Ranking and Selection With Tight Bounds on Probability of Correct Selection

Peter I. Frazier

Operations Research & Information Engineering, Cornell University

Thursday March 21, 2013
Simulation Optimization Workshop
Viña del Mar, Chile

Supported by AFOSR YIP FA9550-11-1-0083 and
NSF CAREER CMMI-1254298

Ranking and Selection with Independent Normal Samples

- We have k alternative systems that can be simulated.
  - e.g., different methods for operating a supply chain.
  - different pricing mechanisms for airline tickets.
  - different inspection policies for shipping containers entering a port.
- Each time we simulate alternative x, we observe

$$y \sim \mathrm{Normal}(\theta_x, \sigma_x^2) \qquad \text{(independent across } x \text{ and time)}$$

  where θ_x is unknown.
  - Initially, let's assume σ_x^2 is known. We will relax this later.
  - Approximate normality can be checked empirically, and if it is not met, samples can be batched together.
- **Goal**: Use simulation efficiently to find $\arg\max_x \theta_x$.

We would like a worst-case guarantee on PCS (Probability of Correct Selection)

- A **policy** is a (possibly adaptive) rule for deciding how many samples to take from each alternative, and for selecting an estimate of $\arg\max_x \theta_x$ based on these samples.
- Given a set of true sampling means $\theta = (\theta_1, \ldots, \theta_k)$, and a policy π,

$$\mathrm{PCS}(\pi, \theta) \qquad \textbf{(Probability of Correct Selection)}$$

  is the probability that π selects an alternative in $\arg\max_x \theta_x$.
- In a perfect world, given some P^*, we would have a procedure satisfying

$$\mathrm{PCS}(\pi, \theta) \geq P^* \quad \text{for all } \theta \in \mathbb{R}^k.$$

- This turns out to be too restrictive.

We will settle for an Indifference-Zone (IZ) Guarantee

Instead, we seek a guarantee of the following form:

- Define the **preference zone** as

$$\mathrm{PZ}(\delta) = \left\{\theta \in \mathbb{R}^k : \theta_{[1]} - \theta_{[2]} \geq \delta\right\},$$

  where $\theta_{[1]} \geq \theta_{[2]} \geq \ldots \geq \theta_{[k]}$, and $\delta > 0$ is fixed.
- The **indifference zone (IZ)** is the complement: $\mathbb{R}^k \setminus \mathrm{PZ}(\delta)$.
- We say a policy π has an **IZ guarantee** with parameters δ and P^* if

$$\mathrm{PCS}(\pi, \theta) \geq P^* \quad \text{for all } \theta \in \mathrm{PZ}(\delta).$$

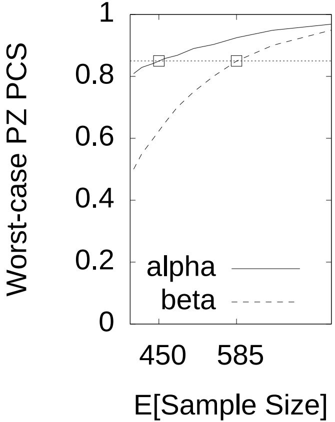
Goal: Find a policy that satisfies the IZ guarantee, while taking as few samples as possible.

There are Many Procedures with an IZ Guarantee

Lots of previous work has constructed policies that satisfy the IZ guarantee.

- Fixed sample size policies: [Bechhofer, 1954]
- Two-stage policies: [Dudewicz and Dalal, 1975, Rinott, 1978]
- Fully sequential policies [Paulson, 1964, Bechhofer and Goldsman, 1987, Hartmann, 1988, Hartmann, 1991, Paulson, 1994, Kim and Nelson, 2001, Nelson et al., 2001, Hong, 2006, Bechhofer et al., 1968]

Loose PCS bounds cause the the number of samples taken to be larger than necessary



- Consider a sampling procedure with a tunable parameter h, that controls how much sampling is done.
- If we increase h, then $\pi(h)$ takes more samples, and provides a bigger PCS.
- Let $\alpha(h) = \inf_{\theta \in \mathrm{PZ}(\delta)} \mathrm{PCS}(\pi(h), \theta)$ be the actual worst-case preference-zone PCS. (solid line, from exhaustive simulation)
- Let $\beta(h) \leq \alpha(h)$ be the best bound on $\alpha(h)$ that we can prove. (dashed line)
- We choose h so that $\beta(h) = P^*$.
- Had we known α, we could have instead chosen a smaller h with $\alpha(h) = P^*$, and taken fewer samples.

Contribution 1: Tight PCS Bounds

We construct a fully sequential elimination policy, called the **Bayes-Inspired IZ (BIZ)** policy, which has two properties:

1. BIZ satisfies the IZ guarantee:

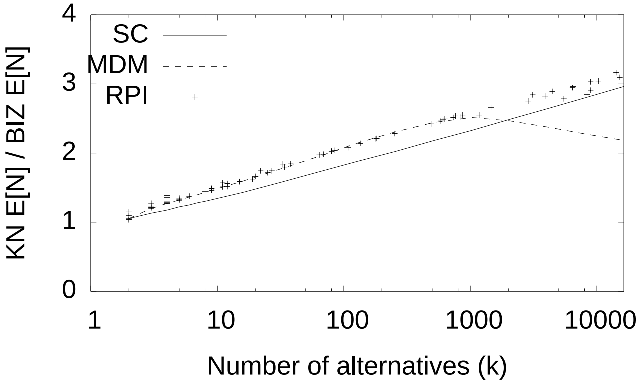
$$\mathrm{PCS}(\mathrm{BIZ}, \theta) \geq P^*, \quad \text{for all } \theta \in \mathrm{PZ}(\delta).$$

2. In continuous time, the lower bound P^* on PCS is tight:

$$\inf_{\theta \in \mathrm{PZ}(\delta)} \mathrm{PCS}(\mathrm{BIZ}, \theta) = P^*$$

Caveat: These results require strong assumptions on the variances: that they are known, and are either common across alternatives, or are rational multiples of a common value.

Contribution 2: BIZ Requires Fewer Samples



- When the number of alternatives is large, BIZ samples much less than existing IZ policies.
- For many of the largest problems, BIZ requires between 2 and 3 times fewer samples than KN [Kim and Nelson, 2001].

First we present results for common known variance

- For the next few slides, we assume the sampling variances are known, and all the same.
- Later, we return to unknown and/or heterogeneous sampling variances.

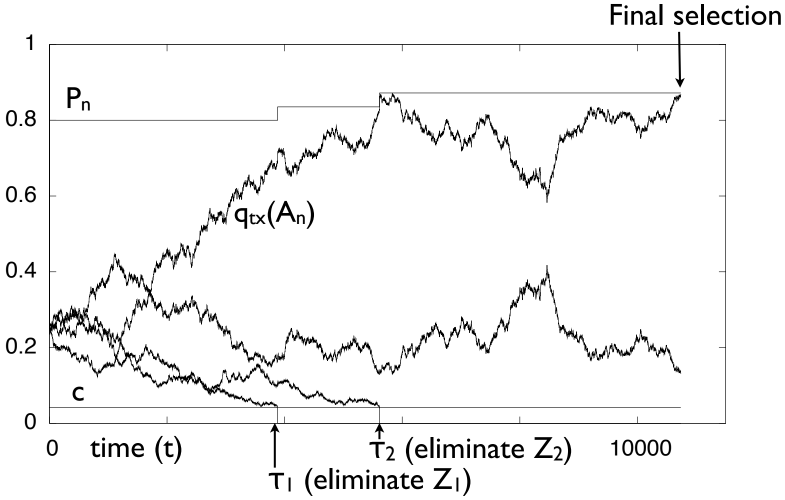
We Construct BIZ with Bayesian Ideas

- Let Y_{tx} be the sum of all observations from alternative x by time t.
- For $A \subseteq \{1, \ldots, k\}$, let

$$q_{tx}(A) = \exp\left(\frac{\delta}{\sigma^2} Y_{tx}\right) \Big/ \sum_{x' \in A} \exp\left(\frac{\delta}{\sigma^2} Y_{tx'}\right).$$

- Under a Bayesian prior distribution that is concentrated on slippage configurations, $q_{tx}(A)$ is the posterior probability that $x = X^*$, given that $X^* \in A$. Here, X^* is the alternative with the largest sampling mean.
- Although we construct BIZ with Bayesian ideas, and manipulate Bayesian PCS in its analysis, it is a non-Bayesian method: You do not need to have a prior to use BIZ, and its IZ guarantee is non-Bayesian.

The BIZ Procedure



The BIZ Procedure

Fix parameters $c \leq 1-(P^*)^{1/(k-1)}$, $\delta > 0$, $P^* > 1/k$.

1. Let $A \leftarrow \{1,\ldots,k\}$, $t \leftarrow 0$, $P \leftarrow P^*$.
2. While $\max_{x\in A} q_{tx}(A) < P$
   - 2a. While $\min_{x\in A} q_{tx}(A) \leq c$
     - Let $x \in \arg\min_x q_{tx}(A)$.
     - Let $P \leftarrow P/(1-q_{tx}(A))$.
     - Remove x from A.
   - 2b. Sample from each $x \in A$ to obtain $Y_{t+1,x}$. Then increment t.
3. Select $\hat{x} \in \arg\max_{x\in A} Y_{tx}$ as our estimate of the best.

Recall:

$$q_{tx}(A) = \exp\left(\frac{\delta}{\sigma^2} Y_{tx}\right) \Big/ \sum_{x'\in A} \exp\left(\frac{\delta}{\sigma^2} Y_{tx'}\right).$$

Generalization of BIZ, which allows sampling in discrete or continuous time

- **Parameters**: $\mathbb{T} \in \{\mathbb{R}_+, \mathbb{Z}_+\}$, $c \leq 1 - (P^*)^{1/(k-1)}$, $\delta > 0$, $P^* > 1/k$.
- **Initalization**: $\tau_0 = 0$, $A_0 = \{1, \dots, k\}$, $P_0 = P^*$.
- **Elimination Time**:

$$\tau_{n+1} = \inf \left\{ t \in \mathbb{T} \cap [\tau_n, \infty) : \min_{x \in A_n} q_{tx}(A_n) \leq c \text{ or } \max_{x \in A_n} q_{tx}(A_n) \geq P_n \right\}.$$

- **Eliminated Alternative and Contention Set**:

$$Z_{n+1} = \arg\min_{x \in A_n} q_{\tau_{n+1}}(A_n), \qquad A_{n+1} = A_n \setminus Z_{n+1}$$

- **Stopping Boundary**:

$$P_{n+1} = P_n \Big/ \left(1 - \min_{x \in A_n} q_{\tau_{n+1}}(A_n) \right).$$

- The **selected alternative** is the single alternative in A_{k-1}.

Tight PCS Bounds

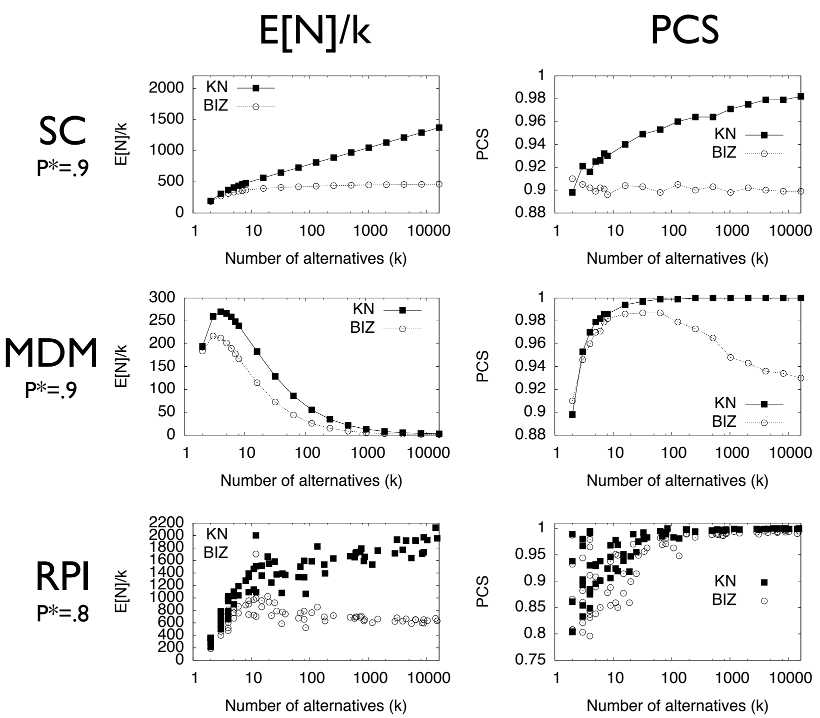
Theorem

Assume that $\sigma_x^2 = \sigma^2$ *is known. Fix any* $\delta > 0$, $P^* \in (1/k, 1)$, $c \leq 1 - (P^*)^{1/(k-1)}$, $\mathbb{T} \in \{\mathbb{R}_+, \mathbb{Z}_+\}$, *and let* π *be the corresponding BIZ policy. Then,*

$$\mathrm{PCS}(\pi, \theta) \geq P^* \quad \forall \theta \in \mathrm{PZ}(\delta)$$

Moreover, if $\mathbb{T} = \mathbb{R}_+$ *(i.e., if BIZ operates in continuous time),*

$$\inf_{\theta \in \mathrm{PZ}(\delta)} \mathrm{PCS}(\pi, \theta) = P^*$$



E[N]/k
PCS
SC
P*=.9
MDM
P*=.9
RPI
P*=.8
KN
BIZ
E[N]/k
PCS
Number of alternatives (k)

Heterogeneous Sampling Variance: Continuous Time

- In general, sampling variances σ_x^2 are heterogeneous and unknown.
- In continuous time, this problem is easily addressed:
  1. The sampling variance σ_x^2 can be estimated perfectly given $(Y_{tx} : 0 \le t \le \varepsilon)$ for any $\varepsilon > 0$.
  2. Replace $Y_{t,x} \sim \mathcal{N}(\theta_x, t\sigma_x^2)$ with $Y_{\sigma_x^2 t,x}/\sigma_x^2 \sim \mathcal{N}(\theta_x, t)$ and we obtain a ranking and selection problem with common sampling variance 1.
  3. Use BIZ for common sampling variance 1 on the transformed Y values.
- The IZ guarantee and the tightness of the PCS bound still hold.

Heterogeneous and Unknown Sampling Variance: Discrete Time

- In discrete time, the problem is harder:
- Idea: use a discrete-time approximation to $Y_{\sigma_x^2 t,x}/\sigma_x^2$ in our calculation of $q_{tx}(A)$, based on the sample variances $\widehat{\sigma}_x^2$.
- When this discrete-time approximation is exact, we retain the IZ guarantee.
- When it does not hold, the IZ guarantee no longer holds, and the procedure is a heuristic.
- Ongoing work: Does the IZ guarantee hold in a limiting sense?

Heterogeneous and Unknown Sampling Variance: Discrete Time

- Let n_{tx} be the number of samples taken from alternative x by time t.
- We sample so as to keep n_{tx} proportional to $\widehat{\sigma}_x^2$ (as much as allowed by discrete time).
- Most obvious approach: Approximate $Y_{\sigma_x^2 t,x}/\sigma_x^2$ with $Y_{n_{tx},x}/\widehat{\sigma}_x^2$.
- A better approach: approximate it with

$$\frac{\sum_{x' \in A} n_{tx'}}{\sum_{x' \in A} \widehat{\sigma}_{tx'}^2} \frac{Y_{n_{tx},x}}{n_{tx}}$$

  This allows the alternative with the highest $q_{tx}(A)$ to also be the one with the highest sample mean, $Y_{n_{tx},x}/n_{tx}$.
- When n_{tx} is exactly proportional to σ_x^2 and $\widehat{\sigma}_x^2 = \sigma_x^2$, this reduces to what we had in continuous time, and the IZ guarantee still holds.

Fix $c \le 1-(P^*)^{1/(k-1)}$, $\delta > 0$, $P^* > 1/k$, $B_1,\ldots,B_k > 0$, n_0 and let

$$\hat{q}_{tx}(A) = \exp\left(\delta \frac{\sum_{x''\in A} n_{tx''}}{\sum_{x''\in A} \widehat{\sigma}^2_{tx''}} \frac{Y_{n_{tx},x}}{n_{tx}}\right) \Bigg/ \sum_{x'\in A} \exp\left(\delta \frac{\sum_{x''\in A} n_{tx''}}{\sum_{x''\in A} \widehat{\sigma}^2_{tx''}} \frac{Y_{n_{tx'},x'}}{n_{tx'}}\right).$$

1. Sample each alternative n_0 times. For each x, let $n_{0x} \leftarrow n_0$, and let Y_{0x} and $\widehat{\sigma}^2_{0x}$ be the sample mean and sample variance respectively. Also let $A \leftarrow \{1,\ldots,k\}$, $P \leftarrow P^*$, $t \leftarrow 1$.
2. While $\max_{x\in A} \hat{q}_{tx}(A) < P$
   - 2a. While $\min_{x\in A} \hat{q}_{tx}(A) \le c$
     - Let $x \in \arg\min_{x\in A} \hat{q}_{tx}(A)$.
     - Let $P \leftarrow P/(1-\hat{q}_{tx}(A))$.
     - Remove x from A.
   - 2b. Let $z \in \arg\min_{x\in A} n_{tx}/\widehat{\sigma}^2_{tx}$.
   - 2c. For each $x \in A$, let $n_{t+1,x} = \mathrm{ceil}\left(\widehat{\sigma}^2_{tx}(n_{tz}+B_z)/\widehat{\sigma}^2_{tz}\right)$.
   - 2d. For each $x \in A$, if $n_{t+1,x} > n_{tx}$, take $n_{t+1,x} - n_{tx}$ additional samples from alternative x. Let $Y_{t+1,x}$ and $\widehat{\sigma}^2_{t+1,x}$ be the sample mean and sample variance respectively of all samples from alternative x thus far. Then increment t.
3. Select $\hat{x} \in \arg\max_{x\in A} Y_{tx}/n_{tx}$ as our estimate of the best.

Tight PCS Bounds: Known Heterogeneous Variance

Theorem

Assume the σ_x^2 are known. Also, if $\mathbb{T} = \mathbb{Z}_+$, assume $\sigma_x^2 = B_x\sigma^2$ for some $\sigma^2 \in \mathbb{R}$ and $B_x \in \mathbb{Z}_+$.
Fix any $\delta > 0$, $P^* \in (1/k, 1)$, $c \leq 1 - (P^*)^{1/(k-1)}$, $\mathbb{T} \in \{\mathbb{R}_+, \mathbb{Z}_+\}$. Let π be the corresponding BIZ policy.
Then,

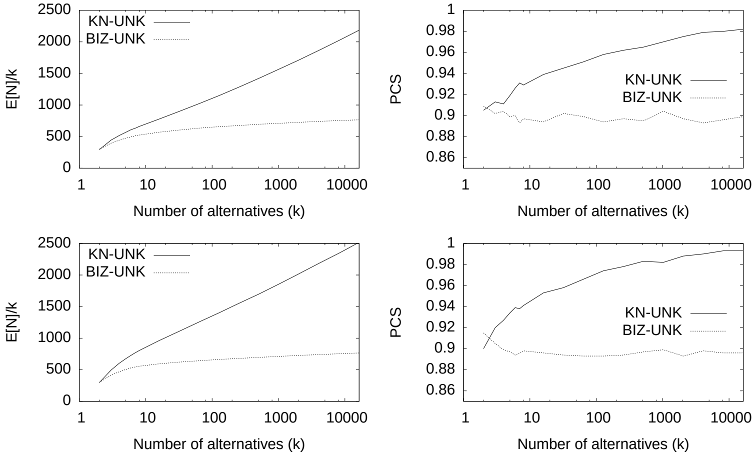
$$\mathrm{PCS}(\pi, \theta) \geq P^* \quad \forall \theta \in \mathrm{PZ}(\delta)$$

Moreover, if $\mathbb{T} = \mathbb{R}_+$ (i.e., if BIZ operates in continuous time),

$$\inf_{\theta \in \mathrm{PZ}(\delta)} \mathrm{PCS}(\pi, \theta) = P^*$$

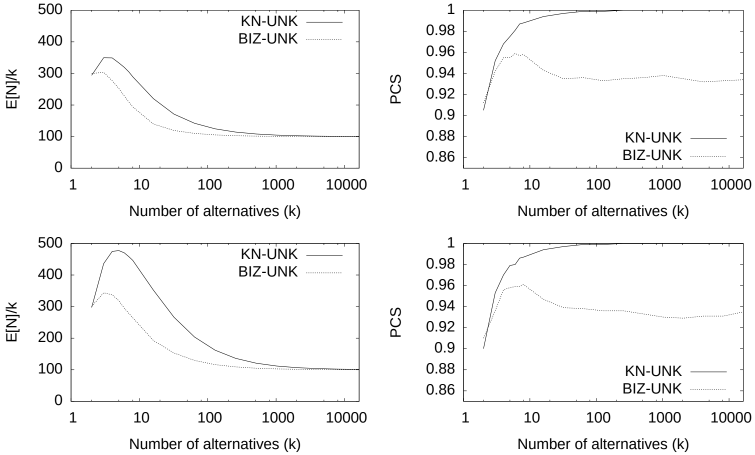
Unknown Heterogeneous Variance

Slippage configuration.
Top row: increasing variance. Bottom row: decreasing variance.
Left column: $E[N]/k$. Right column: PCS.



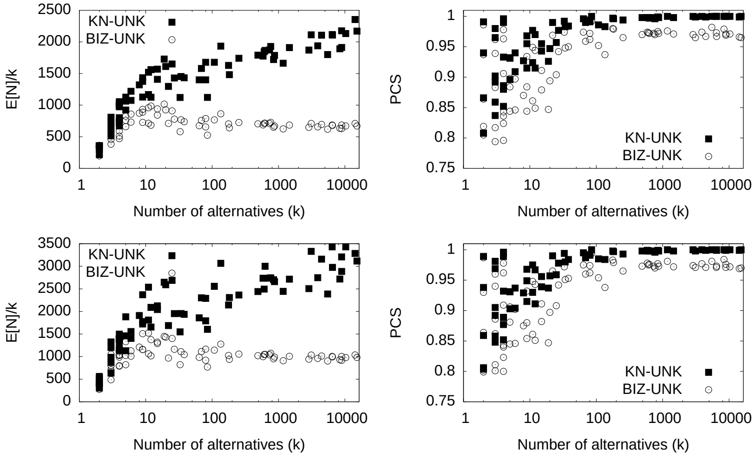
Unknown Heterogeneous Variance

Monotone Decreasing Means configuration.
Top row: increasing variance. Bottom row: decreasing variance.
Left column: $E[N]/k$. Right column: PCS.



Unknown Heterogeneous Variance

Random Problem Instances.
Top row: common variance. Bottom row: heterogeneous variance.
Left column: $E[N]/k$. Right column: PCS.



Conclusion

- BIZ is a fully sequential IZ procedure with elimination that delivers **exactly** P^* (in continuous time, and under the worst θ in the preference zone).
- To my knowledge, this is the **first fully sequential elimination IZ policy with this property** for $k > 2$.
- Theoretical results require unrealistic assumptions on the sampling variances, but empirical results suggest that behavior is robust to violations of these assumptions in the problem regimes tested.

Thank You!

References I

Bechhofer, R. (1954).
A single-sample multiple decision procedure for ranking means of normal populations with known variances.
The Annals of Mathematical Statistics, 25(1):16–39.

Bechhofer, R. and Goldsman, D. (1987).
Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the normal population which has the largest mean.
Communications in Statistics-Simulation and Computation, 16(4):1067–1092.

Bechhofer, R., Kiefer, J., and Sobel, M. (1968).
Sequential Identification and Ranking Procedures.
University of Chicago Press, Chicago.

Dudewicz, E. and Dalal, S. (1975).
Allocation of observations in ranking and selection with unequal variances.
Sankhyā: The Indian Journal of Statistics, Series B, 37(1):28–78.

Hartmann (1991).
An improvement on Paulson's procedure for selecting the population with the largest mean from k normal populations with a common unknown variance.
Sequential Analysis, 10(1-2):1–16.

Hartmann, M. (1988).
An improvement on Paulson's sequential ranking procedure.
Sequential Analysis, 7(4):363–372.

Hong, J. (2006).
Fully sequential indifference-zone selection procedures with variance-dependent sampling.
Naval Research Logistics (NRL), 53(5):464–476.

References II

Kim, S.-H. and Nelson, B. L. (2001).
A fully sequential procedure for indifference-zone selection in simulation.
ACM Trans. Model. Comput. Simul., 11(3):251–273.

Nelson, B., Swann, J., Goldsman, D., and Song, W. (2001).
Simple procedures for selecting the best simulated system when the number of alternatives is large.
Operations Research, 49(6):950–963.

Paulson, E. (1964).
A sequential procedure for selecting the population with the largest mean from k normal populations.
The Annals of Mathematical Statistics, 35(1):174–180.

Paulson, E. (1994).
Sequential procedures for selecting the best one of k koopman-darmois populations.
Sequential Analysis, 13(3).

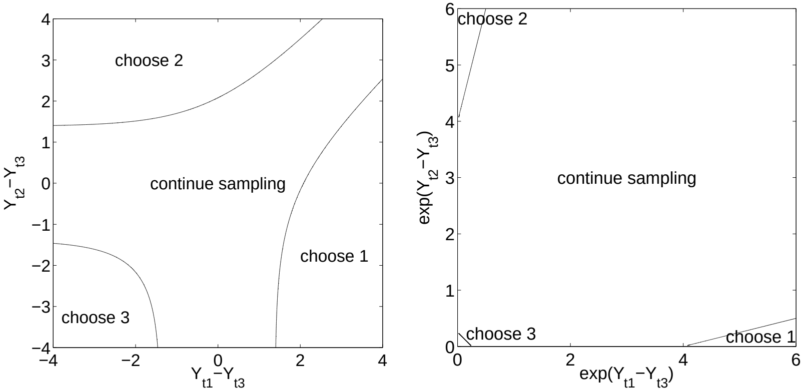
Rinott, Y. (1978).
On two-stage selection procedures and related probability-inequalities.
Communications in Statistics-Theory and Methods, 7(8):799–811.

Heterogeneous Sampling Variance

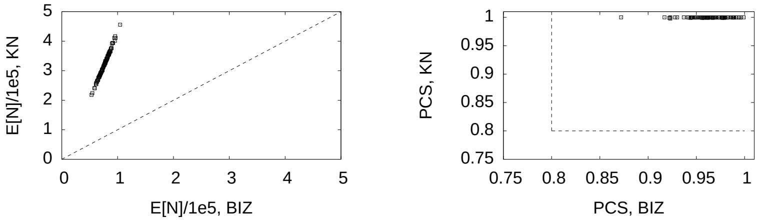
- In continuous time, BIZ can be extended easily to the case where sampling variances are heterogeneous, and the main theoretical results are still true (IZ guarantee with a tight bound).
- In discrete time, the same technique can be applied, but the IZ guarantee no longer holds exactly. Ongoing work: does it hold in a limiting sense?

Continuation Region

- Images show continuation region for $k = 3$,
  in linear coordinates (left) and exponential coordinates (right).
- BKS (BIZ with $c = 0$) stops when Y_t exits the continuation region.



Numerical Comparisons: Random Problem Instances



- $k = 100$.
- $P^* = 0.8$.
- θ was generated randomly from an independent normal prior.
- It was then adjusted so that no alternative other than the best is within δ of the best, i.e., so that $\theta \in \mathrm{PZ}(\delta)$.

BIZ Construction: Elimination

Initialization: $(\tau_0 = 0,\ A_0 = \{1, 2, 3, 4\})$

t=1 $\quad$ 1 2 3 4

t=2 $\quad$ 1 2 3 4

t=3 $\quad$ 1 2 3 4 $(\tau_1 = 3,\ Z_1 = 2,\ A_1 = \{1, 3, 4\})$

t=4 $\quad$ 1 $\cancel{2}$ 3 4

t=5 $\quad$ 1 $\cancel{2}$ 3 4 $(\tau_2 = 5,\ Z_2 = 4,\ A_2 = \{1, 3\})$

t=6 $\quad$ 1 $\cancel{2}$ 3 $\cancel{4}$

t=7 $\quad$ 1 $\cancel{2}$ 3 $\cancel{4}$ $(\tau_3 = 7,\ Z_3 = 1,\ A_3 = \{3\})$

Selection: Select alternative 3 as the best.

Proof Sketch

Let CS be the event of correct selection.
For $\theta \in \mathbb{R}^d$, let Q_θ be a prior that is uniform on the permutations θ. In particular, $Q = Q_{[\delta,0,\dots,0]}$.

Lemma (Symmetry)

$\mathrm{PCS}(\pi,\theta)$ *is invariant to permutations of* θ.
Moreover, $\mathrm{PCS}(\pi,\theta) = Q^\pi_\theta\{\mathrm{CS}\}$.

Lemma (Monotonicity)

For $\theta \in \mathrm{PZ}(\delta)$, $Q^\pi_\theta\{CS\} \geq Q^\pi_{[\delta,0,\dots,0]}\{CS\}$.

Lemma (Bayes PCS of Least-favorable Configuration)

$Q^\pi_{[\delta,0,\dots,0]}\{CS\} \geq P^*$, *with equality if* $\mathbb{T} = \mathbb{R}_+$.

BIZ Construction: Elimination

BIZ is an **elimination** procedure.

- It defines a sequence of stopping times, $0 = \tau_0 \leq \tau_1 \cdots \leq \tau_{k-1} < \infty$.
- For $n < k$,
  - τ_n is the time that the n^{th} alternative is eliminated.
  - $Z_n \in \arg\min_{x \in A_{n-1}} Y_{\tau_n,x}$ is the n^{th} alternative eliminated.
  - $A_n = \{1, \dots, k\} \setminus \{Z_m : m \leq n\}$ are the remaining alternatives.
- At time τ_{k-1}, we stop sampling and select the single remaining alternative as best.

Elimination allows us to quickly eliminate very bad alternatives, reducing sampling effort.

Continuous-Time Observation Process

- In the original problem, we observe independent $\mathcal{N}(\theta_x, \sigma_x^2)$ values from alternative x.
  - The sum of all observations up to the current time is a random walk.
- In our continuous-time generalization, we let $(Y_{tx} : t \in \mathbb{R}_+)$ be a Brownian motion with drift θ_x and volatility $\sigma_x = \sigma$.
- Let $\mathbb{T} \in \{\mathbb{Z}_+, \mathbb{R}_+\}$. We restrict elimination and stopping decisions to be in $\mathbb{T}$.
- When $\mathbb{T} = \mathbb{Z}_+$, the resulting procedure can be implemented in discrete time.