A New Solution to a Classic Problem: Ranking & Selection with Tight Lower Bounds on Probability of Correct Selection

Peter I. Frazier

Operations Research & Information Engineering, Cornell University

Friday February 18, 2011
Center for Applied Mathematics (CAM) Colloquium
Cornell University

Outline

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 write $\theta = [\theta_1, \ldots, \theta_k]$.
- Approximate normality can be checked empirically, and if it is not met, samples can be batched together.
- **Goal:** Use simulation efficiently to find $\arg\max_x \theta_x$.

Typical Ad Hoc Approach

- Pick some number N of out the air.
  - Often, N is some round number that feels like it has the right order of magnitude, e.g., 10^2, 10^3, 10^4.
- Take N samples from each alternative.
- Calculate the sample mean $\overline{Y}_x = \sum_{n=1}^{N} Y_{nx}$ for each alternative x.
- Select $\hat{x} \in \arg\max_x \overline{Y}_x$ as the alternative we estimate to be the best.

Issue with the Typical Ad Hoc Approach

- **Issue:** We have no guarantees on the performance of the estimator $\hat{x}$.
- How likely is it that $\hat{x}$ selects the best alternative?
  We do not know.
- If we knew θ we could compute it, but θ is unknown.

Definitions: Policies and PCS

- 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)
- 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.)

Definitions: IZ Guarantee

- Let $\delta > 0$ be the smallest difference the experimenter feels is worth detecting. This is a fixed parameter.
- 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.
- The **indifference zone (IZ)** is the complement of 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.**

[Bechhofer, 1954]

- The IZ formulation of the R&S problem is due to [Bechhofer, 1954], which also presented the following policy for the case $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2$.
- Bechhofer's procedure: From each alternative take

$$\left\lceil \frac{2h^2\sigma^2}{\delta^2} \right\rceil$$

  samples, where h is a fixed constant defined in terms of the normal distribution.
- This policy satisfies the IZ guarantee.

[Bechhofer, 1954]

Comments on Bechhofer's procedure:

- This policy is reasonable for the **slippage configuration**, $\theta = [\delta, 0, \ldots, 0]$, but samples more than necessary under other configurations.
- Subsequent research has focused on sequential policies that can stop early or drop bad alternatives under easier configurations.

Indifference Zone Formulation

Many papers have constructed policies satisfying the IZ guarantee. Here is a partial list:

- 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, Kim and Nelson, 2006]

Fully Sequential Policies

- Fully sequential policies improve on the performance of [Bechhofer, 1954], but . . .
- . . . techniques used to prove the IZ-guarantee relied on bounds that are loose for large numbers of alternatives k.

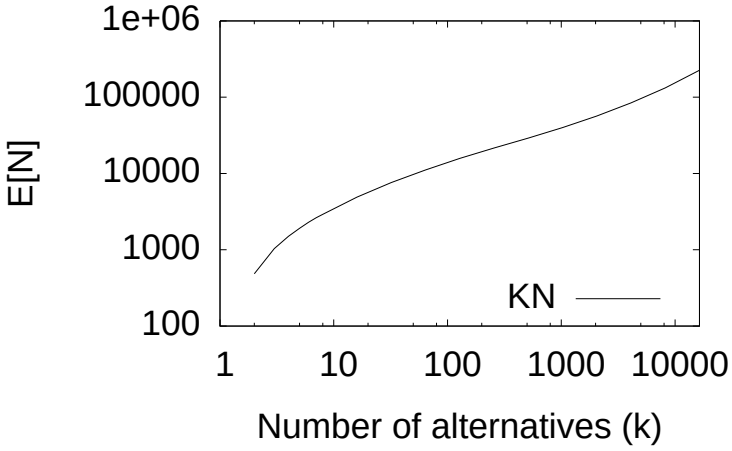
Overdelivery & Too Many Samples

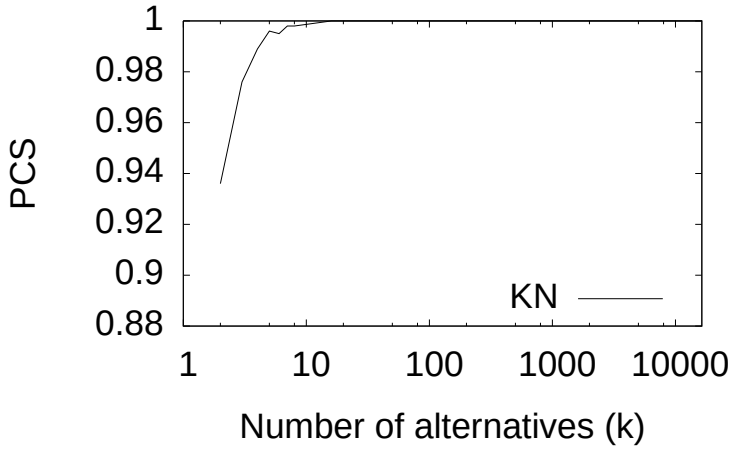
- Let π be a procedure with an IZ guarantee for a fixed P^* and δ.
- P^* is the PCS that π guarantees it will deliver.
- For any $\theta \in \mathrm{PZ}(\delta)$,

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

  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.

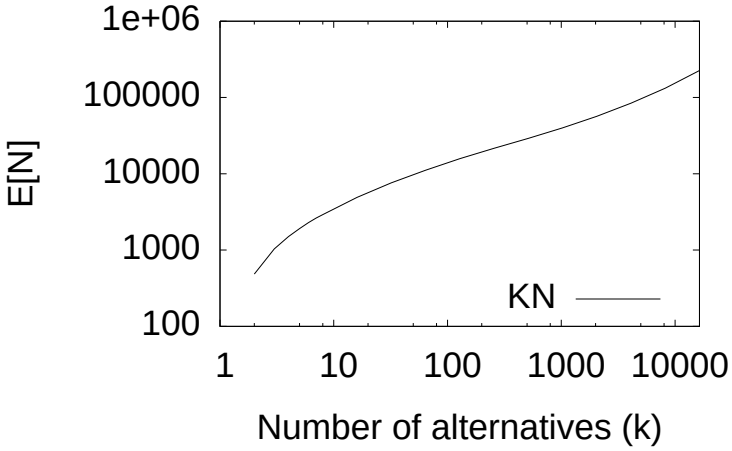
Example: Too Many Samples

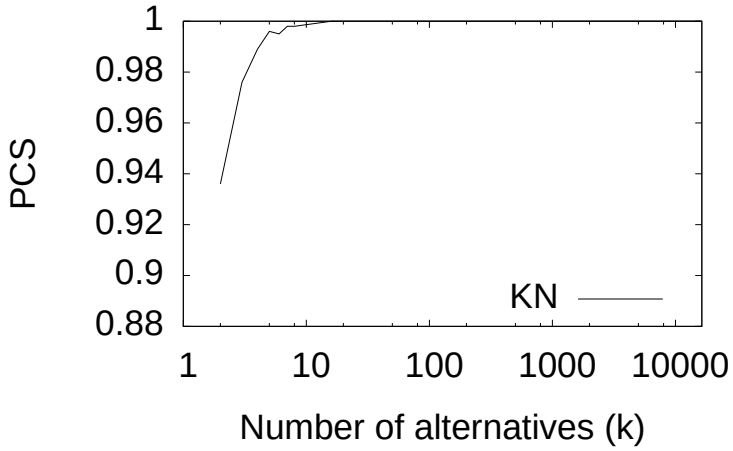




- This is the **monotone-decreasing-means configuration**, $\theta = [-\delta, -2\delta, \ldots, -k\delta]$.
- Parameters are: $\delta = 1$, $\sigma_x = 10$, $P^* = 0.9$.
- The policy is the KN policy of [Kim and Nelson, 2001], which is state-of-the-art for problems with not much variation in σ_x^2.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

Example: Too Many Samples





- $\theta = [-\delta, -2\delta, \ldots, -k\delta]$, $\delta = 1$, $\sigma_x = 10$, $P^* = 0.9$.
- We expect $E[N]$ should grow slowly with k:
  - When $k = 1000$, the worst alternative is $999/\sigma_x \approx 100$ standard deviations worse than the best, and so should be easily identified as worse with just one sample.
- $E[N]$ grows too quickly with k because KN is constructed with union (Bonferonni) bounds that grow loose as k grows large.
- This defect is shared by nearly every sequential R&S procedure.

Goal of this Research

- Given any P^* and δ, construct a fully sequential policy π for which
  1. We can prove that π has the IZ guarantee with parameters δ and P^*:

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

  2. The lower bound P^* on PCS is tight:

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

- It seems reasonable to expect that such a policy will sample less than existing policies with the IZ guarantee, when k is large.

Outline

simple-BIZ

We first present a simple version of the **Bayes-inspired IZ (BIZ)** policy that does not eliminate alternatives, and that assumes a common sampling variable $\sigma_x^2 = \sigma^2$, called "simple-BIZ".

- simple-BIZ samples each alternative at each time t.
  - In the jargon of R&S, we say simple-BIZ is a "non-elimination procedure."
- Upon stopping, simple-BIZ selects the alternative with the largest sample mean as the best.

Bayesian Prior

- BIZ is derived using Bayesian ideas, even though the guarantees on its performance are non-Bayesian.
- Let Q be a prior probability measure on $\mathrm{PZ}(\delta)$ under which

$$X_* \sim \mathrm{Uniform}(1, \ldots, k)$$
$$\theta_x = \begin{cases} a + \delta & \text{if } x = X_* \\ a & \text{if } x \neq X_*, \end{cases}$$

  for some fixed $a \in \mathbb{R}$.
- Q is concentrated on slippage configurations.
- We assume that the sampling variance is known and homogeneous, $\sigma_x^2 = \sigma^2$.

Bayesian Posterior

- Let $Y_t = (Y_{t1}, \ldots, Y_{tk})$, where Y_{tx} is the sum of all observations from alternative x by time t.
  - In discrete time, $(Y_{tx} : t \in \mathbb{Z}_+)$ is a random walk with independent $\mathrm{Normal}(\theta_x, \sigma^2)$ increments.
  - In continuous time, $(Y_{tx} : t \in \mathbb{R}_+)$ is a Brownian motion with drift θ_x and volatility σ.
- Given data Y_t, the Bayesian posterior distribution on the best alternative X_* is

$$q_{tx} = Q\{X_* = x \mid Y_t\} = \frac{\exp\left(\frac{\delta}{\sigma} Y_{tx}\right)}{\sum_{x'=1}^{k} \exp\left(\frac{\delta}{\sigma} Y_{tx'}\right)}$$

- The posterior on X_* does not depend on a.

simple-BIZ

- q_{tx} is the Bayesian posterior probability that alternative x is best.
- If we stop at t and select $\hat{x} \in \arg\max_x q_{tx}$, our Bayesian PCS is $\max_x q_{tx}$.
  (Bayesian PCS differs from the (non-Bayesian) PCS already discussed).
- simple-BIZ stops sampling when this Bayesian probability of correct selection exceeds our target P^*:

$$\tau = \inf\left\{t \in \mathbb{T} : \max_x q_{tx} \geq P^*\right\},$$

  where $\mathbb{T} = \mathbb{R}_+$ for continuous time, or $\mathbb{T} = \mathbb{Z}_+$ for discrete time.
- simple-BIZ can be defined without referencing its Bayesian interpretation by defining q_{tx} directly in terms of $\exp(\frac{\delta}{\sigma} Y_{tx})$.

Main Result (simple-BIZ)

Theorem

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

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

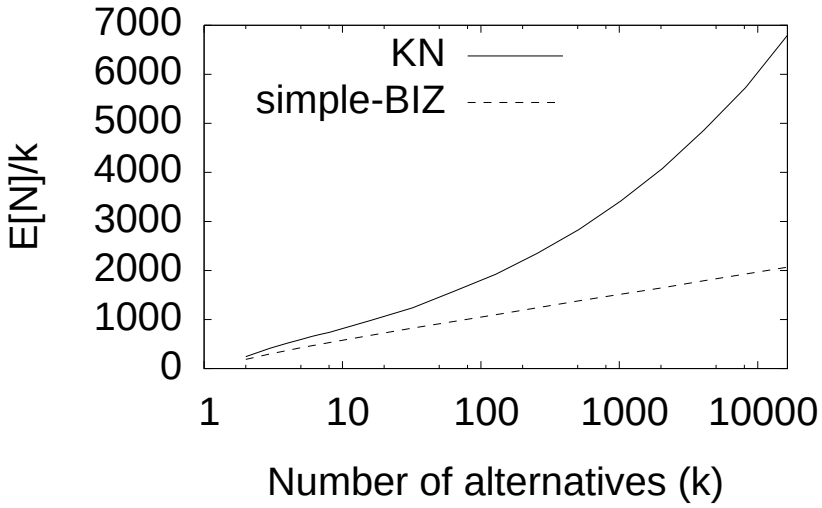
Moreover, if $\mathbb{T} = \mathbb{R}_+$, then

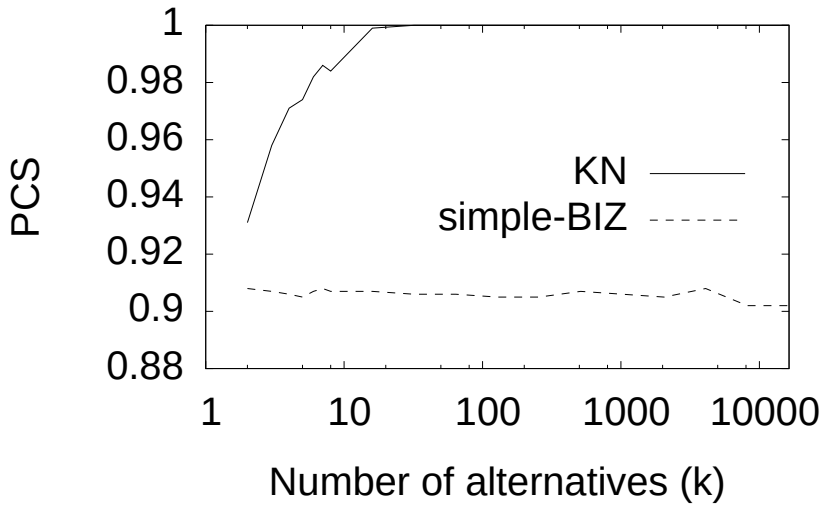
$$\inf_{\theta \in \mathrm{PZ}(\delta)} \mathrm{PCS}(\pi, \theta) = P^*$$

In other words:

- **simple-BIZ satisfies the IZ-guarantee.**
- **In continuous-time, the bound on PCS is tight.**

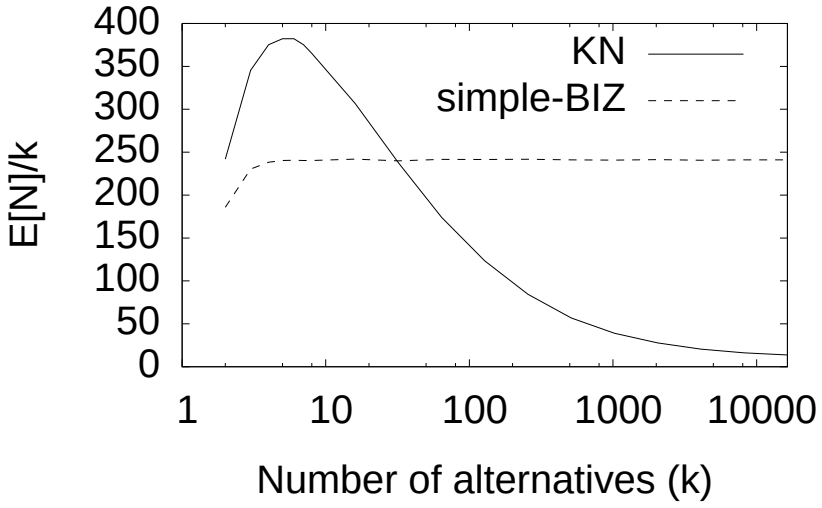
Numerical Comparisons: Slippage Configuration

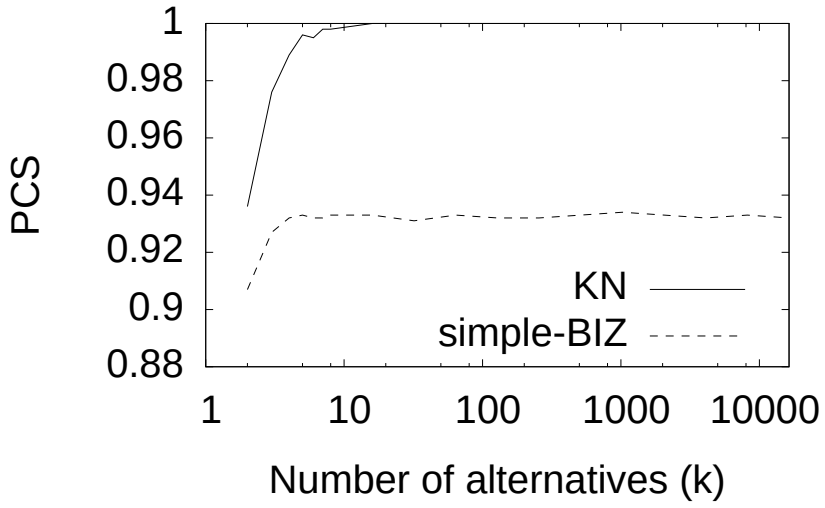




- The slippage configuration is $\theta = [0, \ldots, 0, \delta]$.
- KN is the previously discussed policy from [Kim & Nelson 2001]
- Evaluation is in discrete time, with $P^* = 0.9$, $\sigma = 10$, $\delta = 1$.
- Estimated with $\geq 10,000$ independent replications.

Numerical Comparisons: Monotone Decreasing Means





- The monotone decreasing means configuration is $\theta = [-\delta, -2\delta, \ldots, -k\delta]$.
- Evaluation is in discrete time, with $P^* = 0.9$, $\sigma = 10$, $\delta = 1$, estimated with $\geq 10,000$ independent replications.

Ideas in Proof

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

Lemma (Symmetry)

Let $\theta \in \mathbb{R}^k$. Then $\mathrm{PCS}(\pi, \theta) = Q_\theta^\pi \{\mathrm{CS}\}$.

Lemma (Monotonicity)

Let $\theta \in \mathrm{PZ}(\delta)$. Then $Q_\theta^\pi \{\mathrm{CS}\} \geq Q^\pi \{\mathrm{CS}\}$

Lemma (Stopping)

If $\mathbb{T} = \mathbb{Z}_+$, then $Q^\pi\{\mathrm{CS}\} \geq P^*$.
If $\mathbb{T} = \mathbb{R}_+$, then $Q^\pi\{\mathrm{CS}\} = P^*$.

Outline

BIZ with Elimination

- Under the monotone decreasing means configuration, the performance of simple-BIZ suffered because it sampled all of the alternatives at each time t, even the ones that were clearly not best.
- We generalize simple-BIZ to eliminate bad alternatives early.
- This focuses sampling effort on alternatives in contention, reducing the overall number of samples.
- Our theoretical results still hold (IZ guarantee, tight bounds in continuos-time).

Elimination in General

Elimination is a common strategy used by R&S procedures.

- We maintain a set of alternatives A_t that are still "in contention."
- At time t, we sample only from those alternatives in A_t.
- For $t > s$, $A_t \subseteq A_s$, so we can drop alternatives but cannot add then.

Bayesian Prior and Posterior Distribution

- We use the same prior Q: $X_* \sim \mathrm{Uniform}(1, \ldots, k)$, $\theta_x = a + \delta 1_{\{x = X_*\}}$.
- For $A \subseteq \{1, \ldots, k\}$, define

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

- 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 in Discrete-Time

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

1. Let $A \leftarrow \{1, \ldots, k\}$, $t \leftarrow 0$, $\alpha \leftarrow 1 - P^*$.
2. While $\max_{x \in A} q_{tx}(A) < 1 - \alpha$
   - 2a. While $\max_{x \in A} 1 - q_{tx}(A) \geq b$
     - Let $x \in \arg\max_x 1 - q_{tx}(A)$.
     - Let $\alpha \leftarrow \alpha/(1 - q_{tx}(A))$.
     - 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.

Choosing $b \geq 1$ causes us to never eliminate, recovering simple-BIZ.

BIZ in General (Discrete and Continuous Time)

- We define a sequence of stopping times, $0 = \tau_0 \leq \tau_1 \cdots \leq \tau_{k-1} \leq \infty$. τ_n is the time that the nth alternative is eliminated.
- Let $Z_n \in \arg\min_{x \in A_{n-1}} Y_{\tau_n, x}$, breaking ties uniformly at random. Z_n is the nth alternative eliminated, and is only defined if $\tau_n < \infty$.
- Let $A_n = \{1, \ldots, k\} \setminus \{Z_m : m \leq n\}$ be the alternatives that remain after the first n are eliminated.
- Let $\tau_{n+1} = \inf\{t \in \mathbb{T} : t \geq \tau_n, \max_{x \in A_n} 1 - q_{tx}(A_n) \geq b\}$.
- Let $\alpha_0 = 1 - P^*$, $\alpha_{n+1} = \alpha_n / (\max_{x \in A_n} 1 - q_{tx}(A_n))$.
- Let $N_t = \min\{n : t \geq \tau_n\}$ be the number eliminated by time t.
- Let $\tau = \inf\{t \in \mathbb{T} : \max_{x \in A_{N_t}} q_{tx}(A_{N_t}) \geq 1 - \alpha_{N_t}\}$. This is the time we select an alternative.

Main Result (BIZ)

Theorem (Pending Check)

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

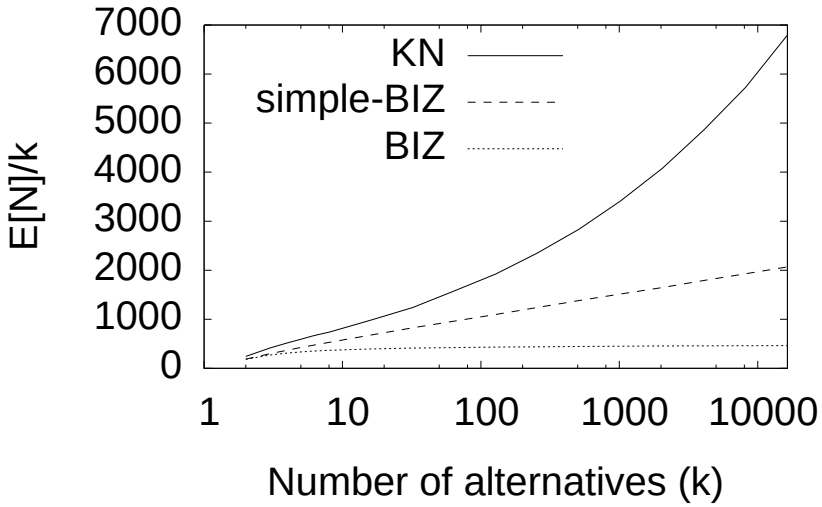
$$\mathrm{PCS}(\pi, \theta) \geq P^* \quad \forall \theta \in \mathrm{PZ}(\delta)$$

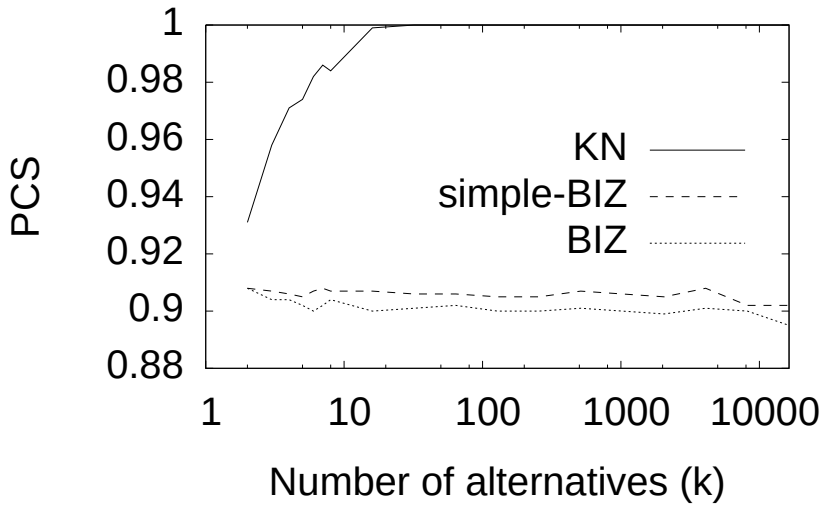
Moreover, if $\mathbb{T} = \mathbb{R}_+$, *then*

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

This generalizes the previous result, since choosing $b \geq 1$ gives the simple-BIZ policy.

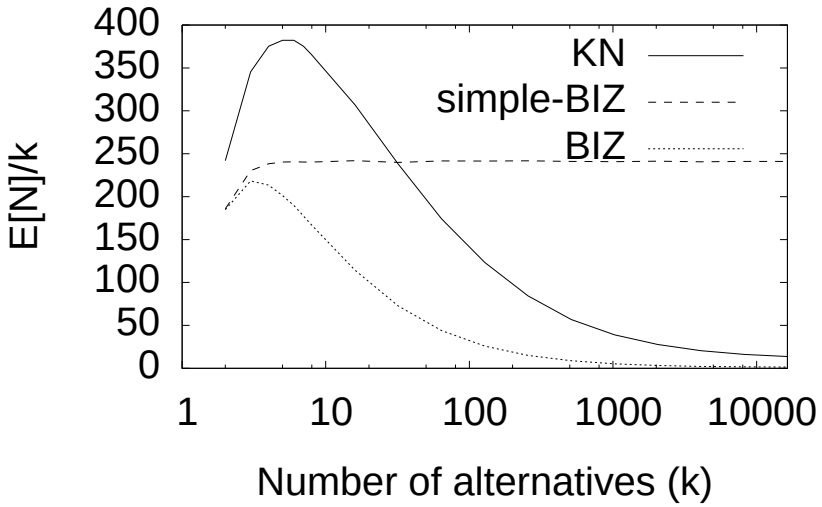
Numerical Comparisons: Slippage Configuration

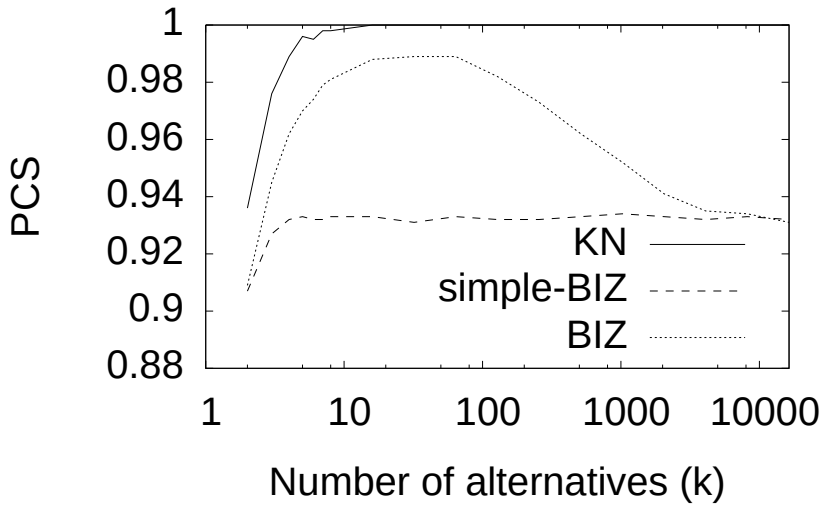




- Settings are the same as before: $\theta = [0, \ldots, 0, \delta]$, $P^* = 0.9$, $\sigma = 10$, $\delta = 1$, estimated with $\geq 10,000$ independent replications.
- BIZ uses $b = (1 - P^*)^{1/(k-1)}$.

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 $b = (1 - P^*)^{1/(k-1)}$.

Outline

Heterogeneous and Unknown Sampling Variances

- Thus far we have assumed that the sampling variance σ_x^2 are known and homogeneous, $\sigma_x^2 = \sigma^2$.
- This is unrealistic in most simulation applications (although it may be more realistic in other applications).
- What can be done? Consider 3 cases: continuous time, discrete-time variances known, and discrete-time unknown variances.

Sampling Variances: Discrete Time, Known Variances

Suppose $\sigma_1, \ldots, \sigma_k$ are known but unequal.

- Instead of taking 1 sample per time-step, take $m_x \geq 1$. This replaces Y_{tx} with $\frac{1}{m_x} Y_{tm_x,x}$, which has $\mathrm{Normal}(\theta_x, \sigma_x^2/m_x)$ increments.
- Let $\lambda = \max_x \sigma_x^2/m_x$.
- To each composite observation of alternative x, simulate and normal noise with mean 0 and variance σ_x^2/m_x.
- The resulting problem has common known sampling variance of λ.

This method might be reasonable if the original $\sigma_1, \ldots, \sigma_k$ are close to each other, but seems unsatisfying in other cases.

Sampling Variances: Discrete Time, Unknown Variances

Suppose $\sigma_1, \ldots, \sigma_k$ are unknown.

- Perform an initial stage of sampling where we take, e.g., 30 samples from each alternative, estimate the each sampling variance, and then use BIZ as if these estimates were correct.
- This will work reasonably well IF:
  1. The cost of the first stage is small compared to the overall simulation effort required (sampling variance is high compared to differences in θ).
  2. The performance of BIZ is insensitive to small changes in the sampling variances.

Sampling Variances: Continuous Time

- In continuous time, we can estimate σ_x^2 perfectly given $(Y_{tx} : 0 \leq t \leq \varepsilon)$ for any $\varepsilon > 0$. Let $0 = t_0 < t_1 < \ldots < \varepsilon$. Then,

$$\sum_{n=1}^{\infty} (Y_{t_n,x} - Y_{t_{n-1},x})^2 / (t_n - t_{n-1}) = \sigma_x^2 \quad \text{almost surely.}$$

- After estimating σ_x^2, let $\lambda = \max_x \sigma_x$, $m_x = \sigma_x^2/\lambda$, $Z_{tx} = \frac{1}{m_x} Y_{tm_x,x}$. Then, Z_{tx} has drift θ_x and volalitily λ.
- We have transformed a variance-unknown problem to a problem with common known variance without adding any simulated noise.

Conclusion

- BIZ is a fully sequential IZ policy that has **tight bounds on PCS**.
- To my knowledge, this is the first sequential IZ policy with this property for $k > 2$.
- By not overdelivering on PCS, BIZ requires **many fewer samples** to make a selection than existing R&S procedures.
- BIZ makes it feasible to do R&S with statistical guarantees for much larger values of k: $k = 10^3$, 10^4, 10^5, ...
- This method of Bayesian analysis with least-favorable priors to do worst-case analysis may have other applications.
- Ongoing work: hetoregenous sampling variance, choice of b, other allocation rules, ...

Thank You!

References

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.

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.

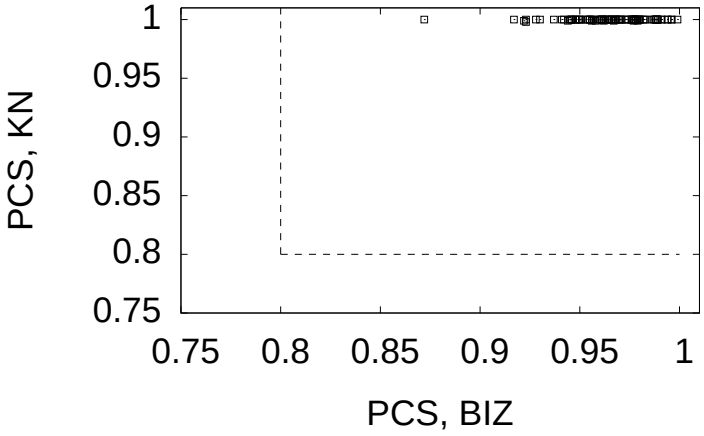
Hartmann, M. (1988).
An improvement on paulson s sequential ranking procedure.
Sequential Analysis, 7(4):363–372.

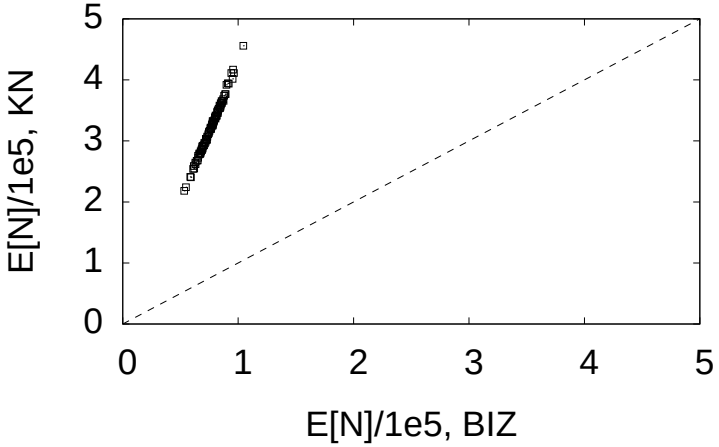
Hong, J. (2006).
Fully sequential indifference-zone selection procedures with variance-dependent sampling.
Naval Research Logistics (NRL), 53(5):464–476.

Kim, S. and Nelson, B. (2006).
On the asymptotic validity of fully sequential selection procedures for steady-state simulation.
Operations Research, 54(3):475–488.

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.

Numerical Comparisons: Random Problem Instances





- 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 $\theta \in \mathrm{PZ}(\delta)$.

Continuation and Stopping Region

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