closed form, but we can evaluate it, right?

Easy, right?

$J(\theta)$ has no closed form, but we can evaluate it, right?

Generate some data, calculate thumbprint, compare similarity.

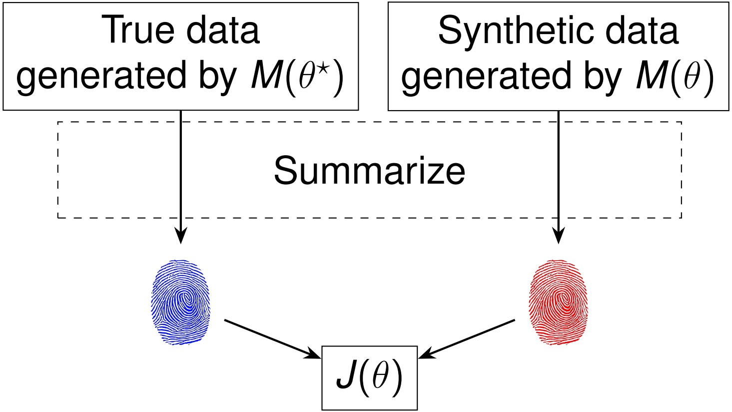
Easy, right?

$J(\theta)$ has no closed form, but we can evaluate it, right?

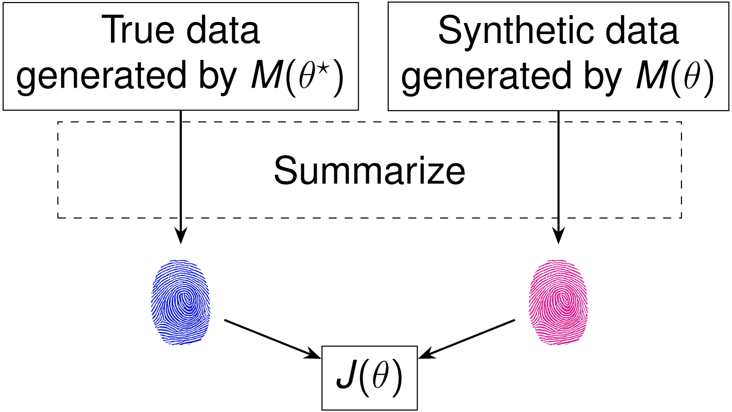
Generate some data, calculate thumbprint, compare similarity.

Use gradient descent or any other standard optimization tool (and hope the function is convex...)

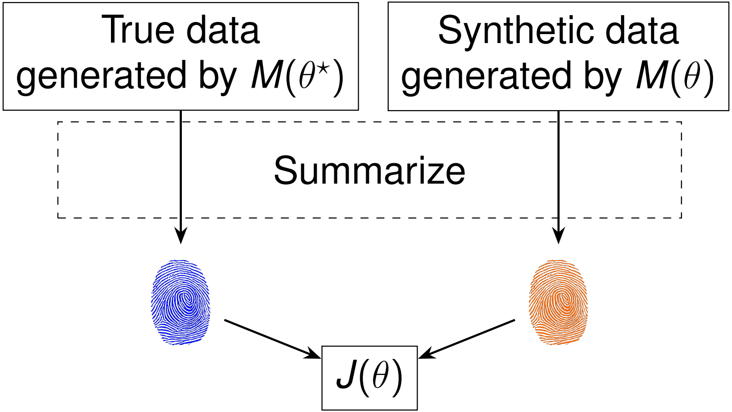
Problem!


True data
generated by $M(\theta^\star)$
Synthetic data
generated by $M(\theta)$
Summarize
$J(\theta)$

Problem!


True data
generated by $M(\theta^\star)$
Synthetic data
generated by $M(\theta)$
Summarize
$J(\theta)$

Problem!


True data generated by $M(\theta^\star)$
Synthetic data generated by $M(\theta)$
Summarize
$J(\theta)$

Problem!

Each "evaluation" of $J(\theta)\dots$

Problem!

Each "evaluation" of $J(\theta)\ldots$

is a sample from $F_{J(\theta)}$!

Treat Indirect Inference as a Bandit Problem

Each parameter θ_i in $\Theta = \{\theta_1, \ldots, \theta_K\}$ is an arm with reward distribution $F_{J(\theta_i)}$.

Treat Indirect Inference as a Bandit Problem

Each parameter θ_i in $\Theta = \{\theta_1, \ldots, \theta_K\}$ is an arm with reward distribution $F_{J(\theta_i)}$.

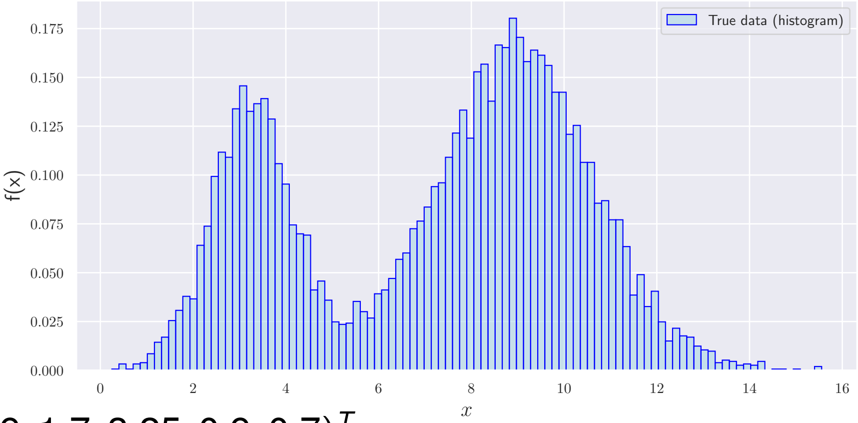
Use bandit optimization methods to find best arm (i.e. best $\theta \in \Theta$).

Bandit Optimization for Indirect Inference works!

Mixture normal model:

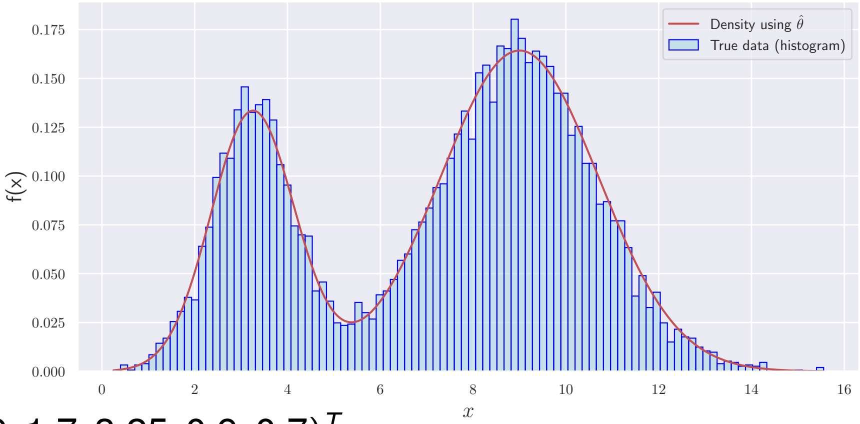
$$f_\theta(x) = pN(x; \mu_1, \sigma_1^2) + (1-p)N(x; \mu_2, \sigma_2^2)$$

Bandit Optimization for Indirect Inference works!



$$\theta^{\star} = (9, 1.7, 3.25, 0.9, 0.7)^T$$

Bandit Optimization for Indirect Inference works!



$$\hat{\theta} = (9, 1.7, 3.25, 0.9, 0.7)^T$$

First steps...

First steps...

- Discretization

First steps...

- Discretization
- More difficult for finer grids due to similar rewards.

Next steps?

Continuous action spaces!

Next steps?

Continuous action spaces!

- Extant theory

Next steps?

Continuous action spaces!

- Extant theory
- Function approximation

Thank You!

Felix Steinberger Eriksson
felixse@kth.se