Selecting the Best System Using the Three Stage and the Four-Stage Selection Approaches\textsuperscript{1}

Mohammad H. Almomani

School of Mathematical Sciences
Universiti Sains Malaysia
11800 Penang, Malaysia
mh_momani@yahoo.com

Rosmanjawati Abdul Rahman

School of Mathematical Sciences
Universiti Sains Malaysia
11800 Penang, Malaysia
rosmanjawati@usm.my

Abstract

Statistical selection approaches are used to select the best stochastic system from a finite set of alternatives. The best system will be the system with minimum or maximum performance measure. We consider the problem of selecting the best system when the number of alternative systems is huge. Three-Stage and Four-Stage selection approaches are proposed to solve this problem. The main strategy in these two selection approaches involves a combination method of cardinal and ordinal optimization. Ordinal optimization procedure is used to reduce the number of systems in the search space such that to be appropriate for cardinal optimization procedures. Three-Stage selection approach consists three procedures; Ordinal Optimization, Subset Selection and Indifference-Zone. While, Four-Stage selection approach consists four procedures; Ordinal Optimization, Optimal Computing Budget Allocation, Subset Selection and Indifference-Zone. In this paper, we compare the performance between the two selection approaches; the Three-Stage and Four-Stage approaches. The numerical results show that the performance of Four-Stage selection approach is better compare to Three-Stage selection approach in different aspects.

\textsuperscript{1}This work is partly supported by Universiti Sains Malaysia Short Term Grant (304/PMATHS/639045).
1 Introduction

We consider the following optimization problem

\[ \min_{\theta \in \Theta} J(\theta) \]  

where \( \Theta \) is an arbitrary feasible solution set which is finite and huge. Let \( J \) be the expected performance measure of some complex simulation system, written as \( J(\theta) = E[L(\theta, Y)] \), where \( \theta \) is a vector representing the system design parameters, \( Y \) represents all the random effect of the system and \( L \) is a deterministic function that depends on \( \theta \) and \( Y \). In this paper, without loss of generality we assume the best system is the system that has the smallest mean, which is unknown and to be inferred from simulation. Therefore, our goal is selecting the system that has the smallest sample mean. Suppose that there are \( n \) systems, and let \( Y_{ij} \) represents the \( j^{th} \) output from the system \( i \), where \( i = 1, 2, \ldots, n \). Let \( Y_i = \{Y_{ij}, j = 1, 2, \ldots\} \) denote the output sequence from the system \( i \). We assume \( Y_{ij} \) are independent and identically normal distributed with unknown means \( \mu_i = E(Y_{ij}) \) and variances \( \sigma^2_i = Var(Y_{ij}) \), and the \( Y_1, Y_2, \ldots, Y_n \) are mutually independent. In practice, the \( \sigma^2_i \) are unknown, so we estimate it using the sample variances \( s^2_i \) for \( Y_{ij} \). However, since we assume that the smallest mean is better, therefore if the ordered \( \mu_i \) values are denoted by \( \mu_{[1]} \leq \mu_{[2]} \leq \ldots \leq \mu_{[n]} \), then the system having mean \( \mu_{[1]} \) is referred to as the best system. Clearly, the Correct Selection (CS) occurs when the system selected by the selection approach is the same as the actual best system.

Ranking and Selection (R&S) procedures are used usually to select the best system or a subset that contain the best systems when the number of alternatives is small, with a pre specified significance level, see Kim and Nelson [24]. Problem occurs for a large scale problem when it needs a huge computational time. Since for each sample of the performance value require one simulation run, therefore, for a large scale problem will require a large number of samples. This is very time-consuming and may be impossible. In this situation, we would change our objective to finding good systems rather than estimating accurately the performance value for these systems. This idea lies in Ordinal Optimization (OO) procedure. The objective of OO procedure is to isolate a subset of good systems with high probability and at the same time, to reduce
Three-stage and four-stage approaches

the required simulation time, see Ho et al. [22]. Since the OO procedure can reduce computation cost, the use of Optimal Computing Budget Allocation (OCBA) can further reduce the cost by allocate the computing budget among different systems instead of allocating it among equally simulated systems, see Chen et al. [11].

There are two measures of selection quality; the Probability of Correct Selection ($P(CS)$) and the Expected Opportunity Cost ($E(OC)$) of a potentially incorrect selection, see He et al. [21]. These two measures are important, because in the selection problems we are using the simulation method to get the performance measure for each alternative system. This will cause a potential for incorrect selection. So we need these two measures to determine the quality of the selection. The traditional selection approaches use the $P(CS)$ as a measure of selection quality, with selecting the best system with high values of $P(CS)$ as their goal. However, the performance for these selection approaches become high by maximizing the $P(CS)$. On the other hand, another measure of selection quality, the $E(OC)$ has become important in business and engineering applications, that leads to the studies to reduce the opportunity cost of a potentially incorrect selection. For more details of $E(OC)$, see Gupta and Miescke [19, 20], Chick and Inoue [12, 13].

In this paper, we review two selection approaches for solving problem in (1). First, we discuss the Three-Stage (3-Stage) approach as proposed by Almomani and Alrefaei [3] in context of the $P(CS)$ as a measure of selection quality. The 3-Stage approach consists of three stages. In the first stage, OO procedure is used to select a subset $G$ randomly from the feasible solution set, where the probability that $G$ contains one of the best $m\%$ alternatives is high. In stage two, a Subset Selection ($SS$) procedure is applied on the subset $G$ to select a subset $I$ that contains the best system with high probability. Next, in stage three, Indifference-Zone ($IZ$) procedure is used on set $I$ to select the best system. We also discuss the 3-Stage approach in context of the $E(OC)$ as discussed in Almomani et al. [4].

The second selection approach as proposed in Almomani and Abdul Rahman [1] is a Four-Stage (4-Stage) involved with four procedures the OO, OCBA, SS and IZ. In the first step, the OO is used to select a subset that overlap with the set of the actual best $m\%$ system. Then, OCBA procedure is used to allocate the available simulation samples in a way that maximize the probability of correct selection. This is followed by SS procedure to get a smaller subset that contains the best system among the subset that was selected before. Finally, using the IZ procedure to select the best system among the survivors systems in the previous stage. Extension from that, we also study the 4-Stage approach in context of the $E(OC)$ as discussed in Almomani and Abdul Rahman [2]. We also make a comparison on the performance of 3-Stage and 4-Stage selection approaches from different aspects especially
on their elapsed (execution) time, number of simulation samples and measures of selection quality.

The rest of this paper is organized as follows; In Section 2, we give a background for the OO with OCBA, and R&S procedures. Both algorithms for 3-Stage and 4-Stage selection approaches are discussed in Section 3. This is followed by a numerical example in Section 4. Finally, concluding remarks are presented in Section 5.

2 Background

In this section, we briefly review the Ordinal Optimization with Optimal Computing Budget Allocation (OO + OCBA) procedures, and Ranking and Selection (R&S) procedures.

2.1 Ordinal Optimization and Optimal Computing Budget Allocation

The OO’s goal is isolating a subset of good systems with high probability and reducing the required simulation time for discrete event simulation. The aim of this procedure, as proposed by Ho et al. [22] is to find good systems, rather than estimating the performance value of these systems accurately. Therefore, OO procedure is used to select a subset that overlaps with the set of the actual best $m\%$ systems with high probability.

Suppose the correct selection is to select a subset $G$ of $g$ systems from a feasible solution set $\Theta$, that contains at least one of the top $m\%$ best systems. Since $\Theta$ is very huge, then the probability of correct selection is given by $P(CS) \approx (1 - (1 - \frac{m}{100})^g)$. Furthermore, suppose that the correct selection is to select a subset $G$ of $g$ systems that contains at least $r$ of the best $s$ systems. If we assume $S$ to be the subset that contains the actual best $s$ systems, then the probability of correct selection can be obtained using the hyper geometric distribution as

$$P(CS) = P(|G \cap S| \geq r) = \sum_{i=r}^{g} \binom{r}{i} \binom{s}{g-i} \binom{n-s}{g-i}.$$

Since the number of alternatives is very large then the $P(CS)$ can be approximated by the binomial random variable, as $P(CS) \approx \sum_{i=r}^{g} \binom{r}{i} \left(\frac{m}{100}\right)^i \left(1 - \frac{m}{100}\right)^{g-i}$. More details of OO can be found in Deng et al. [16], Dai [14], Xiaolan [33], Deng and Ho [15], Lee et al. [27], Li et al. [28], Zhao et al. [34] and Ho et al. [23].

Meanwhile, the OCBA technique was proposed to improve the performance of OO procedure by determining the optimal numbers of simulation samples for each system, instead of equally simulating all systems. The OCBA used to determine the best simulation lengths for all simulation systems to reduce the total computation time. The goal of OCBA is to allocate the total simulation samples from all systems in a way that maximizes the probability of selecting
the best system within a given computing budget, see Chen et al.[11], Chen [7], Chen et al.[10], Chen et al. [9], Banks [5] and Chen [8]. Let \( B \) be the total sample that available for solving the optimization problem given in (1). Our target is to allocate these computed simulating samples to maximize the \( P(CS) \). In mathematical notation

\[
\max_{T_1, \ldots, T_n} P(CS)
\]

s.t. \( \sum_{i=1}^{n} T_i = B \)

\( T_i \in \mathbb{N} \quad i = 1, 2, \ldots, n \)

where, \( \mathbb{N} \) is the set of non-negative integers, \( T_i \) is the number of samples allocated to system \( i \), and \( \sum_{i=1}^{n} T_i \) denotes the total computational samples, and assuming that the simulation times for different systems are roughly the same. To solve this problem Chen et al. [11] proposed the following theorem.

**Theorem 2.1.** Given a total number of simulated samples \( B \) to be allocated to \( n \) competing systems whose performance is depicted by random variables with means \( J(\theta_1), J(\theta_2), \ldots, J(\theta_n) \), and finite variances \( \sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2 \) respectively, as \( B \to \infty \), the approximate probability of correct selection can be asymptotically maximized when

1. \( \frac{T_i}{T_j} = \left( \frac{\sigma_i}{\sigma_j} \right)^2 \); where \( i, j \in \{1, 2, \ldots, n\} \) and \( i \neq j \neq b \).

2. \( T_b = \sigma_b \sqrt{\sum_{i=1}^{n} \frac{T_i^2}{\sigma_i^2}} \)

where \( \delta_{bh} \) the estimated difference between the performance of the two systems \( (\delta_{bh} = \bar{T}_b - \bar{T}_h) \), and \( \bar{T}_h \leq \min \bar{T}_h \) for all \( i \). Here \( \bar{T}_i = \frac{1}{T_i} \sum_{j=1}^{T_i} Y_{ij} \), where \( Y_{ij} \) is a sample from \( Y_{i} \) for \( j = 1, \ldots, T_i \).

**Proof:** See Chen et al. [11].

### 2.2 Ranking and Selection

Selecting the system with the smallest or largest expected performance is one of major problems that arise in simulation. When the number of alternative systems is small, we can use R&S procedures to select the best system or a subset that contain the best systems. There are two different R&S procedures; the Indifference-Zone (IZ) and Subset Selection (SS).

The IZ procedures goal is selecting the best system among \( n \) systems when the number of alternatives less than or equal 20. Suppose we have \( n \) alternative systems that are normally distributed with unknown means \( \mu_1, \mu_2, \ldots, \mu_n \), and
suppose that these means are ordered as \( \mu_1 \leq \mu_2 \leq \cdots \leq \mu_n \). We want to select the system that has the best minimum mean \( \mu_1 \). The IZ is defined to be the interval \([\mu_1, \mu_1 + \delta^*]\), where \( \delta^* \) is a predetermined small positive real number and is called indifference zone. We are interested in selecting an alternative \( i^* \) such that \( \mu_{i^*} \in [\mu_1, \mu_1 + \delta^*] \). Let the CS is selecting an alternative whose mean belongs to the indifference zone. We prefer the CS to take place with high probability, say with a probability not smaller than \( P^* \) where \( 1/n \leq P^* \leq 1 \). For this problem, Rinott [29] has presented a procedure that is applicable when the data are normally distributed and all systems are simulated independently of each others. Meanwhile, Tamhane and Bechhofer [31] have presented a simple procedure that is valid when variances may not be equal.

On the other hand, when the number of alternatives are relatively large SS procedure can be used to solve the selection problems. This procedure will screen out the feasible solution set and eliminate non-competitive systems to construct a subset that contains the best system with high probability. The SS procedure required that \( P(CS) \geq P^* \), where the CS is selecting a subset that contains the actual best system, with \( P^* \) as a predetermined probability. This SS procedure was dating back to Gupta [18], who presented a single stage procedure for producing a subset containing the best system with a specified probability. Extension of this work which is relevant to the simulation settings include Sullivan and Wilson [30] who derived a two stage SS procedure to determine a subset of maximum size \( m \) that, with a specified probability will contain systems that are all within a pre-specified amount of the optimum. Another comprehensive review of R&S procedures can be found in Bechhofer et al. [6], Goldsman and Nelson [17] and Kim and Nelson [24, 25, 26].

3 Three-Stage and Four-Stage Selection Approaches

In this section, we review two selection approaches that are proposed to select a good stochastic system with high probability when the number of elements in the feasible solution set is huge. First approach is the Three-Stage (3-Stage) as proposed in Almomani and Alrefaei [3] and Almomani et al. [4]. They discussed the approach in context of the two measures of selection quality; the \( P(CS) \) and \( E(OC) \). Whereas the second approach is called the Four-Stage (4-Stage) as discussed in Almomani and Abdul Rahman [1] and Almomani and Abdul Rahman [2]. They also proposed the approach in context of the \( P(CS) \) and \( E(OC) \).
3.1 Three-Stage approach

The 3-Stage approach as presented in Almomani and Alrefaei [3] have been discussed in context of $P(\text{CS})$. They proved that, this selection approach works practically well, in order to select the best system with high $P(\text{CS})$. Meanwhile, Almomani et al. [4] showed that the 3-Stage approach provides the minimum value of $E(\text{OC})$ as a potentially incorrect selection. This approach consists three stages; in the first stage, a subset $G$ is selected randomly from the feasible solution set. Let $g$ denotes the size of the subset $G$, where $g$ is a relatively small number, but the probability that $G$ contains one of the best $m\%$ alternatives is high $(1 - \alpha_1)$. Then, the SS procedure is applied on the subset $G$ to select a subset $I$ that contains the best system with high probability $(1 - \alpha_2)$, where $|I| \leq 20$. Note that, in order to use the SS procedure the size of the subset $G$ should be relatively small. Finally, the IZ procedure is applied on the set $I$ to select the best system with the probability as $(1 - \alpha_3)$. Note also that, the size of set $I$ is less than or equal 20, so that we can use the IZ procedure. The algorithm is described in Almomani and Alrefaei [3] as follows:

**Algorithm (3-Stage):**

**Step 0:** Determine the number of alternatives $g$ in $G$, i.e. $|G| = g$, where $G$ is the subset that will be selected in the first stage that satisfies $P(G$ contains at least one of the best $m\%$ system $) \geq 1 - \alpha_1$. Determine the number of samples $t_0 \geq 2$, and the indifference zone $\delta^*$, $t = t_{(1 - \alpha_2/2)^{1/2}}$, from the $t$-distribution.

**Step 1:** Select a subset $G$ of size $g$ randomly from the feasible solution set $\Theta$. Take random samples of $t_0$ observations $y_{ij}$ ($1 \leq j \leq t_0$) for each system $i$, $i = 1, \ldots, g$.

**Step 2:** Calculate the first stage sample means and variances $\bar{y}_i^{(1)}$ and $s_i^2$ for all $i = 1, \ldots, g$

$$\bar{y}_i^{(1)} = \frac{\sum_{j=1}^{t_0} y_{ij}}{t_0} \quad \text{and} \quad s_i^2 = \frac{\sum_{j=1}^{t_0} (y_{ij} - \bar{y}_i^{(1)})^2}{t_0 - 1}$$

**Step 3:** Set $I = \{i : 1 \leq i \leq g \text{ and } \bar{y}_i^{(1)} \leq \bar{y}_j^{(1)} - [W_{ij} - \delta^*], \forall i \neq j\}$, where $W_{ij} = t \left( \frac{s_i^2}{t_0} + \frac{s_j^2}{t_0} \right)^{1/2}$ for all $i \neq j$, and $[x]^* = x$ if $x < 0$ and $[x]^* = 0$ otherwise.

**Step 4:** If $I$ contains a single index, then this system is the best system. Otherwise, for all $i \in I$ compute the second stage sample size $N_i = \max\{t_0, \lceil \frac{h t_0}{\delta^*} \rceil \}$, where $h = h(1 - \alpha_3/2, t_0, |I|)$ be the Rinott [29] constant and can be obtained from tables of Wilcox [32].
Step 5: Take a random sample of $N_i - t_0$ additional observations for all systems $i \in I$ and compute the overall sample means $\bar{y}_i^{(2)} = \frac{\sum_{j=1}^{N_i} y_{ij}}{N_i}$ for all $i \in I$.

Step 6: Select the system $i \in I$ with the smallest $\bar{y}_i^{(2)}$ as the best system.

3.2 Four-Stage approach

The 4-Stage approach as proposed by Almomani and Abdul Rahman [1] is used to select the best stochastic system when the number of alternatives is very large. Almomani and Abdul Rahman [1] showed that this approach achieved the correct selection with high $P(CS)$ and with a small number of simulation samples. Furthermore, Almomani and Abdul Rahman [2] also proved that this approach selected the best system with the minimum $E(OC)$ of a potentially incorrect selection. This selection approach consists of four stages. In the first stage, using the $OO$ procedure, a subset $G$ is selected randomly from the feasible solution set that intersects with the set $m\%$ of actual best systems with high probability $(1 - \alpha_1)$. We use the $OCBA$ procedure in the second stage to allocate the available computing budget. This will follow by the $SS$ procedure to get a smaller subset $I$ that contains the best system among the subset that is selected before, with high probability $(1 - \alpha_2)$, where $|I| \leq 20$. Finally, using the $IZ$ procedure to select the best system from set $I$ with high probability $(1 - \alpha_3)$. The algorithm of this selection approach is described as follows:

Algorithm (4-Stage):

Step 0: Determine $g$ and $k$ where $|G| = g$ and $|G'| = k$. Here, $G$ is defined as the selected subset from $\Theta$, that satisfies $P(G \text{ contains at least one of the best } m\% \text{ systems}) \geq 1 - \alpha_1$, and $G'$ is defined as the selected subset from $G$, where $g \geq k$. Let the number of initial simulation samples $t_0 \geq 2$, and $t = t \left(1 - \alpha_2/2\right)^{-1/2} t_0$, where $l$ is the iteration number, and determine the total computing budget $B$.

Step 1: Select a subset $G$ of size $g$ randomly from $\Theta$. Take a random sample of $t_0$ observations $y_{ij}$ $(1 \leq j \leq t_0)$ for each system $i$ in $G$, where $i = 1, \ldots, g$.

Step 2: Calculate the sample mean $\bar{y}_i^{(1)}$ and variances $s_i^2$ as

$$\bar{y}_i^{(1)} = \frac{\sum_{j=1}^{T_i} y_{ij}}{T_i} \quad \text{and} \quad s_i^2 = \frac{\sum_{j=1}^{T_i} (y_{ij} - \bar{y}_i^{(1)})^2}{T_i - 1}$$

where $i = 1, \ldots, g$. 
Step 3: Order the systems in $G$ according to their sample averages

\[ \bar{y}_1^{(1)} \leq \bar{y}_2^{(1)} \leq \ldots \leq \bar{y}_g^{(1)} \]

Step 4: Select the best $k$ systems from the set $G$, and represent this subset as $G'$.

Step 5: If $\sum_{i=1}^{g} T_i^l \geq B$, then go to Step 8. Otherwise, select randomly a subset $G''$ of the $g - k$ alternatives from $\Theta - G'$, let $(G = G' \cup G'')$.

Step 6: Increase the computing budget by $\Delta$ and compute the new budget allocation, $T_1^{l+1}, T_2^{l+1}, \ldots, T_g^{l+1}$, by using Theorem 2.1.

Step 7: Perform additional $\max\{0, T_i^{l+1} - T_i^l\}$ simulations for each system $i, i = 1, \ldots, g$, let $l \leftarrow l + 1$. Go to Step 2.

Step 8: Set $I = \{i : 1 \leq i \leq k \text{ and } \bar{y}_i^{(1)} \leq \bar{y}_j^{(1)} - [W_{ij} - \delta^*], \forall i \neq j\}$, where

\[ W_{ij} = t \left( \frac{s_i^2}{T_i} + \frac{s_j^2}{T_j} \right)^{1/2} \]

for all $i \neq j$, and $[x]^\pm = x$ if $x < 0$ and $[x]^\pm = 0$ otherwise.

Step 9: If $I$ contains a single index, then this system is the best system. Otherwise, for all $i \in I$, compute the second sample size $N_i = \max\{T_i, \lceil (h \alpha_3)^2 \rceil\}$, where $h = h(1 - \alpha_3/2, t_0, |I|)$ be the Rinott [29] constant and can be obtained from tables of Wilcox [32].

Step 10: Take additional $N_i - T_i$ random samples of $y_{ij}$ for each system $i \in I$, and compute the overall sample means for $i \in I$ as $\bar{y}_i^{(2)} = \frac{\sum_{j=1}^{N_i} y_{ij}}{N_i}$.

Step 11: Select system $i \in I$ with the smallest $\bar{y}_i^{(2)}$ as the best system.

4 Numerical example

As example, we consider the $M/M/1$ queuing systems where the inter arrival times and the service times are exponentially distributed and the system has one server. Our goal is selecting one of the best $m\%$ systems that has the minimum average waiting time per customer from $n$ $M/M/1$ queuing systems, where $n$ is a large number. As a measure of selection quality, we use the $P(CS)$ and the $E(OC)$ of a potentially incorrect selection. We define the Opportunity Cost ($OC$) as the difference between unknown means of the selected best system and the actual best system. Applying the two of selection approaches; the 3-Stage and 4-Stage under some assumptions, then we make a comparison between these selection approaches regarding their simulation elapsed time, number of simulation samples, the $P(CS)$ and the $E(OC)$. 
Let the arrival rate $\lambda$ is a fixed number, and the service rate $\mu$ is belong to the interval $[a, b]$ such that $\lambda = 1$ and $\mu \in [7, 8]$. Suppose we have 3000 of $M/M/1$ queuing systems, and we discretize the problem by assuming that $\mu \in \Theta = 7 + i/3000$, where $i = 0, 1, \ldots, 3000$. Therefore, the best queuing system that has a minimum average waiting time, would be the 3000th queuing system with $\mu_{3000} = 8.0$. In the first experiment, we apply the 3-Stage and 4-Stage approaches with assumptions that $n = 3000, g = 200, \alpha_2 = \alpha_3 = 0.005, \delta^* = 0.05, \text{and } t_0 = 20$. Consider that we want to select one of the best (1%) systems, so our target would be the systems from 2971 to 3000. The correct selection would be selecting the system that belongs to $\{2971, 2972, \ldots, 3000\}$. For this, and the analytical probability of the correct selection can be calculated as $P(CS) \geq 1 - \left(1 - \frac{1}{100}\right)^{200} + 0.005 + 0.005 \geq 0.85$.

Table 1 and Table 2 contain the results of this experiment for the first 10 replications. From the tables, “Best” means the index of the system that has been selected by the selection approaches as the best system, $|I|$ is the number of systems from Step 3 and Step 8 for each the 3-Stage and 4-Stage algorithms respectively, $\sum_{i=1}^{g} T_i$ is the total sample size in Step 5 for the 4-Stage algorithm, which is also representing the total initial sample size in the 3-Stage algorithm. Meanwhile, $\sum_{i \in I} N_i$ is the total sample size in Step 4 and Step 9 for each in the 3-Stage and 4-Stage algorithms respectively. Column “Time” represents the elapsed (execution) time of the programme (method), in millisecond (ms), where $ms = 1/1000$ second. The “Approach $E(OC)$” represents by $E(\bar{y}_b - \bar{y}_*)$, whereas the “Analytical $E(OC)$” is equal to $E(\bar{w}_b - \bar{w}_*)$. The $\bar{w}_*$ and $\bar{y}_*$ each represent the unknown average waiting time and the sample mean respectively for the actual best system $i^*$, where $i^* = 3000$ and $\bar{w}_* = 0.142857143$. Whereby $\bar{w}_b$ and $\bar{y}_b$ for each is the unknown average waiting time and the sample mean respectively for the selected best system $b$. Note that, we can calculate the $\bar{w}_*$ and $\bar{w}_b$ using formula $\bar{w}_i = \frac{1}{\mu_i - \lambda_i}$; where $i = \{i^*, b\}$, and $\mu_i$ and $\lambda_i$ are respectively the service rate and the arrival rate for the system $i$. After the simulation is performed, $\bar{y}_b$ can be calculated according to the system output. In particular, in these two tables (Table 1 and Table 2) we take the absolute value of the difference between the sample mean for the best system and the actual best system for each the 3-Stage and 4-Stage approaches and report them as in the seventh column “Approach $E(OC)$”.

Clearly, we can see that the 4-Stage approach works well under these assumption settings, compared to the 3-Stage approach. Also note that, the sizes of set $I$, $(|I|)$, are more than 20 in Table 1, implying that the 1Z procedure cannot be used on set $I$, and as a result the 3-Stage approach with those settings is not applicable to select the best system. Furthermore, the “Best” using the 3-Stage approach is too far from the correct selection. Thus, we need to change some assumptions in order to reduce the size of set $I$ and also in the same time to improve the efficiency of the 3-Stage selection approach.
Three-stage and four-stage approaches

Table 1: The numerical results for 3-Stage approach when $n = 3000, g = 200, m\% = 1\%, t_0 = 20$

| Run | Best | $|I|$ | $\sum_{i=1}^g T_i$ | $\sum_{i\in I} N_i$ | Time | Approach $E(OC)$ | Analytical $E(OC)$ |
|-----|------|-----|---------------------|---------------------|------|-----------------|-------------------|
| 1   | 635  | 72  | 4000                | 4684                | 13916| 0.000451923     | 0.018130246       |
| 2   | 2000 | 112 | 4000                | 7220                | 13889| 0.004078115     | 0.007142857       |
| 3   | 2072 | 145 | 4000                | 9367                | 14006| 0.001121827     | 0.006604794       |
| 4   | 2665 | 67  | 4000                | 4338                | 13616| 0.003358024     | 0.002315855       |
| 5   | 2612 | 186 | 4000                | 12132               | 15284| 0.002312955     | 0.002689141       |
| 6   | 1474 | 88  | 4000                | 5809                | 14201| 0.000151010     | 0.011194413       |
| 7   | 2547 | 148 | 4000                | 9711                | 14416| 0.001015146     | 0.003149573       |
| 8   | 428  | 93  | 4000                | 6037                | 13984| 0.000397490     | 0.019938603       |
| 9   | 2327 | 50  | 4000                | 3366                | 14432| 0.0004025164    | 0.004729810       |
| 10  | 2578 | 131 | 4000                | 8488                | 14977| 0.003102494     | 0.002929620       |

Table 2: The numerical results for 4-Stage approach when $n = 3000, g = 200, m\% = 1\%, t_0 = 20, k = 20, \Delta = 50, B = 10000$

| Run | Best | $|I|$ | $\sum_{i=1}^g T_i$ | $\sum_{i\in I} N_i$ | Time | Approach $E(OC)$ | Analytical $E(OC)$ |
|-----|------|-----|---------------------|---------------------|------|-----------------|-------------------|
| 1   | 2996 | 10  | 18743               | 1611                | 15108| 0.004189763     | 0.000027216       |
| 2   | 2971 | 18  | 39479               | 24249               | 19713| 0.002972592     | 0.000197552       |
| 3   | 2979 | 3   | 21429               | 4185                | 15698| 0.001213327     | 0.000143000       |
| 4   | 2988 | 9   | 17336               | 683                 | 14747| 0.000652923     | 0.00081679       |
| 5   | 2993 | 16  | 22922               | 5484                | 15628| 0.001791123     | 0.00047635       |
| 6   | 2999 | 17  | 22464               | 6889                | 15785| 0.000703828     | 0.00006803       |
| 7   | 2977 | 13  | 16995               | 1019                | 14623| 0.003093975     | 0.000156634      |
| 8   | 2985 | 19  | 23436               | 7679                | 16026| 0.000137651     | 0.000102114      |
| 9   | 2992 | 16  | 18837               | 2685                | 15160| 0.002445665     | 0.000054443       |
| 10  | 1101 | 17  | 23093               | 6398                | 16403| 0.000744456     | 0.014202697       |

In the second experiment, using the same assumption settings as before, we increase the initial sample size $t_0$ from 20 to 300. We expected that when we increase the value of $t_0$, it will decrease the size of set $I$ to be less than 20. Table 3 contains the results of this experiment for the 3-Stage approach. We present only the first 10 replications. From Table 3, it is clear that when we increase $t_0$, the size of set $I$ becomes less than 20. Therefore the $IZ$ procedure can be applied next on the set $I$. Unfortunately, by increasing the $t_0$, will increase the simulation samples $\sum_{i=1}^g T_i$ and $\sum_{i\in I} N_i$, and will lead to an increase in elapsed time. However, 3-Stage approach work well with $t_0 = 300$ in this
Table 3: The numerical results for 3-Stage approach when $n = 3000, g = 200, m\% = 1\%, t_0 = 300$

| Run | Best | $|I|$ | $\sum_{i=1}^{g} T_i$ | $\sum_{i \in I} N_i$ | Time | Approach $E(OC)$ | Analytical $E(OC)$ |
|-----|------|------|-------------------|-------------------|------|------------------|-------------------|
| 1   | 2973 | 19   | 38700             | 185638            | 0.001103857 | 0.000183910      |
| 2   | 1417 | 17   | 5100              | 180825            | 0.000048268 | 0.011646642      |
| 3   | 2971 | 18   | 44400             | 181381            | 0.000949636 | 0.000197552      |
| 4   | 2939 | 15   | 31500             | 191605            | 0.001044223 | 0.000416175      |
| 5   | 2999 | 19   | 38700             | 192880            | 0.000195119 | 0.000006803      |
| 6   | 2814 | 7    | 11100             | 191883            | 0.001045452 | 0.001276613      |
| 7   | 2975 | 9    | 24000             | 195421            | 0.000760501 | 0.000170271      |
| 8   | 2978 | 2    | 21600             | 192821            | 0.001273740 | 0.000149817      |
| 9   | 2991 | 9    | 29100             | 187716            | 0.001507465 | 0.000061251      |
| 10  | 2981 | 14   | 40200             | 181830            | 0.001160250 | 0.000129369      |

Table 4 contains the results of these experiments to select one of the best (1%) systems for the 3-Stage and 4-Stage approaches. We apply the same experiments as discussed above, for the $t_0 = 300$ and $t_0 = 20$ over 100 replications. From the table, $\bar{T}$ is the average of the elapsed time, $\sum_{i=1}^{g} T_i$ and $\sum_{i \in I} N_i$ for each referring to the average of $\sum_{i=1}^{g} T_i$ and $\sum_{i \in I} N_i$ respectively, and $E(OC)$ is the average of Expected Opportunity Cost.

Table 4: The performance of 3-Stage and 4-Stage approaches when $n = 3000, g = 200, m\% = 1\%$ over 100 replications

<table>
<thead>
<tr>
<th>Selection Approaches</th>
<th>$t_0$</th>
<th>$T$</th>
<th>$\sum_{i=1}^{g} T_i$</th>
<th>$\sum_{i \in I} N_i$</th>
<th>$P(CS)$</th>
<th>$E(OC)$</th>
<th>Analytical $E(OC)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-Stage</td>
<td>300</td>
<td>186668.74</td>
<td>60000</td>
<td>29305</td>
<td>80%</td>
<td>0.000931827</td>
<td>0.001615970</td>
</tr>
<tr>
<td>4-Stage</td>
<td>20</td>
<td>16160.54</td>
<td>20793</td>
<td>4372</td>
<td>83%</td>
<td>0.001818919</td>
<td>0.001800993</td>
</tr>
</tbody>
</table>

From the results we make a comparison on the performance of the 3-Stage and 4-Stage approaches. We find that the averages of the elapsed time, and simulation samples $\sum_{i=1}^{g} T_i$ and $\sum_{i \in I} N_i$, for the 3-Stage approach are very high compared to the 4-Stage approach. On the other hand, we also note that in both approaches; the 3-Stage and 4-Stage, the values of $P(CS)$ are high and very close to the analytical $P(CS)$ values. Furthermore, both the 3-Stage and
Three-stage and four-stage approaches

4-Stage approaches have small values of the $E(OC)$, with the value of $E(OC)$ for the 3-Stage approach is less compared to the 4-Stage approach.

Figure 1 shows the $E(OC)$ for each approach over 100 replications. It shows that the $E(OC)$ in the 3-Stage approach in most replications are less compared in the 4-Stage approach. The reason for this situation is that in the 3-Stage approach we used a huge number of simulation samples and it will cause the sample mean for the selected system (best) and the actual best system is very closed to each other.

![Figure 1: The $E(OC)$ for the 3-Stage and 4-Stage approaches when $n = 3000, g = 200, m\% = 1\%$ over 100 replications](image)

Now, we apply the new parameters settings, $n = 10000, g = 100, \alpha_2 = \alpha_3 = 0.005$ and $\delta^* = 0.01$, into the same $M/M/1$ queuing systems to select one of the best (3%) systems. Now our goal will be to selecting the system that belongs to \{9701, 9702, \ldots, 10000\}, and the analytical probability of the correct selection can be calculated as $P(CS) \geq 1 - \left( 1 - \frac{3}{100} \right)^{100} + 0.005 + 0.005 \geq 0.94$.

Here, the service rate $\mu \in \Theta = 7 + i/10000$, where $i = 0, 1, \ldots, 10000$, and the actual best system will be 10000 with average waiting time is 0.142857143. We report the results in Table 5 for each the 3-Stage and 4-Stage approaches with two different cases of $t_0 = 400$ and $t_0 = 50$. The experiment involved with over 100 replications.

We get the same conclusion as in Table 4, which is the averages of elapsed time and the total simulation samples are high in the 3-Stage approach compared to the 4-Stage approach. In addition, we note that both approaches; the 3-Stage and 4-Stage, select the best system with high $P(CS)$ and small $E(OC)$. Also, we find that the values of $E(OC)$ are closed to the analytical values.

Figure 2 shows the $E(OC)$ for the both approaches over 100 replications.
Table 5: The performance of 3-Stage and 4-Stage approaches when $n = 10000, g = 100, m\% = 3\%$ over 100 replications

<table>
<thead>
<tr>
<th>Selection Approaches</th>
<th>$t_0$</th>
<th>$T$</th>
<th>$\sum_{i=1}^{g} T_i$</th>
<th>$\sum_{i \in I} N_i$</th>
<th>$P(CS)$</th>
<th>$E(OC)$</th>
<th>Analytical $E(OC)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-Stage</td>
<td>400</td>
<td>842563.11</td>
<td>40000</td>
<td>69371</td>
<td>89%</td>
<td>0.000417065</td>
<td>0.000986310</td>
</tr>
<tr>
<td>4-Stage</td>
<td>50</td>
<td>114087.95</td>
<td>46999</td>
<td>23226</td>
<td>91%</td>
<td>0.000931048</td>
<td>0.000842583</td>
</tr>
</tbody>
</table>

Clearly, the values of the $E(OC)$ in the 3-Stage approach are less compare to the 4-Stage approach, because we use a large number of the total simulation samples in 3-Stage approach. In Figure 1, notice that most replications in 4-Stage approach show high values in $E(OC)$ compared to the 3-Stage approach. Meanwhile, this situation is less in Figure 2. This is due to the high total number of simulation samples that used in 3-Stage approach.

Figure 2: The $E(OC)$ for the 3-Stage and 4-Stage approaches when $n = 10000, g = 100, m\% = 3\%$ over 100 replications

5 Conclusion

In this paper, we compare between two selection approaches, to select the best stochastic system from a finite and huge set of alternatives. The first approach is the 3-Stage which consists three procedures; the $OO$, $SS$ and $IZ$, whereas the 4-Stage approach, consists four procedures; $OO$, $OCBA$, $SS$ and $IZ$. We apply both the 3-Stage and 4-Stage approaches on the example of the
Three-stage and four-stage approaches

1969

$M/M/1$ queuing systems under two different parameter settings to compare their performance. We discuss the performance comparisons on their elapsed time, number of simulation samples, and the measures of selection quality; the $P(CS)$ and $E(OC)$. From the numerical results we find that, the 3-Stage and 4-Stage approaches are not equally effective in selecting the best system for the large scale problems. Clearly, under the same parameter settings, the 3-Stage approach does not work well for the same number of initial samples $t_0$ that have been used in the 4-Stage approach. This is due to the inappropriate size of set $I$, to apply the $IZ$ procedure. To solve this problem we should increase the value of $t_0$ when we used 3-Stage approach, but this will cause increasing in the total number of the simulation samples and the elapsed time.

Regarding the $P(CS)$, we found that the $P(CS)$ in the 3-Stage approach can reach to the same $P(CS)$ as in the 4-Stage approach but only by using huge number of simulation samples. Although, from both Figure 1 and Figure 2, they show that the $E(OC)$ for the 3-Stage approach are less compared to the 4-Stage approach but it does not implying that the performance of 3-Stage approach is better than the 4-Stage approach. The reason for this is the used of high number in simulation samples in the 3-Stage approach. In general, the performance of the 4-Stage approach is better than the performance of the 3-Stage approach, since it involved with a small number of simulation samples.

ACKNOWLEDGEMENTS. The researchers would like to thank School of Mathematical Sciences, USM and the USM fellowship scheme for financial support.

References


Three-stage and four-stage approaches


Received: October, 2011