Sequential Bayes-Optimal Policies for Multiple Comparisons with a Control

Jing Xie, Peter I. Frazier

School of Operations Research and Information Engineering, Cornell University, USA

Cornell University

Introduction

Multiple Comparisons with a Control (MCC)

- determines which alternative systems under consideration have mean performance exceeding a known threshold;
- explores the unknown objective set $\mathbb{B} = \{x : \theta_x \geq d_x\}$.

Bayes-Optimal Fully Sequential Policies for Allocating Simulation Effort

- can be characterized and computed efficiently, using techniques from multi-armed bandits and optimal stopping.
- are flexible in the sense that they
  - allow limitations on the ability to sample to be modeled with either a *random stopping time* or *sampling costs* or both;
  - allow sampling distributions from any *exponential family*.

Problem Formulation

- Prior distributions from the *conjugate* exponential family are placed on the parameters of the sampling distributions. Parameters of these priors / posteriors form a stochastic process $(S_n)_{n\geq 0}$.
- Suppose there is some random time horizon T beyond which we will be unable to sample. T is geometrically distributed with parameter $1-\alpha$; allow $T = +\infty$ a.s., in which case $\alpha = 1$.
- Our estimate of $\mathbb{B}$ given the available information after n samples, i.e., $\mathcal{F}_n$, is $B_n = \{x : \mathbb{P}\{\theta_x \geq d_x \mid \mathcal{F}_n\} \geq 1/2\}$. When sampling stops, we receive a *reward* equal to the total number of alternatives correctly classified by this estimate.
- A **policy** π is composed of a **sampling rule** for choosing an adapted sequence of sampling decisions $(x_n)_{n\geq 1}$, and a **stopping rule** for choosing an adapted stopping time τ.
- Our goal is to find a policy that maximizes the **expected total reward**, i.e., (c_x is the sampling cost for alternative x)

$$\sup_{\pi} \mathbb{E}^{\pi}\left[\sum_{x\in B_{\tau\wedge T}} \mathbf{1}_{\{x\in\mathbb{B}\}} + \sum_{x\notin B_{\tau\wedge T}} \mathbf{1}_{\{x\notin\mathbb{B}\}} - \sum_{n=1}^{\tau\wedge T} c_{x_n}\right].$$

The Optimal Solution

We solve this sequential Bayesian MCC problem using **dynamic programming (DP)**. The **value function** is

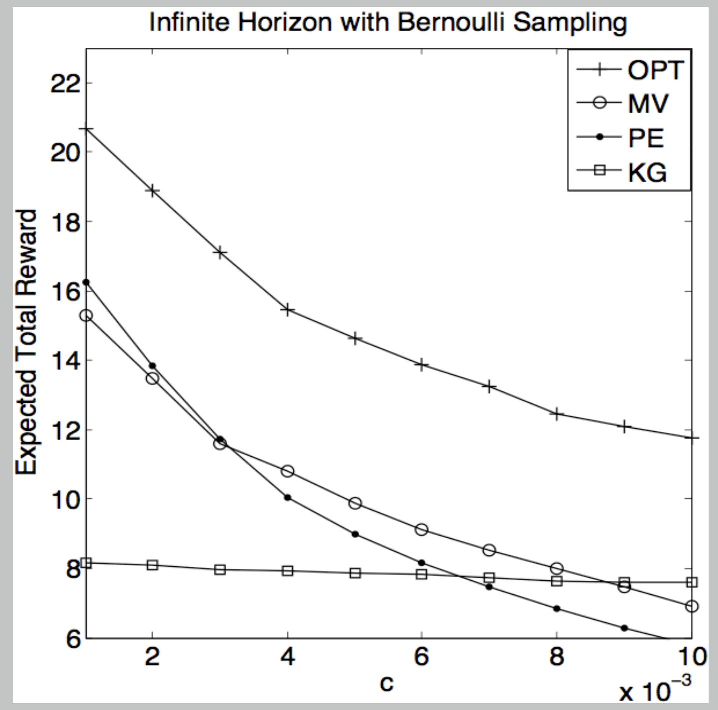
$$V(s) = \sup_{\pi} \mathbb{E}^{\pi}\left[\sum_{n=1}^{\tau} \alpha^{n-1}\mathcal{R}_{x_n}(S_{n-1,x_n}) \,\middle|\, S_0 = s\right],$$

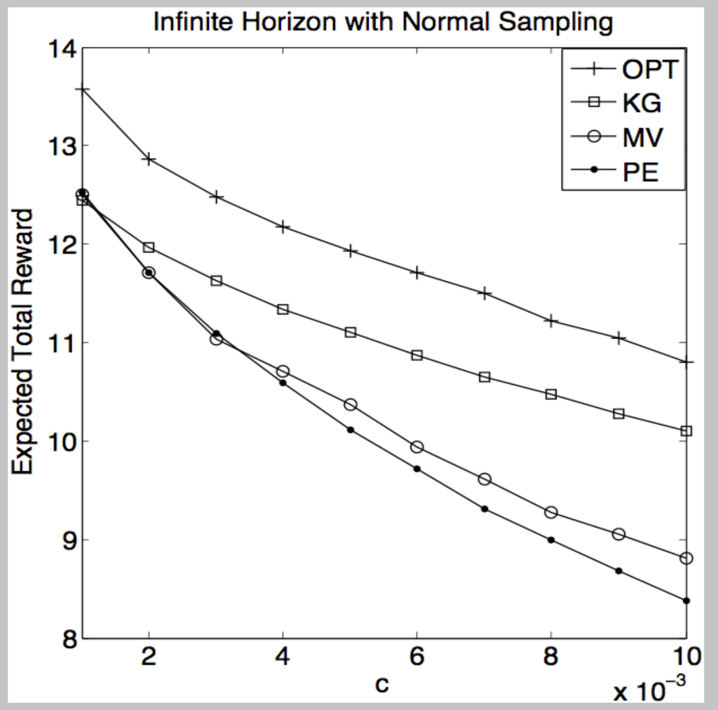
where $\mathcal{R}.(\cdot)$ are single-period reward functions.

- Optimal policies $(x_n^*)_{n\geq 1}$ and τ^* are theoretically specified.
- Approximations are applied to the numerical implementations.

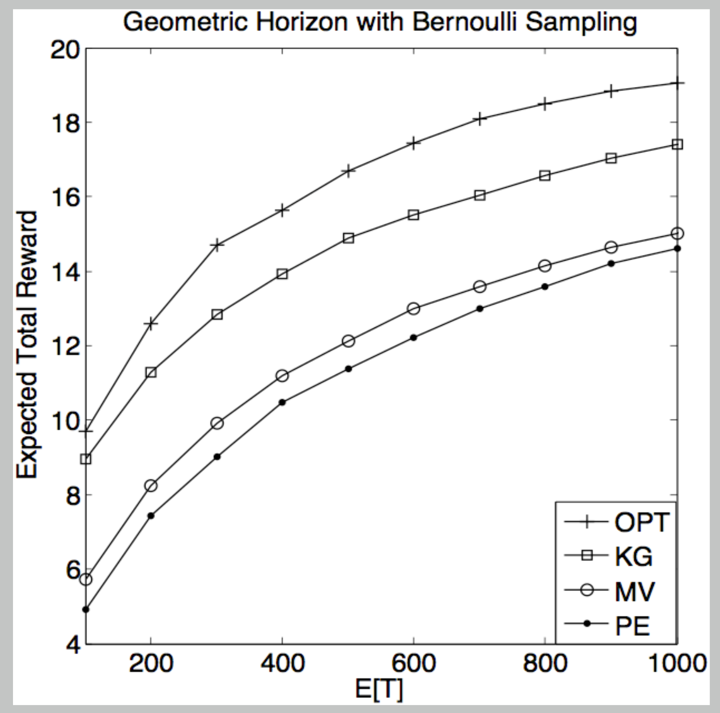
Bayes-Optimal (OPT) vs Other Policies

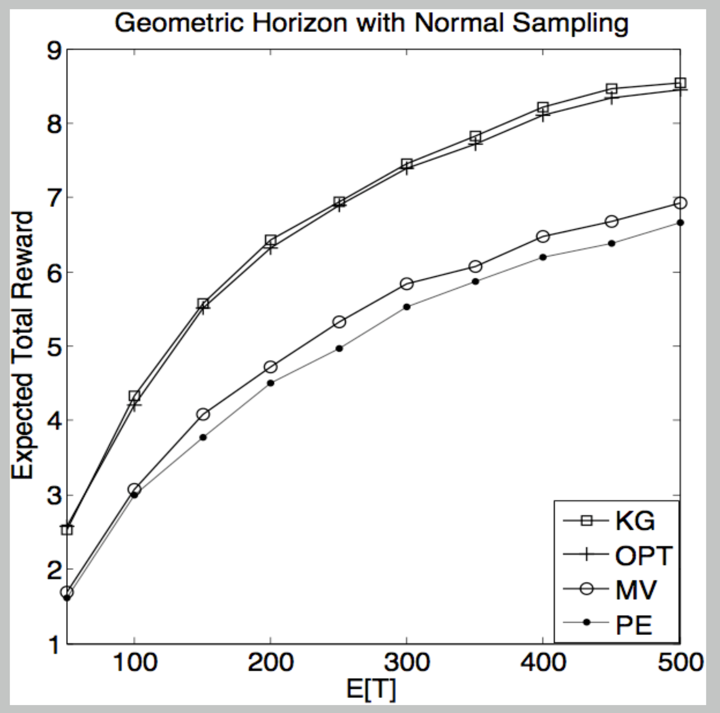
Performance in Infinite Horizon with Sampling Costs





Performance in Geometric Horizon without Costs

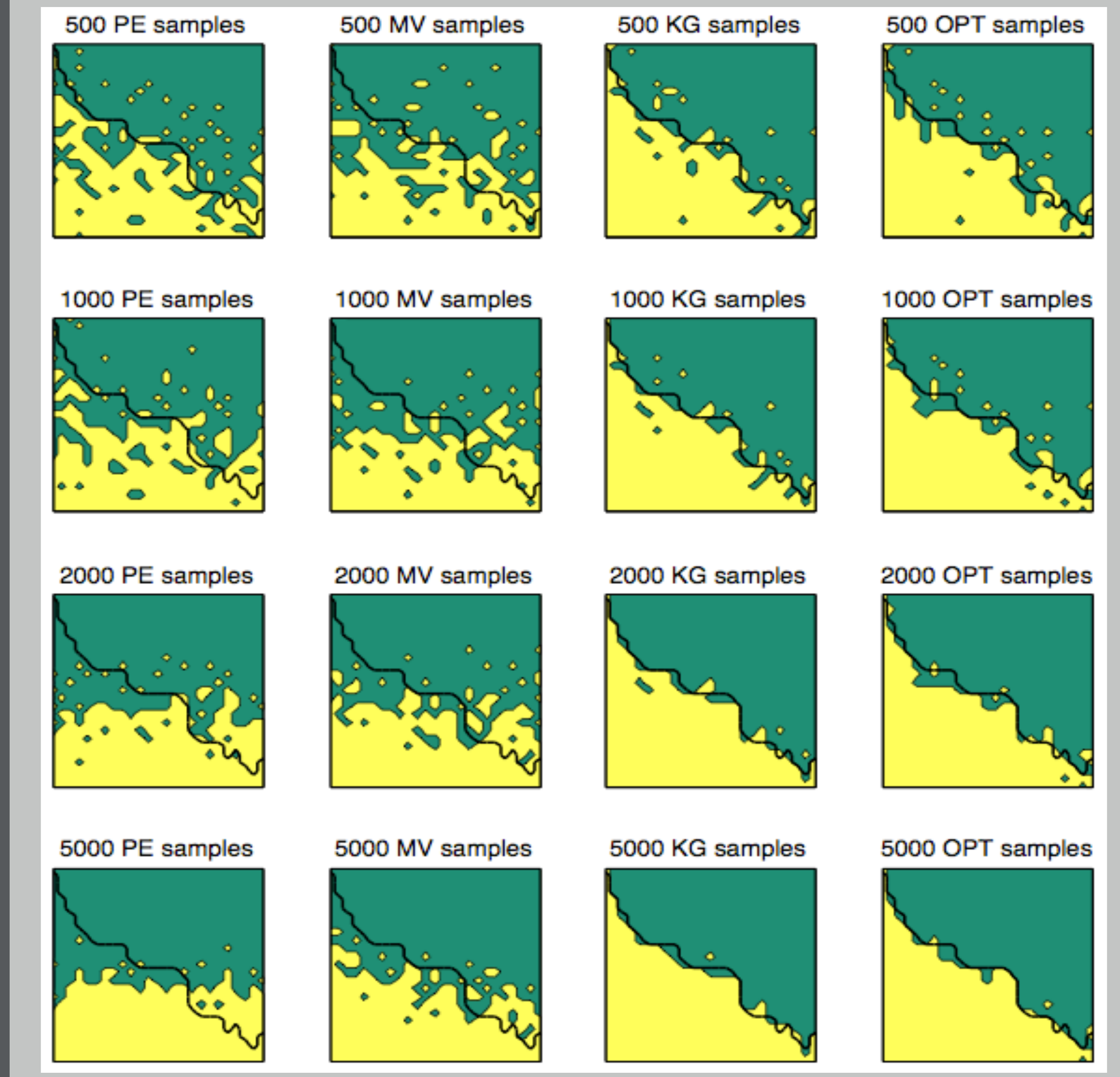




- Pure Exploration (PE): $x_n \sim \mathrm{Uniform}(1,\ldots,k)$.
- Max Variance (MV): $x_{n+1} \in \mathrm{argmax}_x\{\sigma_{nx}\}$.
- Knowledge Gradient (KG): $x_{n+1} \in \mathrm{argmax}_x\{\mathcal{R}_x(S_{nx})\}$.

Ambulance Quality of Service Application

- Administrators of a city's emergency medical services would like to know which of several methods under consideration for positioning ambulances satisfy the minimum requirement of **70% of calls answered on time**.
- The **ambulance allocation plans** are distributed along the x-axis and the **call arrival rates** are distributed along the y-axis. A pair like this is considered an *alternative* and there are 625 alternatives. We assume a *normal sampling distribution* and a *geometric horizon with no sampling costs*.



The black curves are the boundary between $\mathbb{B}$ (the set of qualified alternatives) and its complement.

The yellow regions are the estimates B_n under the corresponding sampling policy, given the stated number of samples.

Conclusions

- We provide **new tools** for simulation analysts facing MCC problems. These new tools
  - dramatically improve efficiency over naive sampling methods;
  - make it possible to efficiently and accurately solve previously intractable MCC problems.
- **Other applications** include determining through simulation under which conditions the current policies of a logistics company are sufficient to maintain quality of service, and finding which projects have a positive net expected value.

jx66@cornell.edu, pf98@cornell.edu