

ASYMPTOTIC VALIDITY OF A FULLY SEQUENTIAL ELIMINATION PROCEDURE FOR INDIFFERENCE-ZONE RANKING AND SELECTION WITH TIGHT BOUNDS ON PROBABILITY OF CORRECT SELECTION

Saul Toscano-Palmerin
Peter I. Frazier

Cornell University
257 Rhodes Hall
206 Rhodes Hall
Ithaca, NY 14853, USA

ABSTRACT

We consider the indifference-zone (IZ) formulation of the ranking and selection problem in which the goal is to choose with high probability an alternative with similar mean than the one with largest mean. Conservatism leads classical IZ procedures to take too many samples in problems with many alternatives. The Bayes-inspired Indifference Zone (BIZ) procedure, proposed in Frazier (2014), is less conservative than previous procedures, but its proof of validity requires strong assumptions, specifically it assumes that the variances are known and have an integer multiple structure. In this paper, we consider a slightly modification of the original BIZ procedure in order to simplify the analysis, and we present a new proof of its asymptotic validity that relaxes these assumptions. Specifically, we prove the validity of the algorithm when the variances are known and the difference between the best alternative and the second best goes to zero.

1 INTRODUCTION

There are many applications where we have to choose the best alternative among a finite number of simulated alternatives. For example, in inventory problems, we may want to choose the best inventory policy (s, S) for a finite number of values of s and S . The ranking and selection problem has to decide how many samples are needed to find the alternative with the largest mean. A good solution to the problem is efficient and accurate, i.e. the procedure proposed for the problem must balance between the number of samples required and quality of the selection.

This paper considers the indifference-zone (IZ) formulation of the ranking and selection problem, in which the distance between the best system and the other systems is sufficiently large and the best system is chosen with probability larger than some threshold given by the user. This property is called the IZ guarantee, and the preference zone (PZ) is defined as the set of system configurations under which the difference between the best system and the others is at least some given $\delta > 0$. The paper Bechhofer (1954) is considered the seminal work, and early work is presented in the monograph Bechhofer, Kiefer, and Sobel (1968). Some compilations of the theory developed in the area can be found in R. E. Bechhofer (1995), Swisher, Jacobson, and Yücesan (2003), Kim and Nelson (2006a) and Kim and Nelson (2007). Beyond these classical approaches, there are the Bayesian approach (Frazier 2012) and the optimal computing budget allocation approach (Chen and Lee 2011).

A good IZ procedure satisfies the IZ guarantee and requires as few samples as possible. The first IZ procedures presented in Bechhofer (1954), Paulson (1964), Fabian (1974), Rinott (1978), Hartmann (1988), Hartmann (1991), Paulson (1994) satisfy the IZ guarantee, but they usually take too many samples when there are many alternatives, in part because their probability of correct selection (PCS) is much larger than the probability specified by the user (Wang and Kim 2013). One reason for this is that these procedures use Bonferroni’s inequality, which leads then to sample more than necessary. More recently, new algorithms were developed in Kim and Nelson (2001), Goldsman, Kim, Marshall, and Nelson (2002), Hong (2006), and they improve the performance but they still use the Bonferroni’s inequality, and so the methods are inefficient when there are many alternatives. Procedures in Kim and Dieker (2011), Dieker and Kim (2012) do not use the Bonferroni’s inequality when there are only three alternatives, but when they compare more than three alternatives, they do use the Bonferroni’s inequality.

Conservative algorithms lead to take more samples than needed, and consequently conservative procedures are unpopular when there are more than a few hundred of alternatives. Three common sources of conservativeness are the Bonferroni’s inequality, the change from discrete process to continuous process, and bound the worst case which is not a common case. Frazier (2014) eliminates one source of conservativeness: the Bonferroni’s inequality. He presented a new sequential elimination IZ procedure, called BIZ (Bayes-inspired Indifference Zone), whose lower bound on worst-case probability of correct selection in the preference zone is tight in continuous time, and almost tight in the discrete time. In numerical experiments, the number of samples required by BIZ is significantly smaller than that of procedures like the KN procedure of Kim and Nelson (2001) and the \mathcal{P}_B^* procedure of Bechhofer, Kiefer, and Sobel (1968), especially on problems with many alternatives. Unfortunately, the proof that the BIZ procedure satisfies the IZ guarantee for the discrete-time case assumes that variances are known and have an integer multiple structure which is not very realistic.

The contribution of this work is that we prove the asymptotic validity of the BIZ procedure for the discrete-time case when the variances are known and δ goes to zero. In order to simplify analysis, we consider a slightly modification of this procedure. Although we only proved it in the known variance setting, it is still a significantly generalization over the assumption that the variances are known and have an integer multiple structure. Furthermore, we do not need to assume that the alternatives follow a normal distribution. The only assumptions are that the alternatives are independent, identically distributed and have finite variance. We conjecture that these techniques can be used to show the validity of this procedure when the variances are unknown, and we present numerical experiments that support this belief. Kim and Nelson (2006b) also proves the asymptotical validity of a IZ procedure in the same limit. Our proof shares some similarities since we both use a central limit theorem.

This paper is organized as follows: In 2, we recall the indifference-zone ranking and selection problem. In 3, we recall the Bayes-inspired IZ (BIZ) procedure from Frazier (2014). In 4, we present the proof of the validity of the algorithm when the variances are known. In 5, we present some numerical experiments. In 6, we conclude.

2 INDIFFERENCE-ZONE RANKING AND SELECTION

Ranking and Selection is a problem where we have to select the best system among a finite set of alternatives, i.e. the system with the largest mean. The method selects a system as the best based on the samples that are observed sequentially over time. We suppose that samples are identically distributed and independent, over time and across alternatives, and each alternative x has mean μ_x . We define $\mu = (\mu_1, \dots, \mu_k)$.

If the best system is selected, we say that the procedure has made the *correct selection* (CS). We define the *probability of correct selection* as

$$\text{PCS}(\mu) = \mathbb{P}_\mu(\hat{x} \in \arg \max_x \mu_x)$$

where \hat{x} is the alternative chosen by the procedure and \mathbb{P}_μ is the probability measure under which samples from system x have mean μ_x and finite variance λ_x^2 .

In the Indifference-Zone Ranking and Selection, the procedure is indifferent in the selection of a system whenever the means of the populations are nearly the same. Formally, let $\mu = [\mu_k, \dots, \mu_1]$ be the vector of the true means, the *indifference zone* is defined as the set $\{\mu \in \mathbb{R}^k : \mu_{[k]} - \mu_{[k-1]} < \delta\}$. The complement of the indifference zone is called the *preference zone* (PZ) and $\delta > 0$ is called the indifference zone parameter. We say that a procedure meets the *indifference-zone (IZ) guarantee* at $P^* \in (1/k, 1)$ and $\delta > 0$ if

$$\text{PCS}(\mu) \geq P^* \text{ for all } \mu \in \text{PZ}(\delta).$$

We assume $P^* > 1/k$ because IZ guarantees can be met by choosing \hat{x} uniformly at random from among $\{1, \dots, k\}$.

3 THE BAYES-INSPIRED IZ (BIZ) PROCEDURE

BIZ is an elimination procedure. This procedure maintains a set of alternatives that are candidates for the best system, and it takes samples from each alternative in this set at each point in time. At beginning, all alternatives are possible candidates for the best system, and over the time alternatives are eliminated. The procedure ends when there is only one alternative in the contention set and this remain alternative is chosen as the best.

Frazier (Frazier 2014) showed that the BIZ procedure with known common variance satisfies the IZ guarantee when the systems follow the normal distribution, with tight bounds on worst-case preference-zone in continuous time. He also proved that this procedure retains the IZ guarantee when the systems follow the normal distribution, and the variances are known and are integer multiples of a common value. The continuous time version of this procedure also satisfies the IZ guarantee, with a tight worst-case preference-zone PCS bound.

The discrete-time BIZ procedure for unknown and/or heterogeneous sampling variances is given below. It takes a variable number of samples from each alternative, and n_{tx} is this number. This algorithm depends on a collection of integers $B_1, \dots, B_k, P^*, c, \delta$ and n_0 . n_0 is the number of samples to use in the first stage of samples, and 100 is the recommended value for n_0 . B_x controls the number of samples taken from system x in each stage. The procedure presented is a slightly modification of the original BIZ procedure where $z \in \arg \max_{x \in A} \hat{\lambda}_{tx}^2$, instead of $z \in \arg \min_{x \in A} n_{tx} / \hat{\lambda}_{tx}^2$.

For each $t, x \in \{1, \dots, k\}$, and subset $A \subset \{1, \dots, k\}$, we define a function

$$q'_{tx}(A) = \exp\left(\delta \beta_t \frac{Z_{tx}}{n_{tx}}\right) \bigg/ \sum_{x' \in A} \exp\left(\delta \beta_t \frac{Z_{tx'}}{n_{tx'}}\right), \quad \beta_t = \frac{\sum_{x' \in A} n_{tx'}}{\sum_{x' \in A} \hat{\lambda}_{tx'}^2}$$

where $\hat{\lambda}_{tx}^2$ is the sample variance of all samples from alternative x thus far and $Z_{tx} = Y_{n_{tx}, x}$.

Algorithm: Discrete-time implementation of BIZ, for unknown and/or heterogeneous variances.

Require: $c \in [0, 1 - (P^*)^{\frac{1}{k-1}}]$, $\delta > 0$, $P^* \in (1/k, 1)$, $n_0 \geq 0$ an integer, B_1, \dots, B_k strictly positive integers.

Recommended choices are $c = 1 - (P^*)^{\frac{1}{k-1}}$, $B_1 = \dots = B_k = 1$ and n_0 between 10 and 30. If the sampling variances λ_x^2 are known, replace the estimators $\hat{\lambda}_{tx}^2$ with the true values λ_x^2 , and set $n_0 = 0$.

- 1: For each x , sample alternative x n_0 times and set $n_{0x} \leftarrow n_0$. Let W_{0x} and $\hat{\lambda}_{0x}^2$ be the sample mean and sample variance respectively of these samples. Let $t \leftarrow 0$. Let $z \in \arg \max_{x \in A} \hat{\lambda}_{tx}^2$.
- 2: Let $A \leftarrow \{1, \dots, k\}$, $P \leftarrow P^*$.
- 3: **while** $x \in \arg \max_{x \in A} q'_{tx}(A) < P$ **do**
- 4: **while** $\min_{x \in A} q'_{tx}(A) \leq c$ **do**
- 5: Let $x \in \arg \min_{x \in A} q_{tx}(A)$.
- 6: Let $P \leftarrow P / (1 - q_{tx}(A))$.

- 7: Remove x from A .
- 8: **end while**
- 9: For each $x \in A$, let $n_{t+1,x} = \text{ceil}\left(\hat{\lambda}_{tx}^2(n_{tz} + B_z)/\hat{\lambda}_{tz}^2\right)$.
- 10: 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 $W_{t+1,x}$ and $\hat{\lambda}_{t+1,x}^2$ be the sample mean and sample variance respectively of all samples from alternative x thus far.
- 11: Increment t .
- 12: **end while**
- 13: Select $\hat{x} \in \arg \max_{x \in A} Z_{tx}/n_{tx}$ as our estimate of the best.

This algorithm generalizes the BIZ procedure with known common variance. In that case, we have that $B_1 = \dots = B_k = 1$ and $n_{tx} = t$. The algorithm 2 can be generalized to the continuous case (See Frazier (2014)).

4 ASYMPTOTIC VALIDITY WHEN THE VARIANCES ARE KNOWN

In this section we prove that the BIZ procedure satisfies asymptotically the IZ guarantee when the variances are known. This means that we consider a collection of ranking and selection problems parametrized by $\delta > 0$. For the problem given δ , we suppose that the vector of the true means $\mu = [\mu_k, \dots, \mu_1]$ is equal to δa for some fixed $a \in \mathbb{R}^k$ that does not depend on δ and $a_k > a_{k-1} \geq \dots \geq a_1$, $a_k - a_{k-1} > 1$. Moreover, the variances of the alternatives are finite, strictly greater than zero and do not depend on δ . We also suppose that samples from system $x \in \{1, \dots, k\}$ are identically distributed and independent, over time and across alternatives. We also define $\lambda_z^2 := \max_{i \in \{1, \dots, k\}} \lambda_i^2$.

Any ranking and selection algorithm can be viewed as mapping from paths of the k -dimensional discrete-time random walk $(Y_{tx} : t \in \mathbb{N}, x \in \{1, \dots, k\})$ onto selection decisions. The BIZ procedure's mapping from paths onto selections decisions can be understood as the composition of three simpler maps.

The first is the mapping from the raw discrete-time random walk $(Y_{tx} : t \in \mathbb{N}, x \in \{1, \dots, k\})$ onto a time changed version of this random walk, written as $(Z_{tx} : t \in \mathbb{N}, x \in \{1, \dots, k\})$, where we recall $Z_{tx} = Y_{n_x(t),t}$ is the sum of the samples from alternative x observed by stage t .

The second maps this time-changed random walk through a non-linear mapping for each t, x and subset $A \subset \{1, \dots, k\}$, to obtain $(q'_{tx}(A) : t \in \mathbb{N}, A \subset \{1, \dots, k\}, x \in A)$, where

$$q'_{tx}(A) = \exp\left(\delta \beta_t \frac{Z_{tx}}{n_{tx}}\right) \bigg/ \sum_{x' \in A} \exp\left(\delta \beta_t \frac{Z_{tx'}}{n_{tx'}}\right) := q'((Z_{tx}, t : x \in A), \delta, t)$$

where we note that $n_x(t)$ and β_t are deterministic in the version of the known-variance BIZ procedure that we consider here.

The third maps the paths of $(q'_{tx}(A) : t \in \mathbb{N}, A \subset \{1, \dots, k\}, x \in A)$ onto selection decisions. Specifically, this mapping finds the first time that $\max_{x \in A} q'_{tx}(A)$ is larger than some threshold P , and then it chooses the system with the largest empirical mean Z_{tx}/n_{tx} .

Later, we can approximate the previous process by the continuous time process

$$q_{tx}(A) = q\left(\left(Z_{\frac{t}{\delta^2}x}, \frac{t}{\delta^2} : x \in A\right), \delta, A \subset \{1, \dots, k\}\right) \quad (1)$$

where $t \geq 0$. Then we will show the difference in selection decisions vanishes as $\delta \rightarrow 0$.

After that, we would like to use a kind of central limit theorem for $Z_{\frac{t}{\delta^2}x}$ when δ goes to zero to work with a Brownian motion instead. In order to do this, we are going to rewrite equation (1) in terms of a centralized version of Z :

$$\mathcal{C}_x(\delta, t) = \frac{Y_{n_x(t),x} - t\lambda_x^2\mu_x}{\frac{\lambda_x^2}{\lambda_z}\delta}$$

and the resulting expression is

$$q_{tx}(A) = q\left(\left(C_x(\delta, t) \frac{\lambda_x^2}{\delta\lambda_z^2} + \frac{\lambda_x^2}{\lambda_z^2} \left(n_0 + \frac{t}{\delta^2}\right) \delta a_x, \frac{t}{\delta^2} : x \in A\right), \delta, A \subset \{1, \dots, k\}\right).$$

Thus, we can see the selection decision resulting from the BIZ algorithm as the result of applying some mapping $f(\cdot, \delta)$ to the paths $t \mapsto \mathcal{C}_x(\delta, t)$.

The last part of our proof is to observe that (1) $\mathcal{C}(\delta, \cdot)$ converges to a multivariate Brownian motion W as δ goes to 0; (2) the function f has a continuity property that causes

$$f(\mathcal{C}(\delta, \cdot), \delta) \Rightarrow g(W)$$

where g is the selection decision from applying the BIZ procedure in continuous time; and (3) the BIZ procedure satisfies the IZ guarantee when applied in continuous time.

First, we are going to see that the the standardized sum of the output data converges to a Brownian motion in the sense of $D_\infty := D[0, \infty)$, which is the set of functions from $[0, \infty)$ to \mathbb{R} that are right-continuous and have left-hand limits, with the Skorohod topology. The definition and the properties of this topology may be found in Chapter 3 of Billingsley (1999).

We briefly recall the definition of convergence of random paths in the sense of D_∞ . Suppose that we have a sequence of random paths $(\mathcal{X}_n)_{n \geq 0}^\infty$ such that $\mathcal{X}_n : \Omega \rightarrow D_\infty$ where $(\Omega, \mathcal{F}, \mathbb{P})$ is our probability space. We say that $\mathcal{X}_n \Rightarrow \mathcal{X}_0$ in the sense of D_∞ if $P_n \Rightarrow P_0$ where $P_n : \mathcal{D}_\infty \rightarrow [0, 1]$ are defined as $P_n[A] = \mathbb{P}[\mathcal{X}_n^{-1}(A)]$ for all $n \geq 0$ and \mathcal{D}_∞ are the Borel subsets for the Skorohod topology.

The following lemma shows that the standardized sum of the output data with t changed by t/δ^2 converges to a Brownian motion in the sense of D_∞ .

Lemma 1 For each $x \in \{1 \dots, k\}$, we define

$$\mathcal{C}_x(\delta, t) = \frac{Y_{n_x(t),x} - t\lambda_x^2\mu_x}{\frac{\lambda_x^2}{\lambda_z}\delta}$$

where $Y_{n,x}$ is the sum of the first n samples from alternative x . Let $x \in \{1 \dots, k\}$, then

$$\mathcal{C}_x(\delta, \cdot) \Rightarrow W_x(\cdot)$$

as $\delta \rightarrow 0$ in the sense of $D[0, \infty)$, where W_x is a standard Brownian motion.

Proof. By Theorem 19.1 of Billingsley (1999),

$$\frac{Y_{n_x(t),x} - \text{floor}\left(\frac{\lambda_x^2}{\lambda_z^2} \left(\cdot \frac{1}{\delta^2}\right)\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z} \sqrt{\frac{1}{\delta^2}}} \Rightarrow W_x(\cdot)$$

in the sense of $D[0, \infty)$.

Fix $w \in \Omega$. Observe that

$$\frac{Y_{\text{floor}\left(\frac{\lambda_x^2}{\lambda_z^2} \left(\cdot \frac{1}{\delta^2}\right)\right),x} - \text{floor}\left(\frac{\lambda_x^2}{\lambda_z^2} \left(\cdot \frac{1}{\delta^2}\right)\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z} \sqrt{\frac{1}{\delta^2}}} - \frac{Y_{\text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2} \left(\cdot \frac{1}{\delta^2}\right)\right),x} - \text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2} \left(\cdot \frac{1}{\delta^2}\right)\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z} \sqrt{\frac{1}{\delta^2}}} \rightarrow 0$$

uniformly in $[0, s]$ for all $s \geq 0$ and then by Theorem A.2

$$\frac{Y_{\text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right), x} - \text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z^2} \sqrt{\frac{1}{\delta^2}}} \Rightarrow W_x(\cdot)$$

in the sense of $D[0, \infty)$.

Since $\frac{\frac{\lambda_x^2}{\lambda_z^2} t \frac{1}{\delta^2} - \text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2} t \frac{1}{\delta^2}\right)}{\frac{\lambda_x^2}{\lambda_z^2} \sqrt{\frac{1}{\delta^2}}} \rightarrow 0$ uniformly on $[0, s]$ for every $s \geq 0$, then by Theorem A.2

$$\frac{Y_{\text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right), x} - \left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z^2} \sqrt{\frac{1}{\delta^2}}} \Rightarrow W_x(\cdot).$$

Finally, observe that for fixed $\omega \in \Omega$,

$$\begin{aligned} & \frac{Y_{\text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right), x} - \left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z^2} \sqrt{\frac{1}{\delta^2}}} - \frac{Y_{\text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right) + n_0 \frac{\lambda_x^2}{\lambda_z^2}\right), x} - \left(n_0 \frac{\lambda_x^2}{\lambda_z^2} + \frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z^2} \sqrt{\frac{1}{\delta^2}}} \\ &= \frac{Y_{\text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right)\right), x} - Y_{\text{ceil}\left(\frac{\lambda_x^2}{\lambda_z^2}\left(\cdot, \frac{1}{\delta^2}\right) + n_0 \frac{\lambda_x^2}{\lambda_z^2}\right), x} + \left(n_0 \frac{\lambda_x^2}{\lambda_z^2}\right) \mu_x}{\frac{\lambda_x^2}{\lambda_z^2} \sqrt{\frac{1}{\delta^2}}} \\ &\rightarrow 0 \end{aligned}$$

uniformly in $[0, t]$ for all $t \geq 0$, and so by Theorem A.2 the result follows. \square

Now, we use the product topology in $D^k[0, \infty)$ for $k \in \mathbb{N}$. This topology may be described as the one under which $(Z_n^1, \dots, Z_n^k) \rightarrow (Z_0^1, \dots, Z_0^k)$ if and only if $Z_n^i \rightarrow Z_0^i$ for all $i \in \{1, \dots, k\}$. See the Miscellany of Billingsley (1968). The following corollary follows from the previous result and independence.

Corollary 1 We have that

$$\mathcal{C}(\delta, \cdot) := (\mathcal{C}_x(\delta, \cdot))_{x \in A} \Rightarrow W(\cdot) := (W_x(\cdot))_{x \in A}$$

as $\delta \rightarrow 0$ in the sense of D_∞^k .

Observe that $\mathcal{C}(\delta, \cdot) \in D[0, \infty)^k$ and the convergence of Lemma 1 is in $D[0, \infty)^k$, consequently the following natural step is to generalize the continuous-time procedure for the heterogeneous variance setting, see Appendix B, to the space $D[0, \infty)^k$. Specifically, for each $F \in D[0, \infty)^k$, we define $q_{tx}^{F, \delta}(A)$ as

$$q_{tx}^{F, \delta}(A) = q' \left(\left(F_x(t) \frac{\lambda_x^2}{\delta \lambda_z^2} + \frac{\lambda_x^2}{\lambda_z^2} \left(n_0 + \frac{t}{\delta^2} \right) \delta a_x, \frac{t}{\delta^2} : x \in A \right), \delta, A \subset \{1, \dots, k\} \right).$$

Note that if we replace F by $\mathcal{C}(\delta, t)$, we get $q_{tx}(A)$ in the previous equation. Using the continuous-time procedure for the heterogeneous variance setting with the paths $(q_{tx}^{F, \delta}(A) : t \geq 0)$, we see that there exists

a selection decision for each $F \in D[0, \infty)^k$. Since we are interested in the right selection, (i.e. in choosing k) for each $F \in D[0, \infty)^k$, we define the following function

$$f(F, \delta) = \begin{cases} 1 & \text{if } k \text{ is chosen} \\ 0 & \text{otherwise} \end{cases}$$

Observe that f is a function that depends in the path $\{q_{tx}^{F, \delta}(A) : t \geq 0\}$. We are doing all this because we want to see that $f(\mathcal{C}(\delta, \cdot), \delta)$ looks like the right or wrong selection decision of an algorithm that depends on a standard Brownian motion instead of $\mathcal{C}(\delta, \cdot)$, and this will give us the proof because we will be able to use the theory developed by Frazier (2014) to prove the theorem. Thus, the following step is to find the limit of $q_{tx}^{F, \delta}(A)$ when δ goes to zero, which is called $q_{tx}^F(A)$ and equal to

$$q_{tx}^F(A) := \exp\left(\frac{F_x(t)}{\lambda_z} + \frac{1}{\lambda_z^2} t a_x\right) / \sum_{x' \in A} \exp\left(\frac{F_{x'}(t)}{\lambda_z} + \frac{1}{\lambda_z^2} t a_{x'}\right).$$

Using the same previous idea, for each $F \in D[0, \infty)^k$, we define the new functions

$$g(F) = \begin{cases} 1 & \text{if } k \text{ is chosen} \\ 0 & \text{otherwise} \end{cases}.$$

So, we want to prove that

$$f(\mathcal{C}(\delta, \cdot), \delta) \Rightarrow g(W)$$

as $\delta \rightarrow 0$ in distribution.

In order to prove this, we first proved a continuity property in Lemma 2, which allows us to use Theorem 5.5 of Billingsley (1968), which implies the desired result. A proof of Lemma 2 may be found in a full version which will be submitted soon to arXiv.

Lemma 2 Let $\{\delta_n\} \subset (0, \infty)$ such that $\delta_n \rightarrow 0$. If $D_s \equiv \{Z \in D[0, \infty)^k : \text{if } \{Z_n\} \subset D[0, \infty)^k \text{ and } \lim_n d_\infty(Z_n, Z) = 0, \text{ then the sequence } \{f(Z_n, \delta_n)\} \text{ converges to } \{g(Z)\}, \text{ then } \mathbb{P}(W \in D_s) = 1.$

By the extension of the CMT (Billingsley 1968), we have the following corollary.

Corollary 2 We have that

$$f(C(\delta, t), \delta) \Rightarrow g(W(t))$$

in distribution as $\delta \rightarrow 0$.

Theorem 1 If samples from system $x \in \{1 \dots, k\}$ are identically distributed and independent, over time and across alternatives, then $\lim_{\delta \rightarrow 0} Pr\{\text{BIZ selects } k\} \geq P^*$ provided $\mu_k = a_k \delta, \mu_{k-1} = a_{k-1} \delta, \dots, \mu_1 = a_1 \delta$, $a_k > a_{k-1} \geq \dots \geq a_1$, $a_k - a_{k-1} > 1$, and the variances are finite and do not depend on δ .

Furthermore,

$$\inf_{a \in PZ(1)} \lim_{\delta \rightarrow 0} \mathbb{P}(CS_\delta) = P^*$$

where $PZ(1) = \{a \in \mathbb{R}^k : a_k - a_{k-1} > 1, a_k > a_{k-1} \geq \dots \geq a_1\}$.

Proof. By the comments given at the beginning of this section, we know that we can work with the algorithm defined by

$$q'_{tx}(A) = q' \left(\left(Z_{\frac{t}{\delta^2}x}, \frac{t}{\delta^2} : x \in A \right), \delta, A \subset \{1, \dots, k\} \right)$$

where $t \in \delta^2 \mathbb{N}$, instead of the algorithm defined by

$$q'_{tx}(A) = q'((Z_{tx}, t : x \in A), \delta, A \subset \{1, \dots, k\})$$

where $t \in \mathbb{N}$.

Now, we define

$$\hat{\tau}_{n+1}(\delta) = \inf \left\{ t \in \{ \tau_n \delta^2, (\tau_n + 1) \delta^2, \dots \} : \min_{x \in A_n} q'_{t/\delta^2, x}(A_n) \leq c \text{ or } \max_{x \in A_n} q'_{t/\delta^2, x}(A_n) \geq P_n \right\}$$

We denote the corresponding continuous hitting times by $(\tau_n(\delta))_n$, which are defined as

$$\tau_{n+1} = \inf \{ t \geq \tau_n : \min_{x \in A_n} q_{tx}(A_n) \leq c \text{ or } \max_{x \in A_n} q_{tx}(A_n) \geq P_n \}.$$

Using that $\mathcal{C}(\delta, \cdot)$ is right-continuous, we can prove that $\hat{\tau}_n(\delta) - \tau_n(\delta) \rightarrow 0$ with probability 1 as $\delta \rightarrow 0$, and so we can use $\mathcal{C}(\delta, \tau_n(\delta))$ instead of $\mathcal{C}(\delta, \hat{\tau}_n(\delta))$ in the limit.

Let CS_δ be the event of doing a correct selection given the configuration $\mu_k = a_k \delta, \mu_{k-1} = a_{k-1} \delta, \dots, \mu_1 = a_1 \delta$. Then by the previous argument and the Corollary 2,

$$\begin{aligned} \lim_{\delta \rightarrow 0} \mathbb{P}(CS_\delta) &= \lim_{\delta \rightarrow 0} \mathbb{P}(f(\mathcal{C}(\delta, t), \delta) = 1) \\ &= \mathbb{P}(g(W) = 1) \\ &\geq P^* \end{aligned}$$

where the last inequality follows from the Theorem 2 of Frazier (2014).

Furthermore, by the same theorem 2,

$$\inf_{a \in PZ(1)} \mathbb{P}(g(W) = 1) = P^*$$

where $PZ(1) = \{a \in \mathbb{R}^k : a_k - a_{k-1} \geq 1\}$.

□

5 NUMERICAL EXPERIMENTS

We show the performance of the BIZ procedure in discrete time with maximum elimination ($c = 1 - (P^*)^{\frac{1}{k-1}}$) when δ goes to zero. We plot both the PCS and the expected total number of samples simulated divided by the number of systems, denoted by $E[N]/k$. We consider this quotient to normalize the number of samples used by the algorithm.

First, we consider the known variance case. Specifically, we consider a slippage configuration with 100 systems normally distributed, where $\mu_k = \delta, \mu_{k-1} = 0, \dots, \mu_1 = 0$, δ is within the interval $[0.1, 10]$, and $\lambda_{100} = 1, \lambda_{99} = 1 + \frac{(0.5)(98)}{99}, \dots, \lambda_1 = 0.5$. Here, $P^* = 0.9$ and $n_0 = 0$. Figure 1 shows that in this case the IZ guarantee is always satisfied, and the bound on probability of correct selection is tight when δ goes to zero. When δ is big enough, the PCS is almost one because the difference between the best system and the others is large enough to be easily identifiable by the BIZ procedure.

We now consider the unknown variance case. Specifically, we consider a slippage configuration with 100 systems normally distributed, where $\mu_{100} = \delta, \mu_{99} = 0, \dots, \mu_1 = 0$, δ is within the interval $[0.1, 10]$, and $\lambda_{100} = 10, \lambda_{99} = \dots = \lambda_1 = 1$. Here, $P^* = 0.9$ and $n_0 = 15$. Figure 2 gives us evidence that our theorem should also be true when the variances are unknown. We should note that this is a difficult example, and

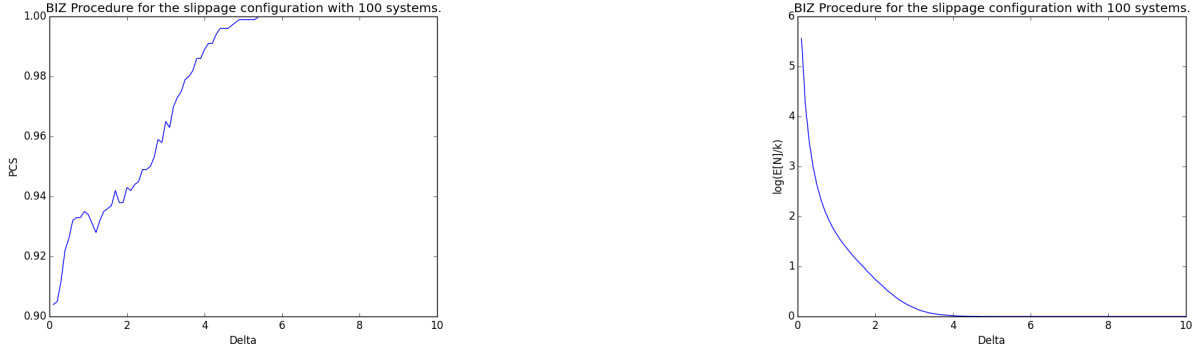


Figure 1: Known variances and $P^* = 0.9, n_0 = 0$. In this example, our theorem is true: the IZ guarantee is always satisfied and the inequality is tight as δ goes to zero.

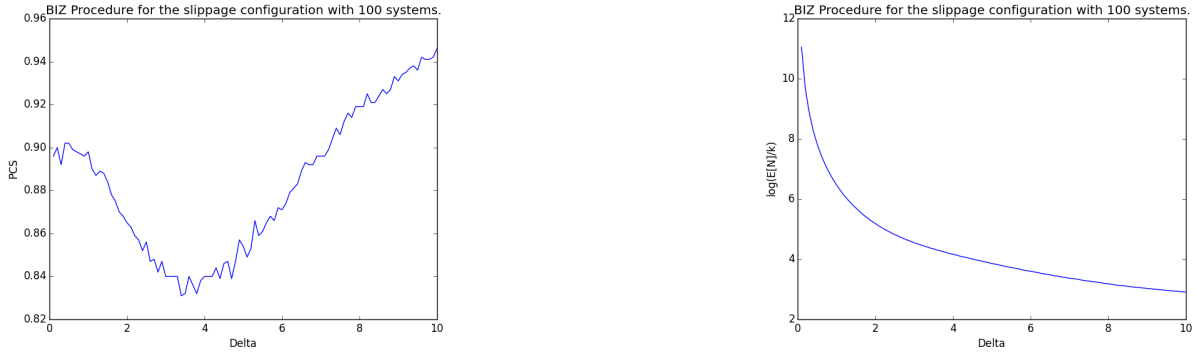


Figure 2: Unknown variances and $P^* = 0.9, n_0 = 15$. In this example, our theorem is true: the IZ guarantee is satisfied and the inequality is tight as delta goes to zero. If we want that the IZ guarantee be satisfied in all the cases, we can just increment n_0 .

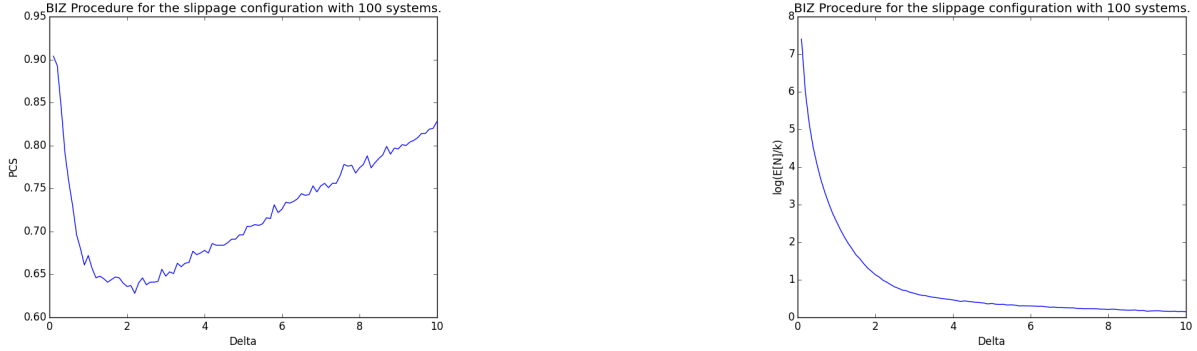


Figure 3: Known variances, $P^* = 0.9$, . This was a hard example to find and it was specially chosen to illustrate our theorem. In practice, we would choose $n_0 > 0$ and the IZ guarantee will be satisfied.

still we have a good performance in most of the cases. However, if we want to always satisfy the IZ guarantee, we can just increment the parameter n_0 .

Now, we consider a pathological case when the variances are known. Specifically, we consider a slippage configuration with 100 systems normally distributed, where $\mu_{100} = \delta, \mu_{99} = 0, \dots, \mu_1 = 0$, δ is within the interval $[0.1, 10]$, and $\lambda_{100} = 10, \lambda_{99} = \dots = \lambda_1 = 1$. Here, $P^* = 0.9$ and $n_0 = 0$. This configurations was specially chosen to illustrate our theorem, and so it is harder than typical configurations. In fact, in most of the configurations the algorithm always satisfies the IZ guarantee. Furthermore, in practice, we would have run the algorithm with $n_0 > 0$, and the BIZ algorithm would have worked well even in this pathological example. Figure 3 shows that PCS converges to 0.9 as δ goes to zero. Specifically, PCS is equal to 0.904 when $\delta = 0.1$. This figure also shows that the bound is tight on probability of correct selection. We should also note that the number of samples required increases very fast. Here, the points plotted have central confidence intervals of length at most 0.014.

6 CONCLUSION

We have proved the asymptotic validity of the Bayes-inspired Zone procedure (Frazier 2014) when the variances are known, which is a new sequential elimination procedure. This algorithm is relevant because it takes fewer samples than other IZ procedures, especially for problems with large numbers of alternatives. Even though this proof does not guarantee that the algorithm will work for any sample, we know that it will work if the alternatives are not very different, which are the most difficult cases.

ACKNOWLEDGMENTS

Peter Frazier and Saul Toscano-Palmerin were partially supported by NSF CAREER CMMI-1254298, AFOSR FA9550-12-1-0200, and AFOSR FA9550-15-1-0038. Peter Frazier was also partially supported by NSF IIS-1247696 and the Atkinson Center for a Sustainable Future Academic Venture Fund. Saul Toscano-Palmerin was also partially supported by the Mexican Secretariat of Public Education (SEP).

REFERENCES

- Bechhofer, R. E. 1954. “A Single-Sample multiple-decision procedure for selecting the multinomial event which has the highest probability”. *The Annals of Mathematical Statistics* 25 (1): 16–39.
- Bechhofer, R. E., J. Kiefer, and M. Sobel. 1968. *Sequential Identification and Ranking Procedures*. Chicago: University of Chicago Press.
- Billingsley, P. 1968. *Convergence of Probability Measures*. New York: John Wiley and Sons.
- Billingsley, P. 1999. *Convergence of Probability Measures*. 2nd ed. New York: John Wiley and Sons.

- Chen, C.-H., and L. H. Lee. 2011. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*, Volume 1. Singapore: World Scientific.
- Dieker, A., and S.-H. Kim. 2012. "Selecting the best by comparing simulated systems in a group of three when variances are known and unequal". In *Proceedings of the 2012 Winter Simulation Conference*, edited by IEEE.
- Fabian, V. 1974. "Note on Anderson's sequential procedures with triangular boundary". *The Annals of Mathematical Statistics* 2 (1): 170–176.
- Frazier, P. 2012. "TUTORIAL: OPTIMIZATION VIA SIMULATION WITH BAYESIAN STATISTICS AND DYNAMIC PROGRAMMING". *Proceedings of the 2012 Winter Simulation Conference*.
- Frazier, P. I. 2014. "A Fully Sequential Elimination Procedure for Indifference-Zone Ranking and Selection with Tight Bounds on Probability of Correct Selection". *Operations Research*.
- Goldsman, D., S. Kim, W. Marshall, and B. Nelson. 2002. "Ranking and selection for steady-state simulation: Procedures and perspectives". *INFORMS Journal on Computing* 14 (1): 2–19.
- Hartmann, M. 1988. "An improvement on Paulson's sequential ranking procedure". *Sequential Analysis* 7 (4): 363–372.
- Hartmann, M. 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* 10:1–16.
- Hong, J. 2006. "Fully sequential indifference-zone selection procedures with variance-dependent sampling". *Naval Research Logistics* 53 (5): 464–476.
- Kim, S., and B. Nelson. 2006a. "Selecting the best system". In *Handbook in Operations Research and Management Science: Simulation*, edited by e. S.G. Henderson, B.L. Nelson, 501–534. Elsevier, Amsterdam.
- Kim, S., and B. Nelson. 2007. "Recent advances in ranking and selection". In *Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come.*, edited by I. Press, 162–172. Piscataway, New Jersey.
- Kim, S.-H., and A. Dieker. 2011. "Selecting the best by comparing simulated systems in a group of three.". *Proceedings of the 2011 Simulation Conference*:3987–3997.
- Kim, S.-H., and B. L. Nelson. 2001. "A fully sequential procedure for indifference-zone selection in simulation". *ACM Trans. Model. Comput. Simul.* 11 (3): 251–273.
- Kim, S.-H., and B. L. Nelson. 2006b. "On the Asymptotic Validity of Fully Sequential Selection Procedures for Steady-State Simulation". *Operations Research* 54 (3): 475–488.
- 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).
- R. E. Bechhofer, T.J. Santner, D. G. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons*. New York: J.Wiley and Sons.
- Rinott, Y. 1978. "On Two-Stage Selection Procedures and Related Probability-Inequalities". *Communications in Statistics-Theory and Methods* 7 (8): 799–811.
- Swisher, J., S. Jacobson, and E. Yücesan. 2003. "Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey". *ACM Transactions on Modeling and Computer Simulation* 13 (2): 134–154.
- Wang, H., and S.-H. Kim. 2013. "Reducing the conservativeness of fully sequential indifference-zone procedures". *IEEE Transactions on automatic control* 58 (6): 1613–1619.

AUTHOR BIOGRAPHIES

SAUL TOSCANO-PALMERIN is a Ph.D. student of the School of Operations Research and Information Engineering at Cornell University, Ithaca, NY. His research interest is in simulation optimization,

machine learning and sequential decision-making under uncertainty. His email address is st684@cornell.edu.

PETER I. FRAZIER is an Assistant Professor of the School of Operations Research and Information Engineering at Cornell University, Ithaca, NY. He holds a Ph.D. in operations research and financial engineering industrial engineering from Princeton University. His research interests include optimal learning, sequential decision-making under uncertainty, and machine learning, focusing on applications in simulation optimization, design of experiments, materials science, e-commerce and medicine. He is a member of INFORMS, and SIAM. His e-mail address is pf98@cornell.edu.