Sequential Ranking and Selection: Tight Bounds and Large-Scale Problems

Peter I. Frazier

Operations Research & Information Engineering, Cornell University

Thursday July 7, 2011
Applied Probability Society Meeting
KTH, Stockholm

Supported by AFOSR YIP #FA9550-11-1-0083

Ranking and Selection

- We have k "alternatives" or "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

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

  where θ_x is unknown.
- We assume known constant $\sigma_x^2 = \sigma^2$. Generalization to unknown heterogeneous σ_x^2 is discussed briefly at the end of the talk.
- **Goal**: Use simulation efficiently to find $\arg\max_x \theta_x$.

Ranking and Selection Policies

- A **policy** is a rule for deciding
  - Which alternatives to sample at each point in time.
  - When to stop sampling.
  - Which alternative to select as the best when we stop sampling. (Policies usually select the one with the largest sample mean)
- These decisions can be made adaptively.
- Given a system configuration $\theta = (\theta_1, \ldots, \theta_k)$, and a policy π,

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

is the probability that π selects an alternative in $\arg\max_x \theta_x$.
(The dependence on σ_x^2 is suppressed in the notation.)

Preference and Indifference Zones

- The **preference zone** is a collection of system configurations where the best is better than the second best by at least δ :

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

where $\theta_{[1]} \geq \theta_{[2]} \geq \ldots \geq \theta_{[k]}$ are the ordered components of θ, and $\delta > 0$ is a fixed parameter.

- The **indifference zone (IZ)** are those system configurations outside the preference zone.
- 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.**

Indifference Zone Formulation

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]
- The procedure in [Bechhofer et al., 1968] is recovered as a special case of the policy that we introduce.

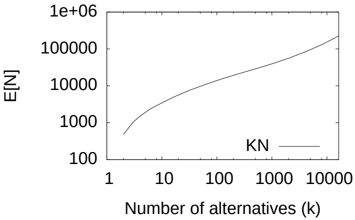
Overdelivery & Too Many Samples

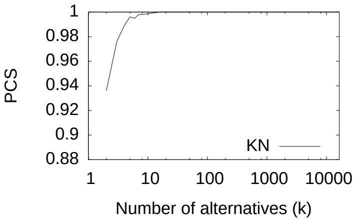
- Let π be a procedure with an IZ guarantee for a fixed P^* and δ.
- P^* is the PCS that π guarantees it will deliver.
- For any $\theta \in \mathrm{PZ}(\delta)$,

  $$\mathrm{PCS}(\pi, \theta) - P^*$$

  is the **overdelivery** on PCS.
- Overdelivery on PCS is inefficient: we could have taken fewer samples and achieved the guaranteed PCS faster.
- For existing policies and large k, this overdelivery causes the number of samples to be **significantly** larger than needed.

Example: Too Many Samples





- This is the **monotone-decreasing-means configuration**, $\theta = [-\delta, -2\delta, \ldots, -k\delta]$.
- Parameters are: $\delta = 1$, $\sigma_x = 10$, $P^* = 0.9$.
- The policy is the KN policy of [Kim and Nelson, 2001], which is state-of-the-art for problems with not much variation in σ_x^2. It has been modified to use a known sampling variance.

Tight Bounds & Fewer Samples

- Given P^* and δ, we construct a fully sequential policy π for which
  1. We prove π has the IZ guarantee:
  $$\mathrm{PCS}(\pi,\theta) \geq P^*, \quad \text{for all } \theta \in \mathrm{PZ}(\delta).$$
  2. We prove the lower bound P^* on PCS is tight in continuous time:
  $$\inf_{\theta \in \mathrm{PZ}(\delta)} \mathrm{PCS}(\pi,\theta) = P^*$$
- When the number of alternatives is large, this policy samples much less than existing IZ policies.
- This policy is inspired by a Bayesian analysis, and we call it the **Bayes-Inspired IZ (BIZ)** policy. (even though the theoretical results apply to non-Bayesian PCS).

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.
  - $A_n = \{1, \ldots, k\} \setminus \{Z_m : m \leq n\}$ 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.

BIZ Construction: Bayesian Prior

- The BIZ policy is inspired by Bayesian ideas.
- Let Q be a prior probability measure on $\mathrm{PZ}(\delta)$ under which

$$X_* \sim \mathrm{Uniform}(1, \ldots, k)$$

$$\theta_x = \begin{cases} \delta & \text{if } x = X_* \\ 0 & \text{if } x \neq X_* \end{cases}$$

- Q is concentrated on **least-favorable configurations**.

BIZ Construction: Bayesian Posterior

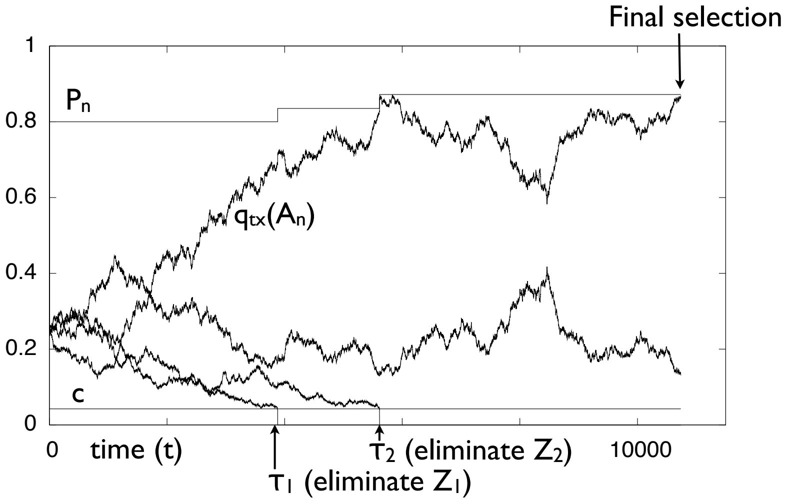
- Recall Y_{tx} is the sum of all observations from alternative x by time t.
- For $A \subseteq \{1,\ldots,k\}$, define

$$q_{tx}(A) = Q\left\{x = X_* \mid X_* \in A, (Y_{tx'})_{x' \in A}\right\}.$$

  This is the **posterior probability that** x **is the best alternative**, given that the best is in the set A.
- One can show for $x \in A$:

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

BIZ



Final selection
1
0.8
0.6
0.4
0.2
0
P_n
$q_{tx}(A_n)$
c
0
time (t)
τ_2 (eliminate Z_2)
10000
τ_1 (eliminate Z_1)

BIZ

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

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

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

Main Result

Theorem

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

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

Moreover, if $\mathbb{T} = \mathbb{R}_+$, then

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

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, $Q = Q_{[\delta,0,\ldots,0]}$.

Lemma (Symmetry)

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

Lemma (Monotonicity)

For $\theta \in \mathrm{PZ}(\delta)$, $Q_\theta^\pi\{CS\} \geq Q_{[\delta,0,\ldots,0]}^\pi\{CS\}$.

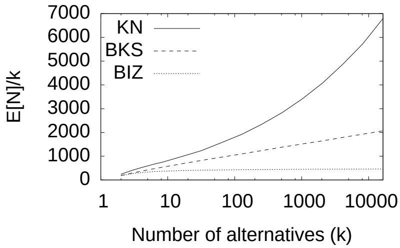
Lemma (Bayes PCS of Least-favorable Configuration)

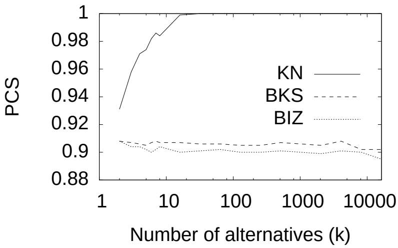
$Q_{[\delta,0,\ldots,0]}^\pi\{CS\} \geq P^*$, *with equality if* $\mathbb{T} = \mathbb{R}_+$.

Recovering [Bechhofer et al., 1968]

- Choosing $c = 0$ recovers the sequential ranking procedure $\mathscr{P}_B^*$ from [Bechhofer et al., 1968].
- In this special case, this policy does not eliminate.
- It was previously unknown that $\mathscr{P}_B^*$ is exact in continuous-time.

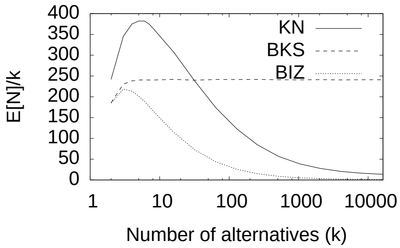
Numerical Comparisons: Slippage Configuration

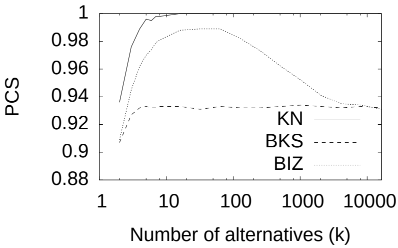




- Settings are the same as before: $\theta = [0, \ldots, 0, \delta]$, $P^* = 0.9$, $\sigma = 10$, $\delta = 1$, estimated with $\geq 10,000$ independent replications.
- BIZ uses $c = 1 - (P^*)^{1/(k-1)}$ (eliminate aggressively).
- BKS is $\mathscr{P}_B^*$ $(c = 0)$ from [Bechhofer et al., 1968].

Numerical Comparisons: Monotone Decreasing Means





- Settings are the same as before: $\theta = [-\delta, -2\delta, \ldots, -k\delta]$, $P^* = 0.9$, $\sigma = 10$, $\delta = 1$, estimated with $\geq 10,000$ independent replications.
- BIZ uses $c = 1 - (P^*)^{1/(k-1)}$ (eliminate aggressively).
- BKS is $\mathscr{P}_B^*$ $(c = 0)$ from [Bechhofer et al., 1968].

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?

Conclusion

- The BIZ policy is a fully sequential policy IZ policy that delivers the target exactly in continuous time.
- To my knowledge, this is the **first fully sequential elimination IZ policy with this property** for $k > 2$.
- This method of Bayesian analysis with least-favorable priors is a general theoretical tool.

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.

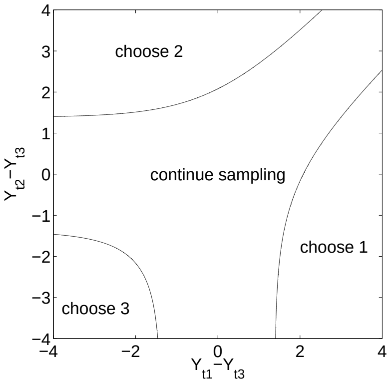
Paulson, E. (1964).
A sequential procedure for selecting the population with the largest mean from k normal populations.
The Annals of Mathematical Statistics, 35(1):174–180.

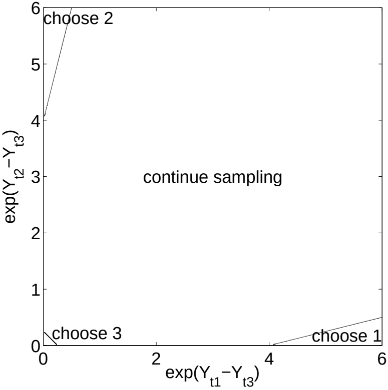
Paulson, E. (1994).
Sequential procedures for selecting the best one of k koopman-darmois populations.
Sequential Analysis, 13(3).

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

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.





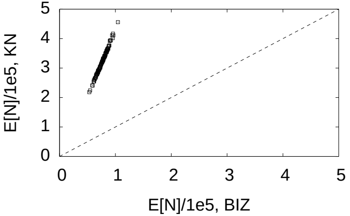
BIZ in Discrete-Time

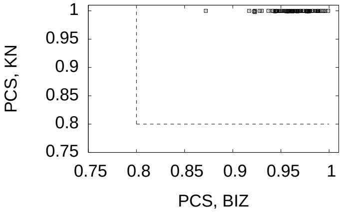
When $\mathbb{T} = \mathbb{Z}_+$, BIZ can be written:

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.

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 1 2 3 4

t=2 1 2 3 4

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

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

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

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

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

Selection: Select alternative 3 as the best.

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 $(Y_{tx} : 0 \leq t \leq \varepsilon)$ 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 bound remain true.

Heterogeneous Sampling Variance: Discrete Time

- In discrete time, the problem is harder:
- Idea: Replace $Y_{t,x}$ with $Y_{n_x(t),x}/\widehat{\sigma_x^2}$ where $n_x(t)$ rounds up $\widehat{\sigma_x^2}t$.
- The IZ guarantee no longer holds.
- Ongoing work: Does the IZ guarantee hold in a limiting sense?