Sequential Bayes-Optimal Policies for Multiple Comparisons with a Control

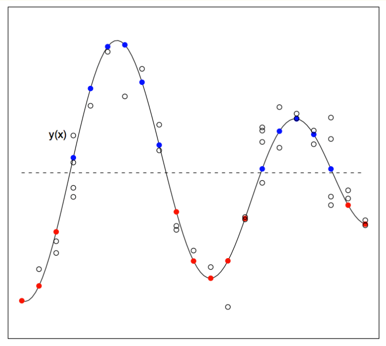
Jing Xie
Peter I. Frazier

School of Operations Research & Information Engineering, Cornell University

Dec 11, 2011
2011 Winter Simulation Conference
Phoenix, Arizona

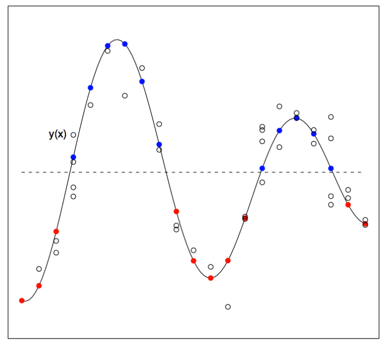
What is Multiple Comparisons with a Control?

- We have a stochastic simulator for k alternative systems.
- Given input $x \in \{1, \ldots, k\}$, it generates random samples from a static distribution with mean μ_x.



What is Multiple Comparisons with a Control?

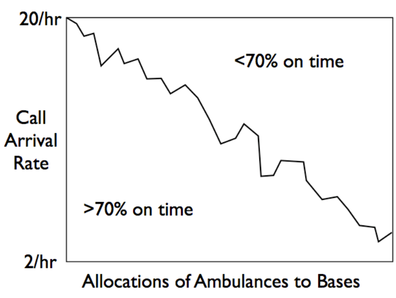
- We have a stochastic simulator for k alternative systems.
- Given input $x \in \{1, \ldots, k\}$, it generates random samples from a static distribution with mean μ_x.
- What is the level set $\mathbb{B} = \{x : \mu_x \geq 0\}$?

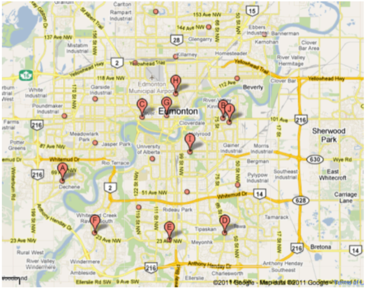


Multiple Comparisons with a Control (MCC):
"Use simulation to determine which alternatives under consideration have mean performance exceeding a known threshold."

MCC appears in Ambulance Positioning

- The city of Edmonton is considering methods for allocating their fleet of 16 ambulances across the 11 bases in the city.
- Which static allocations satisfy the mandated minimums for percentage of emergency calls answered on time?



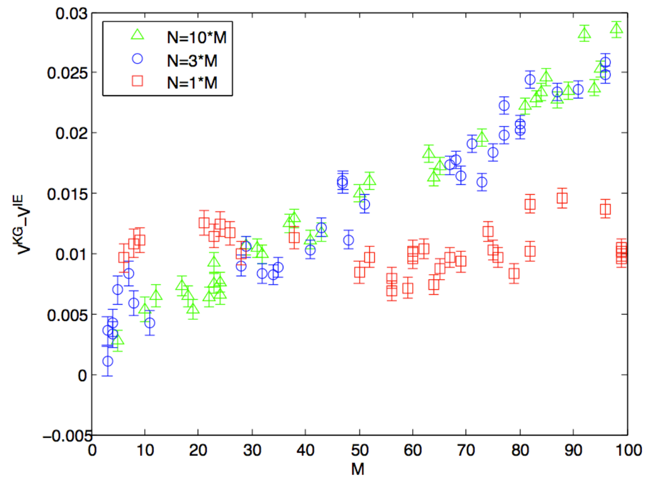


[Thanks to Shane Henderson and Matt Maxwell for providing the ambulance simulation]

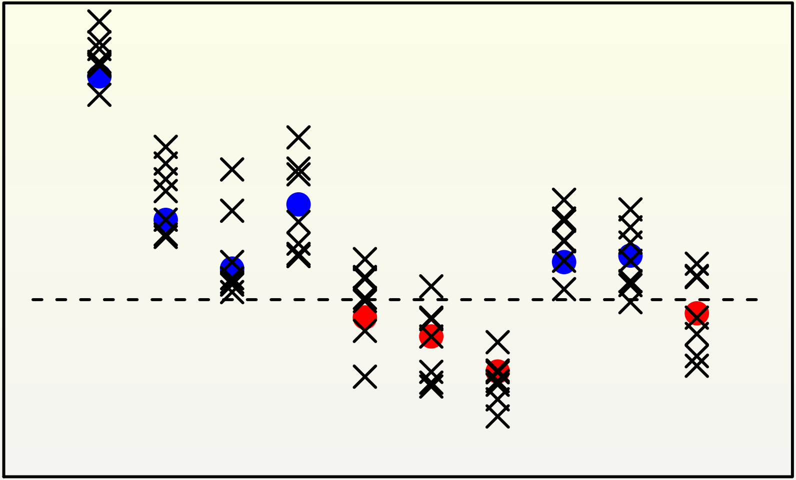
MCC appears in Algorithm Development

A researcher who develops a new algorithm would like to know:

- In which problem settings is average-case performance better with Algorithm A than with Algorithm B?



It seems straightforward to estimate the Level Set...



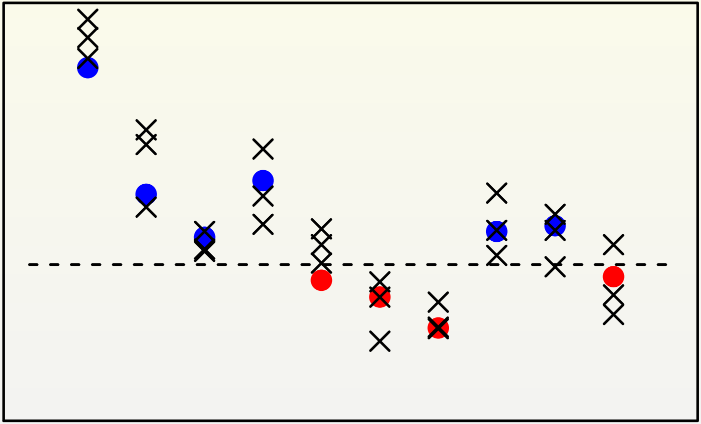
But What If...

- Simulation is expensive?
- Our sampling budget is limited?

But What If...

- Simulation is expensive?
- Our sampling budget is limited?

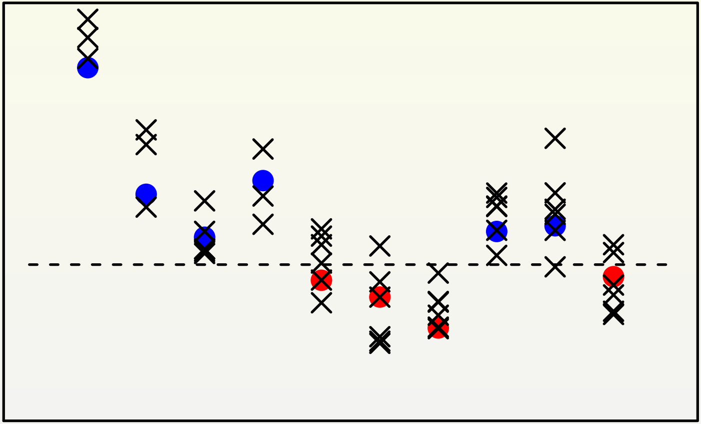
We need to sample more efficiently!



But What If...

- Simulation is expensive?
- Our sampling budget is limited?

We need to sample more efficiently!



How to Allocate Sampling Effort Efficiently?

IDEA: put samples where they help most!

How to Allocate Sampling Effort Efficiently?

IDEA: put samples where they help most!

We design a sampling strategy that provides OPTIMAL average-case performance.

- Using a Bayesian formulation of the MCC problem that explicitly models a limited ability to sample,
- The optimal fully sequential policies can then be solved efficiently and explicitly by a dynamic program.

Formulation of the Bayesian MCC Problem

- We have alternatives $x = 1, \dots, k$.
- Samples from alternative x are $\mathrm{Normal}(\mu_x, \sigma_x^2)$, where μ_x is unknown, while σ_x^2 is assumed known (can be relaxed).
- We have an independent normal Bayesian prior on each μ_x.

Formulation of the Bayesian MCC Problem

- We have alternatives $x = 1, \ldots, k$.
- Samples from alternative x are $\mathrm{Normal}(\mu_x, \sigma_x^2)$, where μ_x is unknown, while σ_x^2 is assumed known (can be relaxed).
- We have an independent normal Bayesian prior on each μ_x.
- We keep sampling until an external deadline (assumed unknown and geometrically distributed) requires us to stop.

Formulation of the Bayesian MCC Problem

- We have alternatives $x = 1, \ldots, k$.
- Samples from alternative x are $\mathrm{Normal}(\mu_x, \sigma_x^2)$, where μ_x is unknown, while σ_x^2 is assumed known (can be relaxed).
- We have an independent normal Bayesian prior on each μ_x.
- We keep sampling until an external deadline (assumed unknown and geometrically distributed) requires us to stop.
- When sampling stops, we estimate the level set $\mathbb{B} = \{x : \mu_x \geq 0\}$ based on the samples and receive a reward R equal to *the number of alternatives correctly classified*.

What Defines an Optimal Policy?

- A policy π is an adaptive rule for choosing where to sample.
- $\mathbb{E}^{\pi}[R|\vec{\mu}]$ is the performance (expected reward) under policy π and true mean vector $\vec{\mu}$.
- $\mathbb{E}^{\pi}[R] = \int \mathbb{E}^{\pi}[R|\vec{\mu}]P(d\vec{\mu})$ is the Bayes- or average-case performance.

What Defines an Optimal Policy?

- A policy π is an adaptive rule for choosing where to sample.
- $\mathbb{E}^{\pi}[R|\vec{\mu}]$ is the performance (expected reward) under policy π and true mean vector $\vec{\mu}$.
- $\mathbb{E}^{\pi}[R] = \int \mathbb{E}^{\pi}[R|\vec{\mu}]P(d\vec{\mu})$ is the Bayes- or average-case performance.

- We wish to find the policy that attains $\sup_{\pi} \mathbb{E}^{\pi}[R]$.
- The solution is characterized theoretically via dynamic programming.
- The curse of dimensionality often makes its computation intractable.

Rewrite this Problem as a Bandit Problem

We decompose the expected reward into an infinite sum of discounted expected one-step rewards

$$\mathbb{E}^{\pi}[R] = R_0 + \mathbb{E}^{\pi}\left[\sum_{n=1}^{\infty} \alpha^{n-1} R_n\right].$$

Here,

- α is the parameter of the geometric distribution of the deadline.
- R_0 is the expected reward if we stop after taking no samples.
- R_n is the expected one-step improvement, due to sampling, of the probability of correctly classifying the alternative sampled.

This is a multi-armed bandit problem.

We Can Now Compute the Optimal Policy

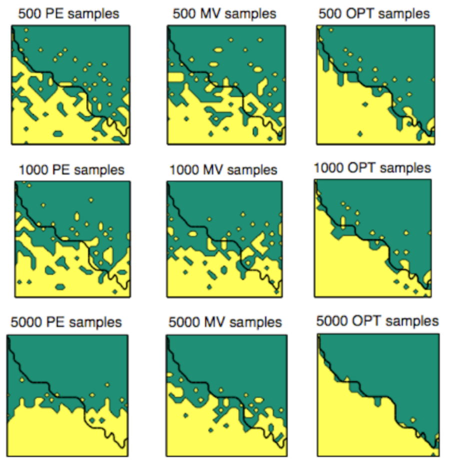
Gittins & Jones 1974 shows that the optimal solution is $\arg\max_x \nu_x(S_{nx})$:

- S_{nx} is a parameterization of the Bayesian posterior on μ_x.
- $\nu_x(\cdot)$ is the Gittins index defined in terms of a single-armed sub-problem:

$$\nu_x(s) = \sup_{\tau>0} \mathbb{E}\left[\frac{\sum_{n=1}^{\tau} \alpha^{n-1} R_n}{\sum_{n=1}^{\tau} \alpha^{n-1}} \,\middle|\, S_{0x} = s \right].$$

We can compute these Gittins indices efficiently because the single-armed sub-problems are much smaller than the full DP.

Results of the Ambulance Service Application



PE = pure exploration (sample at random), MV = max variance (equal allocation), OPT = optimal.

- X-axis: 25 allocations of ambulances;
- Y-axis: 25 arrival rates of emergency calls;

- Black curves: true boundary of the level set $\mathbb{B}$ and its complement;
- Yellow regions: estimates of $\mathbb{B}$ given the stated number of samples under the stated policy.

Generalizations & Conclusions

Our optimality results also allow:

- Other exponential families of sampling distributions, e.g., Bernoulli, normal with unknown variance, Poisson, multinomial, etc;
- Sampling costs together with the option to stop early;
- Different thresholds for each alternative; Other loss functions.

Generalizations & Conclusions

Our optimality results also allow:

- Other exponential families of sampling distributions, e.g., Bernoulli, normal with unknown variance, Poisson, multinomial, etc;
- Sampling costs together with the option to stop early;
- Different thresholds for each alternative; Other loss functions.

We provide new tools for simulation analysts facing MCC problems, which

- can be used in many different simulation applications;
- have a general ability to compute the optimal sampling schemes;
- characterize the level sets with dramatically less simulation effort.

THANK YOU!