Chapter 3
Ranking and Selection: Efficient Simulation Budget Allocation

Chun-Hung Chen, Stephen E. Chick, Loo Hay Lee, and Nugroho A. Pujowidianto

Abstract This chapter reviews the problem of selecting the best of a finite set of alternatives, where best is defined with respect to the highest mean performance, and where the performance is uncertain but may be estimated with simulation. This problem has been explored from several perspectives, including statistical ranking and selection, multiple comparisons, and stochastic optimization. Approaches taken in the literature include frequentist statistics, Bayesian statistics, related heuristics, and asymptotic convergence in probability. This chapter presents algorithms that are derived from Bayesian and related conceptual frameworks to provide empirically effective performance for the ranking and selection problem. In particular, we motivate the optimal computing budget allocation (*OCBA*) algorithm and expected value of information (*EVI*) approaches, give example algorithms, and provide pointers to the literature for detailed derivations and extensions of these approaches.

3.1 Introduction

This chapter deals with the problem of selecting the best of a finite set of alternatives. In the context of simulation optimization, each alternative represents a different potential configuration of (or settings for) relevant decision variables. The performance of each alternative is not known with certainty, but its value can be estimated using stochastic simulation. We consider the setting where the number

C.-H. Chen (✉)
George Mason University, Fairfax, VA, USA
e-mail: cchen9@gmu.edu

S.E. Chick
INSEAD, Fontainebleau, France
e-mail: stephen.chick@insead.edu

L.H. Lee
National University of Singapore, Singapore, Singapore
e-mail: iseleelh@nus.edu.sg

N.A. Pujowidianto
Hewlett-Packard Singapore, Singapore, Singapore
e-mail: nugroho@hp.com

M.C. Fu (ed.), *Handbook of Simulation Optimization*, International Series in Operations Research & Management Science 216, DOI 10.1007/978-1-4939-1384-8_3

of feasible solutions (called "alternatives") is small enough so that each alternative can be simulated at least a few times in order to estimate its performance. In the conclusion, we will point to recent work addressing settings where this is not the case.

In this chapter, we will refer to a replication of a stochastic simulation as a sample. When using simulation to compare the performance of multiple alternatives, one might gradually increase the number of samples for each alternative until the variance of the estimator is sufficiently small (i.e., the confidence intervals for estimation are satisfactorily narrow) so that a satisfactory amount of evidence exists to justify a selection of the "best" alternative. One very simple approach is to use an identical number of samples for each alternative. This approach can be inefficient: if one alternative has very low variance, then it may only require very few samples to accurately estimate its performance.

To improve the efficiency of selecting the best alternative, several approaches have been explored. Intuitively, to ensure a high probability of correctly selecting an optimal alternative, a larger portion of the sampling budget should be allocated to those alternatives that are more critical in the process of identifying good alternatives quickly. Those could be alternatives with high estimated mean performance in combination with a certain degree of uncertainty about the actual mean performance. On the other hand, one might wish to sample less often the alternatives whose estimated means are either poor or have a low degree of uncertainty about their values. Two questions remain. How should one allocate resources to sample from the different alternatives, as a function of their estimated mean performance, and the uncertainty about their mean performance? For how long should one sample until stopping to select an alternative as best?

This chapter focuses on two different approaches to answering these questions, based on the optimal computing budget allocation (*OCBA*) approach and the Bayesian expected value of information (*EVI*) approach. The approaches are motivated and basic algorithms are presented. Derivations can be found in the provided references. Empirical results from these two approaches show that they perform favorably relative to some other approaches that have been proposed, in the sense of providing a relatively high average-case performance over problem instances, for a given number of samples that are observed to estimate the mean performances of the alternatives.

3.1.1 Intuitive Explanations of Simulation Budget Allocation

Consider an inventory control problem where the goal is to maximize the expected profit during a certain horizon by determining the best among five inventory ordering policies. Each alternative inventory policy is specified by two numbers s and S, where $0 \leq s < S$. If the inventory level falls below s, an order is placed to increase the inventory level to S. Otherwise, no order is placed. The profit is estimated via simulation due to randomness in demands and the amount delivered. The goal is to find the alternative with the highest mean profit.

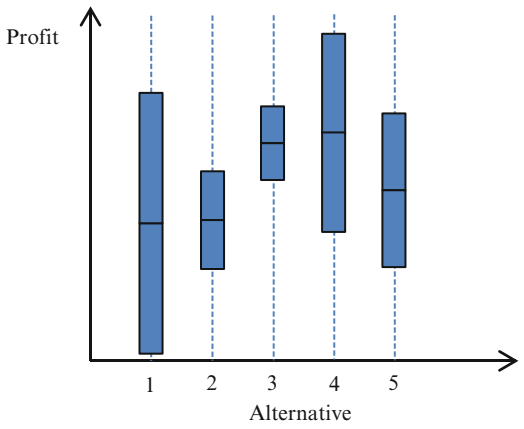

Fig. 3.1 Ninety-nine per cent confidence intervals for five alternatives in a given stage of sequential sampling

One advantage of simulation experiments is that a decision maker can collect samples in a sequential manner. After a given stage of sampling, estimates of the means of each alternative are available, along with an assessment of the uncertainty in the estimates. A selection procedure considers whether additional sampling is required before selecting an alternative, and if so, how to allocate a sampling budget to the different simulated alternatives in the next stage of sampling.

Figure 3.1 gives a representative scenario of possible results from the samples that have been collected through a given stage of sampling. It shows the 99 % confidence intervals along with the accompanying mean estimator (represented as the line in the middle of the confidence interval) for each alternative. Some alternatives seem better, but none are clearly better than all the others: all of the confidence intervals overlap. In situations such as this, it is not straightforward to determine which alternatives can be eliminated and which alternatives should receive more simulation budget.

Intuitively, the decision maker may want to allocate more samples to the alternatives with bigger half width such as alternatives 1 and 4 to reduce the variance of their estimators. On the other hand, it is sensible to allocate more samples to alternatives with larger means such as alternatives 3 and 4 as the objective is to maximize the profit. The question is how much these factors influence what gets sampled in the next stage of sampling. The next section gives an overview of the works attempting to select the best alternative.

3.1.2 Overview of Ranking and Selection (R&S)

R&S procedures aim to identify the best alternative. One R&S procedure might be considered to be better in some sense than another if it requires fewer samples, in expectation, to achieve the same level of evidence for correct selection than

that other procedure does. In this chapter, we will focus on a common context: all alternatives are simulated. This is suitable for simulation optimization problems when the number of alternatives is finite and not so large that it would prevent each alternative from being sampled at least a few times for statistical inference [26].

Many reviews on R&S are available [1, 29, 39]. There are two-stage or few-stage procedures (e.g., [21, 50]), the two-stage procedures with screening (e.g., [48]), and fully-sequential procedures, (e.g., [38]) that can guarantee a desired probability of correct selection. In indifference zone (IZ) procedures, a difference is considered to be significant if it is larger than a specified indifference-zone parameter. The probability of correct selection guarantee in the IZ approach is with respect to the probability of selecting the true best, subject to the condition that the mean of the true best is better than the mean of all of the other alternatives by at least the indifference-zone parameter. Thus, this is based on a worst-case performance metric. This worst case approach can provide frequentist guarantees for correct selection, but might require more samples to be collected to obtain that guarantee than may be practically implementable. As such, they can be statistically conservative.

In this chapter, we present R&S procedures based on average case performance metrics that sample in a highly sequential manner. The goal is either to maximize evidence for correct selection subject to a constraint on the sampling budget or to reach a level of evidence for correct selection with the fewest expected number of samples. Using an average-case analysis rather than a worst-case bound of the IZ approach, we present two distinct approaches. The *OCBA* approach uses a thought experiment that attempts to sequentially maximize the probability that the best alternative can be correctly identified after the next stage of sampling. The *EVI* approach uses a Bayesian description of the uncertainty about the mean of each alternative, a loss function to describe the penalty for not correctly selecting the best alternative, and expected value of information ideas to minimize the expected loss from selecting an alternative after simulation. This expected loss can depend upon which alternative is selected for simulation at each stage. These procedures tend to require much less sampling to achieve the same or better empirical performance for correct selection than procedures which are statistically more conservative [4, 58].

The *EVI* approach differs from the *OCBA* in one key respect. In the *OCBA* approach, the effect of additional sampling is modeled using a Bayesian asymptotic normality result: the distribution that is used to describe uncertainty after the samples are observed is assumed to be normally distributed with the same mean and with a variance that shrinks as samples accumulate. In the *EVI* approach, a decision theoretic framework is used. It explicitly models the fact that the posterior mean will change after the samples are observed. It models the distribution of what the posterior mean will be and explicitly models how changes in the posterior mean will potentially change decisions as to which alternative is best. Changes in a decision from sampling imply a value of information from those samples. The *EVI* approach allocates samples in a way that maximizes, in some sense, a measure of the expected value of information of those samples.

3.1.3 Organization

Section 3.2 formulates the selection problem and provides the basic notation. In addition, it provides a generic algorithm for selection procedures together with explanations on the components of the algorithm. The main insights of the efficient R&S procedures are presented at the end of Sect. 3.2. Section 3.3 describes one class of efficient R&S procedures, called *OCBA*. It encompasses the objective of the allocation, the allocation rule, the algorithm, and how it has been extended in other settings. Similarly, Sect. 3.4 provides the details of the *EVI* procedures. In particular, three different procedures are presented together with their algorithms. Section 3.5 concludes this chapter.

3.2 Problem Formulation and Selection Procedures

The problem of selecting the best is formulated in Sect. 3.2.1, and a generic selection procedure algorithm is provided in Sect. 3.2.2. Section 3.2.3 provides an overview of *OCBA* and *EVI* procedures, which are presented in more detail in Sects. 3.3 and 3.4, respectively.

3.2.1 Problem Formulation of Selecting the Best

We consider the problem of selecting the best of several alternatives based on their means which have to be estimated via stochastic simulation. Without loss of generality, we define the alternative with the largest mean as the best. For readability, we use upper case for random variable, lower case for fixed value or realization, and bold face for vectors.

Let X_{ij} be a random variable whose realization x_{ij} is the output of the jth sample from alternative $i, j = 1, 2, \ldots$. There are k alternatives so that $i = 1, 2, \ldots, k$. Let w_i and σ_i^2 be the unknown mean and variance of alternative i. It will be easier at times to refer to the precision $\lambda_i = 1/\sigma_i^2$ instead of the variance. Let $w_{[1]} \leq w_{[2]} \leq \ldots \leq w_{[k]}$ be the ordered means. In practice, the ordering $[\cdot]$ is unknown, and the best alternative is to be identified by sampling.

A problem configuration is denoted by $\chi = (\boldsymbol{w}, \sigma^2)$ where $\boldsymbol{w} = (w_1, w_2, \ldots, w_k)$ and $\sigma^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2)$. Let n_i be the number of samples from alternative i so far. Let $\bar{x}_i = \sum_{j=1}^{n_i} x_{ij}/n_i$ be the sample mean and $\widehat{\sigma}_i^2 = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2/(n_i - 1)$ be the sample variance. The ordered sample means are $\bar{x}_{(1)} \leq \bar{x}_2 \leq \cdots \leq \bar{x}_{(k)}$. The quantity n_i depends on the decision maker while the quantities $\bar{x}_i$, $\widehat{\sigma}_i^2$ and (i) are updated as more samples are observed.

Let $\mathfrak{D}$ be the alternative that is selected as best by the selection procedure when sampling is completed. Each selection procedure generates estimates $\widehat{w}_i$ of w_i, for

$i = 1, 2, \ldots, k$. This chapter focuses on selection procedures where the estimates are based on the sample mean, i.e., $\widehat{w}_i = \bar{x}_i$. Thus, at the time sampling stops, $\mathfrak{D} = (k)$ and a correct selection occurs when the alternative with the best sample mean is the true best (i.e., $(k) = [k]$).

In Bayesian approaches, unknown quantities are represented as random variables. Let W_i be the random variable that represents the unknown mean of alternative i. The Bayesian framework uses the notion that we can update our knowledge using the conditional distribution of parameters, given the data. Once data are available, the posterior distribution describes the uncertainty of the unknown mean. If we assume that samples are independent and normally distributed with unknown mean and variance, and that a non-informative prior distribution is used for the unknown mean and variance of each alternative, then the posterior marginal distribution for the unknown mean W_i follows a Student's t-distribution $St\left(\bar{x}_i,\, n_i/\sigma_i^2,\, v_i\right)$ where $v_i = n_i - 1$ is the degrees of freedom [20]. The mean is $\bar{x}_i$ for $v_i > 1$ and the variance is $(\sigma_i^2/n_i)v_i/(v_i - 2)$ for $v_i > 2$.

In the above framework, we have assumed that the unknown means of each alternative are independently distributed. Some extensions that are mentioned below describe how to relax some of the assumptions made above.

It will be useful to define two figures of merit that are used by the *OCBA* and *EVI* procedures below. Given the data $\mathscr{E} = \{(x_{i1},\, x_{i2} \ldots,\, x_{in_i})$ for $i = 1,\, 2, \ldots,\, k\}$, the posterior probability of correct selection (*PCS*), or posterior probability that the best alternative is correctly selected, is

$$PCS = \mathrm{P}\left(W_{\mathfrak{D}} \geq W_{[k]} \middle| \mathscr{E}\right). \tag{3.1}$$

We could further write $PCS = \mathrm{P}(\mathfrak{D} = [k])$ in contexts where ties occur with probability 0, as is the case in this setting.

If the vector of means $\boldsymbol{w} = (w_1, w_2,\, \ldots,\, w_k)$, the opportunity cost of selecting alternative i is

$$\mathscr{L}_{LL}(i,\, w) = w_{[k]} - w_i. \tag{3.2}$$

Thus, the expected opportunity cost (*EOC*), given the data $\boldsymbol{\varepsilon}$, when $\mathfrak{D} = (k)$ is selected as best is

$$EOC = \mathrm{E}\left[\mathscr{L}_{LL}(\mathfrak{D},\, W) \middle| \mathscr{E}\right] = \mathrm{E}\left[W_{[k]} - W_{\mathfrak{D}} \middle| \mathscr{E}\right]. \tag{3.3}$$

For both the *PCS* and *EOC*, the probability expectations are with respect to the posterior distribution of the unknown means, given all observed data.

3.2.2 A Generic Algorithm for Selection Procedures

We present a generic sequential selection procedure. Variations on this generic selection procedure result by making different assumptions about the nature of the evidence for correct selection and the approximations made to measure that evidence. In summary, an initialization step is used to provide initial estimates of the mean and variance of each alternative. Then, samples are collected sequentially until a stopping rule is activated. At each stage of sampling, an allocation rule specifies how the samples to be collected during that stage should be allocated among the alternatives.

The following are the steps for a generic algorithm for selection procedures.

1. Specify a first-stage sample size $n_0 > 2$, and a total number of samples $\tau > 0$ to allocate per subsequent stage. Specify stopping rule parameters.
2. Sample $X_{i1}, X_{i2}, \ldots, X_{in_0}$ independently and initialize the number of samples $n_i \leftarrow n_0$ so far for each alternative, $i = 1, 2, \ldots, k$.
3. Determine the sample statistics $\bar{x}_i$ and $\widehat{\sigma}_i^2$ and the order statistics so that $\bar{x}_{(1)} \leq \cdots \leq \bar{x}_{(k)}$.
4. WHILE stopping rule is not satisfied, DO another stage
   a. Use the allocation rule to identify which alternative to sample and determine τ_i, the number of samples to allocate to alternative i.
   b. Observe the additional samples, update the sample statistics and the order statistics.
   c. Update the number of samples collected so far for each alternative, $n_i \leftarrow n_i + \tau_i$.
5. Select the best alternative based on the selection rule.

As described in the algorithm, the decision maker needs to decide the allocation rule, the stopping rule, and the selection rule. The following three paragraphs describe each of the components of the algorithm in general. Sections 3.3 and 3.4 give specific examples of allocation rules and stopping rules.

An *allocation rule* is a mapping from the sampling statistics of the k alternatives to a vector of integers that represents the number of samples to observe from each alternative in the next stage of sampling. Let τ_i be the number of samples allocated to alternative i in the next stage of sampling, for $i = 1, 2, \ldots, k$. We require that there be a total of τ samples, so that $\sum_{i=1}^{k} \tau_i = \tau$. The τ_i's are recalculated at each stage of sampling. For example, in an equal allocation, the allocation rule is to sample evenly from each alternative: it assigns samples so that, at the end of the stage of sampling, the difference between the number of samples between the most-sampled and the least-sampled alternatives, is minimized. If $\tau = k$, then the equal allocation rule sets $\tau_i = 1$ for each i at each stage. If $\tau = 1$, then one equal allocation rule could set $\tau_i = 1$ for the alternative with the smallest index among those alternatives that have been sampled the least so far, and $\tau_i = 0$ for the other alternatives.

The *stopping rule* specifies the condition under which sampling is terminated so that an alternative can be selected as best. It can be based on a total sampling budget or the desired level of evidence for a correct selection. If a total sampling budget is chosen, this reflects a choice to take a deterministic number of samples before selecting an alternative as best independent of the samples seen: sampling continues if and only if $\sum_{i=1}^{k} n_i < \beta$, where β is a user-specified total sampling budget. A selection procedure with such a deterministic sampling budget is "better" if it provides a higher expected level of evidence for correct selection after sampling. Such evidence might be the posterior *PCS*.

If a desired level of evidence for correct selection is chosen as a criterion for stopping, such as a choice to stop sampling when the posterior *PCS* is above a pre-specified threshold, then a "better" procedure is one that requires a fewer number of samples, in expectation, to reach that threshold. Such a stopping rule is called an adaptive stopping rule, because the total number of samples may depend on the values of the samples.

The *selection rule* specifies which alternative to select as best when sampling is completed. A very common sampling rule is to pick the alternative with the largest sample mean, i.e., $\mathfrak{D} = (k)$, where (k) is the alternative with the largest sample mean when the sampling stops, and $\mathfrak{D}$ is the (random) decision variable that represents the selected alternative. This decision rule is known to be optimal in some situations (e.g., if the loss function is the expected opportunity cost [16, 30]). An alternative selection rule is to select the alternative with the largest posterior probability of being the best [2].

Throughout this chapter, the "default" approach is to use a non-informative prior distribution for the unknown parameters [20]. This implies that the decision maker does not favor any specific value for the unknown means. As a result, initial number of samples need to be collected. If there is additional information, the decision maker can use an informative prior distribution to describe that information. Branke et al. [4] and Chick and Frazier [14] show how to handle the analysis with informative prior distributions.

3.2.3 General Concepts for OCBA and EVI

Before going into the details in Sects. 3.3 and 3.4, the basic ideas of four allocation rules are presented. The first allocation rule is *OCBA* for unconstrained optimization, which will be further described in Sect. 3.3. The other three allocation rules attempt to maximize *EVI* and will be discussed in detail in Sect. 3.4.

Basic Idea of *OCBA*

The basic goal of *OCBA* is to maximize the probability of correct selection for a given stage of sampling. It aims to derive closed-form expressions that are easy to implement to allocate multiple samples to multiple alternatives.

This is done using an asymptotic framework to analyze the structure of the optimal allocation when the number of samples tends to infinity. For example, it uses the Bayesian asymptotic normality result where the posterior distribution of W_i follows a normal distribution, $W_i \sim \mathcal{N}\left(\bar{x}_i, \sigma_i^2/n_i\right)$. Based on asymptotic analysis, the simulation budget allocation problem can be formulated as a non-linear deterministic optimization problem that can be solved using classical techniques such as the Karush–Kuhn–Tucker (KKT) conditions. The resulting closed-form allocation rule can then be implemented using a sequential heuristic algorithm.

Basic Idea of *EVI*

The value of information is defined as the expectation of the reward obtained with additional information less the reward obtained without that information (that is, the expected opportunity cost). Ideas for the *EVI* approach have been developed independently for several different distributional assumptions and approaches for how to value the information [16, 22, 31, and others]. The three approaches to *EVI* discussed here include variations on whether the information to be valued is obtained by sampling from one or from multiple alternatives in a given stage of sampling, or whether the information from only a single stage (resulting in a so-called one-step lookahead policy) or from potentially multiple stages of sampling is modeled. The former is typically easier to do than the latter.

The Linear Loss procedure (*LL*) is an *EVI* procedure that can allocate multiple samples to multiple alternatives in each stage of sampling. It aims to minimize the expected opportunity cost (*EOC*), which is the difference in means between the selected alternative and the best alternative. Linear loss is another name for the *EOC* [30]. In the *LL* procedure, the additional information is one extra stage of sampling, and the reward is the posterior mean reward from the alternative that would be chosen as best.

The LL_1 allocation is like the *LL* in that it allocates samples to alternatives at each stage of sampling in order to minimize *EOC* of a potentially incorrect selection, and does so in a myopic one-step lookahead manner. The LL_1 differs from the *LL* allocation in the sense that it requires all samples within a given stage to be taken from a single alternative and so it is called LL_1. This restriction enables a closed form solution for the one-stage value of information when samples are normally distributed (with known or unknown means and variances). The LL_1 allocation was independently developed by Frazier et al. [22]. The resulting approach is called the knowledge gradient of which idea is extensively presented and discussed by Powell and Ryzhov [49].

Economics of Selection Procedure (*ESP*) looks at the case where the decision maker has the option to continue sampling or to stop and select the best. It specifies one alternative to be sampled like LL_1. However, it looks at the future streams of rewards (potentially multiple stages of learning and sampling costs, the benefit from implementing a selected alternative) instead of the one-step lookahead analysis as in the case of *OCBA*, *LL*, and LL_1 which only maximizes the evidence for correct

selection by the end of one stage of sampling. As a result, it accounts for the value of the ability of further sampling after that stage and therefore takes a non-myopic view of the value of sampling.

3.3 Optimal Computing Budget Allocation (*OCBA*)

This section briefly introduces the *OCBA* approach. As n_i increases, $\widehat{w}_i = \bar{x}_i$ becomes a better approximation to w_i in the sense that its corresponding variance becomes smaller. At a given stage, we need to allocate the additional τ samples to each alternative. As motivated in Sects. 3.1 and 3.2, instead of equally simulating all alternatives, we want to choose $\tau_1, \tau_2, \ldots, \tau_k$ more intelligently so that the simulation efficiency can be enhanced.

3.3.1 Maximization of PCS

The simulation budget allocation problem that *OCBA* in Chen et al. [12] aims to maximize the probability of correct selection (*PCS*) subject to the sampling budget of a given stage of sampling τ,

$$\max_{\tau_1, \tau_2, \ldots, \tau_k} PCS \; s.t. \sum_{i=1}^{k} \tau_i = \tau, \tau_i \geq 0. \tag{3.4}$$

Here $\sum_{i=1}^{k} \tau_i = \tau$ denotes the total computational cost assuming the simulation execution times for different alternatives are roughly the same. Formally, τ_i is a non-negative integer. However, the allocation rule is derived assuming that τ_i is a continuous variable.

3.3.2 Asymptotic Allocation Rule

We use the Bayesian asymptotic normality result, $W_i \sim \mathcal{N}\left(\bar{x}_i, \sigma_i^2/n_i\right)$ [20]. After the simulation is performed, $\bar{x}_i$ can be calculated, σ_i^2 can be approximated by the sample variance; *PCS* can then be estimated using a Monte Carlo simulation. However, estimating *PCS* via Monte Carlo simulation is time-consuming. Since the purpose of budget allocation is to improve simulation efficiency, we need a relatively fast and inexpensive way of estimating *PCS* within the budget allocation procedure. Efficiency is more crucial than estimation accuracy in this setting. We adopt a common approximation procedure used in simulation and statistics literature. This approximation is based on the Bonferroni inequality. For brevity, we drop the

notation ε.

$$\begin{aligned} PCS &= \mathrm{P}\left(\bigcap_{i:(i)\neq(k)} \left(W_{(k)} - W_{(i)} \geq 0\right)\right) \\ &\geq 1 - \sum_{i:(i)\neq(k)} \left[1 - \mathrm{P}\left(W_{(k)} - W_{(i)} \geq 0\right)\right] \\ &= 1 - \sum_{i:(i)\neq(k)} \mathrm{P}\left(W_{(k)} < W_{(i)}\right) \\ &= 1 - \sum_{i:(i)\neq(k)} \Phi_N\left(-d_{(j)(k)}\right) = APCS. \end{aligned} \tag{3.5}$$

where Φ_N is the cumulative distribution function of standard normal distribution and $d_{(j)(k)} = \bar{x}_{(k)} - \bar{x}_{(j)}$.

Consider an asymptotic case ($\tau \to \infty$ so that the total sampling budget $\beta \to \infty$ and $\tau_i \to n_i$), Chen et al. [12] show that the approximation of *PCS* given in (3.4) can be maximized when

$$\frac{n_{(i)}}{n_{(j)}} = \left(\frac{\sigma_{(i)}/d_{(i)(k)}}{\sigma_{(j)}/d_{(j)(k)}}\right)^2, (i), (j) \in \{1, 2, \ldots, k\} \text{ and } (i) \neq (j) \neq (k), \tag{3.6}$$

$$n_{(k)} = \sigma_{(k)} \sqrt{\sum_{i:(i)\neq(k)} \frac{n_{(i)}^2}{\sigma_{(i)}^2}}. \tag{3.7}$$

It is interesting to see that (3.6) implies that the number of replications for alternative i is proportional to the square of a noise-to-signal ratio, where the noise refers to the sample standard deviation and the signal refers to the difference between alternative i's sample mean and the best sample mean.

3.3.3 Sequential Heuristic Algorithm for Allocation

With the asymptotic solution in (3.6) and (3.7), we now present a cost-effective sequential approach based on *OCBA* to select the best alternative from k alternatives with a user-specified total sampling budget β. Each alternative is initially simulated with n_0 samples in the first stage, and additional samples are allocated incrementally with τ samples to be allocated in each iteration. Let t be the stage number.

OCBA Algorithm

INPUT k, β, τ, n_0 ($\beta - kn_0$ is a multiple of τ and $n_0 \geq 5$);

INITIALIZE $t \leftarrow 0$;

Perform n_0 samples for all alternatives; $n_{1,t} = n_{2,t} = \ldots = n_{k,t} = n_0$.

LOOP WHILE $\sum_{i=1}^{k} n_{i,t} < \beta$ **DO**

UPDATE Calculate sample means $\bar{x}_i \sum_{j=1}^{n_{i,t}} x_{ij}/n_{i,t}$, and sample standard deviation $\widehat{\sigma}_i = \sqrt{\sum_{j=1}^{n_{i,t}} (x_{ij} - \bar{x}_i)^2/(n_{i,t}-1)}$, $i = 1, 2, \ldots, k$, using the new simulation output; find $(k) = \arg\max_i \bar{x}_i$.

ALLOCATE Increase the sampling budget by τ and calculate the new budget allocation, $n_{1,t+1}, n_{2,t+1}, \ldots, n_{k,t+1}$, according to

i) $\frac{n_{(i),t+1}}{n_{(j),t+1}} = \left(\frac{\widehat{\sigma}_{(i)}/d_{(i)(k)}}{\widehat{\sigma}_{(j)}/d_{(j)(k)}}\right)^2$, for all $(i) \neq (j) \neq (k)$, and

ii) $n_{(k),t+1} = \widehat{\sigma}_{(k)} \sqrt{\sum_{i:(i)\neq(k)} \frac{n_{(i),t+1}^2}{\widehat{\sigma}_{(i)}^2}}$,

SIMULATE Perform $\tau_i = (n_{(i),t+1} - n_{(i),t})^+$ additional simulations for alternative $i = 1, 2, \ldots, k$, where $(x)^+ = \max(0, x)$;

$t \leftarrow t + 1$.

END OF LOOP

The resulting $n_{(i)}$ in the ALLOCATE step is a continuous number that must be rounded to an integer. In the numerical experiments in the next section, $n_{(i)}$ is rounded to the nearest integer such that the summation of additional simulation replications for all solutions equals τ. Note that there may not always exist a solution that satisfies all the three constraints. It actually occurs when at least one solution has been over simulated, i.e., $n_{(i),t+1} < n_{(i),t}$. In this case, we have to relax the constraint. For ease of control of the simulation experiment, we can choose to maintain the constraint $\sum_{i=1}^{k} n_{(i),t+1} = \tau + \sum_{i=1}^{k} n_{(i),t+1}$ and apply some heuristics to round $n_{(i),t+1}$ for all i to nearest integers. Chen and Lee [8] have found numerically that the performance is not sensitive to how we round $n_{(i)}$, probably due to the robustness of a sequential procedure.

Alternative Simpler *OCBA* Procedure

When the computational cost of samples (simulation replications) is relatively large as compared to the computational cost of the ALLOCATE step, we recommend that τ should be small, or even set to 1. When $\tau = 1$ the ALLOCATE and SIMULATE steps can be simplified as follows.

ALLOCATE Increase the sampling budget by $\tau = 1$ and calculate a tentative allocation $n_{1,t+1}, n_{2,t+1}, \ldots, n_{k,t+1}$, according to

i) $\frac{n_{(i),t+1}}{n_{(j),t+1}} = \left(\frac{\widehat{\sigma}_{(i)}/d_{(i)(k)}}{\widehat{\sigma}_{(j)}/d_{(j)(k)}}\right)^2$, for all $(i) \neq (j) \neq (k)$, and

ii) $n_{(k),t+1} = \widehat{\sigma}_{(k)} \sqrt{\sum_{i:(i)\neq(k)} \frac{n_{(i),t+1}^2}{\widehat{\sigma}_i^2}}$,

leave $n_{(i),t+1}$ as a decimal number and find $(i^*) = \arg\max_i \left(n_{(i),t+1} - n_{(i),t}\right)$.

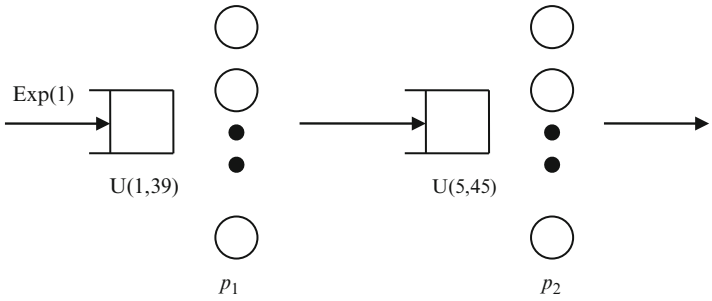

Fig. 3.2 A two-stage queuing system where both p_1 and p_2 must be greater than 10

SIMULATE Perform additional one sample for alternative (i^*);
$n_{(i^*),t+1} = n_{(i^*),t} + 1$; $n_{(i),t+1} = n_{(i),t}$ for $(i) \neq (i^*)$;
$t \leftarrow t + 1$.

Intuitively, we determine which alternative is the most starving one in terms of the need of additional simulation, and then simulate that alternative for one additional replication. This procedure is iteratively continued until the total budget β is exhausted or the estimated *APCS* is sufficiently high. As shown in Chen and Lee [8], this simpler procedure performs equally well in our numerical testing.

3.3.4 Numerical Results

Chen and Lee [8] provide extensive numerical results for *OCBA*. Among them, Fig. 3.2 gives an example of a two-stage queuing system, where we want to allocate 31 parallel servers within a two-stage queue where each stage of the queue can contain no less than 11 servers.

Denote p_1 and p_2 as the numbers of workers allocated to nodes 1 and 2 respectively. Thus, $p_1 + p_2 = 31$, $p_1 \geq 11$, and $p_2 \geq 11$. There are ten alternative combinations of (p_1, p_2). We want to choose the best alternative of (p_1, p_2) so that the average system time for the first 100 customers is minimized. Since there is no closed-form analytical solution for the estimation of the system time, stochastic (discrete-event) simulation can be performed to find the best alternative.

To characterize the performance of different procedures as a function of β, we vary β between 200 and 8,000 for all of the sequential procedures and the estimated achieved *PCS* as a function of β is shown in Fig. 3.3. We estimate the *PCS* by estimating the fraction of the event of correct selection out of the independent experiments that are conducted.

We see that all procedures obtain a higher *PCS* as the available sampling budget increases. We can then record the number of samples that correspond to where the curve crosses at a certain level of *PCS* that we are interested in. It can be seen that

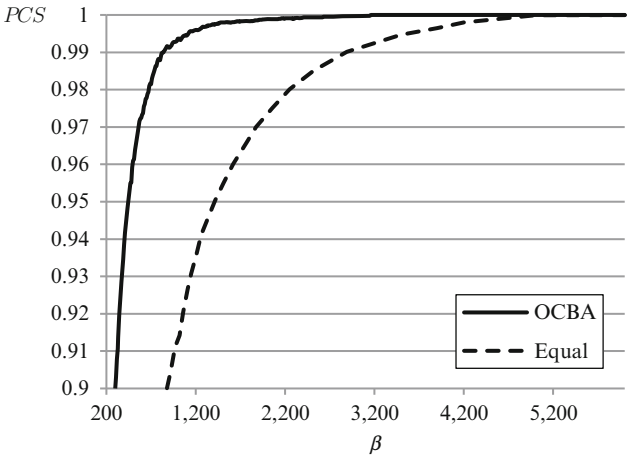

Fig. 3.3 *PCS* vs. β using three sequential allocation procedures

Table 3.1 The sampling budget to attain $PCS = 0.95$ or 0.99

PCS	*OCBA*	Equal allocation
0.95	470	1,450
0.99	850	2,890

Table 3.2 Number of maximum allowable workers to simulate varying numbers of alternatives

Maximum number of workers	Number of alternatives
31	10
41	20
51	30
61	40
71	50
81	60
91	70
101	80
111	90
121	100

OCBA achieves a same *PCS* using the lowest amount of sampling budget. Table 3.1 shows the sampling budget to attain $PCS = 0.95$ and 0.99 for *OCBA* and Equal Allocation.

It is not surprising that the actual sampling cost using *OCBA* depends on the specific problem and the corresponding *PCS* requirement. However, the speedup factor of using *OCBA* versus equal allocation is not very sensitive to problem specifics, except for the number of alternatives.

Instead of providing only 31 servers, we increase the number of servers up to 121, where there must be at least 11 servers at each station. As a result, the number of possible alternatives varies from 10 to 100. Table 3.2 lists some of the possibilities.

Table 3.3 Speedup factor of using *OCBA* compared with the use of equal allocation

Number of alternatives (k)	4	10	20	50	75	100
Speedup factor	1.75	3.42	6.45	12.8	16.3	19.8

In this test, we compare *OCBA* and equal allocation, and focus on the "speedup factors" under *OCBA*. For both procedures, we record the minimum sampling budget where the curve plotting the estimated *PCS* vs. β crosses $PCS = 0.99$: β_{OCBA} and β_{EA}. The "speedup factor" using *OCBA* is given by the ratio β_{EA}/β_{OCBA}. Table 3.3 shows the numerical results for different number of alternatives. We see that *OCBA* is even more efficient as the number of alternatives increases. The higher efficiency is obtained, because a larger alternative space gives the *OCBA* algorithm more flexibility in allocating the sampling budget.

3.3.5 *Minimization of EOC*

Instead of maximizing *PCS*, we turn our attention to the expected opportunity cost (*EOC*). From the simulation efficiency perspective, one has the same question to ask: how should we allocate the simulation samples so that we can select an alternative within the given sampling budget while *EOC* is minimized, instead of maximizing *PCS* as in previous sections?

Deriving an asymptotic solution for minimizing *EOC* is much more complicated than its counterpart for *PCS*. Following the same notion of the greedy approach given in Chen et al. [11] and Hsieh et al. [36], He et al. [33] present a greedy selection procedure, called $OCBA_{LL}$ (or $OCBA - EOC$), to reduce the *EOC* of a potentially incorrect selection by taking a similar *OCBA* approach to selection.

A critical component in the proposed procedure is to estimate how *EOC* changes as n_i changes. Let τ_i be a nonnegative integer denoting the number of additional simulation samples allocated to alternative i in the next stage of sampling. We are interested in assessing how *EOC* would be affected if alternative i was simulated for τ additional replications. He et al. [33] present an *Estimated Expected Opportunity Cost* (*EEOC*), which is an upper bound of *EOC*, as follows

$$EEOC = \sum_{i:(i)\neq(k)} \mathrm{P}\left(W_{(i)} \geq W_{(k)}\right) \mathrm{E}\left[W_{(i)} - W_{(k)} \,\middle|\, W_{(i)} \geq W_{(k)}\right]. \tag{3.8}$$

The *OCBA* approach in this case aims to sequentially minimize *EEOC*. A critical component is to estimate how *EEOC* changes if alternative i is allocated with τ_i additional replications in a given stage. A heuristic approach to the approximation of the predictive posterior distribution yields $W_i \sim \mathcal{N}\left(\frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}, \frac{\widehat{\sigma}_i^2}{n_i + \tau_i}\right)$.

The *EEOC* can then be determined by using the original distributions for the unknown means of alternatives other than i as follows

$$EEOC(i) = \sum_{j:(j)\neq(k)} \int_0^{+\infty} x f^*_{(k),\,(j),(i)}(x)dx, \tag{3.9}$$

where $f^*_{(k),\,(j),(i)}(x)$ is the probability density function (p.d.f.) of the difference between alternative (k) and (j), given that $\tau_{(i)}$ additional replications are allocated to alternative (i), and none is allocated to others. If $(i)=(k)$, then $f^*_{(k),\,(j),(i)}(x)$ is the p.d.f. of $\mathcal{N}\left(\bar{x}_{(j)}-\bar{x}_{(k)},\ \frac{\widehat{\sigma}^2_{(k)}}{n_{(k)}+\tau_{(k)}}+\frac{\widehat{\sigma}^2_{(i)}}{n_{(i)}}\right)$. If $(i)=(j)$, then $f^*_{(k),\,(j),(i)}(x)$ is the p.d.f. of $\mathcal{N}\left(\bar{x}_{(i)}-\bar{x}_{(k)},\ \frac{\widehat{\sigma}^2_{(k)}}{n_{(k)}}+\frac{\widehat{\sigma}^2_{(i)}}{n_{(i)}+\tau_{(i)}}\right)$. Otherwise, no new information is available to distinguish alternatives (k) and (j) and $f^*_{(k),\,(j),(i)}(x)$ is the p.d.f. of $\mathcal{N}\left(\bar{x}_{(j)}-\bar{x}_{(k)},\ \frac{\widehat{\sigma}^2_{(k)}}{n_{(k)}}+\frac{\widehat{\sigma}^2_{(i)}}{n_{(i)}}\right)$.

Since we want to minimize *EOC*, $OCBA_{LL}$ sequentially allocates additional samples to the alternatives that lead to the lowest *EEOC* at each stage. Let r be the number of alternatives to simulate in each stage. The ALLOCATE steps are revised as follows

ALLOCATE Find the set $S(r) \equiv \{j: EEOC(j) \text{ is among the } r \text{ lowest values}\}$. Increase the sampling budget by $\tau_{(i)} = \tau/|S(r)|$ for alternative $(i) \in S(r)$, i.e., $n_{(i),t+1} = n_{(i),t} + \tau_{(i)}$ if $(i) \in S(r)$, $n_{(i),t+1} = n_{(i),t}$ otherwise. If $\tau = 1$, only a single replication is allocated to alternative j which minimizes $EEOC(j)$.

3.3.6 Other Variants

This section discusses some extensions of the *OCBA* approach. We begin by considering an efficient budget allocation procedure for selecting an optimal subset of top-r alternatives rather than the single best alternative [10]. Then we consider problems with multiple performance measures, which can be formulated either as constrained optimization or multi-objective optimization. Other recent developments are then discussed and finally concluded with examples of generalizations of *OCBA* notions.

Subset Selection Problem

Instead of selecting the best alternative as in previous section, we consider a class of subset selection problems in simulation optimization or ranking and selection. In some cases, it is more useful to provide a set of good alternatives than a single best alternative for decision maker to choose, because he/she may have other concerns which are not modeled in the simulation. Such efficient subset selection procedures

are also beneficial to some recent developments in simulation optimization that require the selection of an "elite" subset of good candidate solutions in each iteration of the algorithm, such as evolutionary population-based algorithm. A subset with good performing solutions will result in an update that leads the search in a promising direction.

Specifically, our objective is to find *all* top-r alternatives, where $r > 1$ is the number of alternatives to select. Koenig and Law [40] develop a two-stage procedure for selecting all the r best alternatives along the lines of the procedure in Dudewicz and Dalal [21] (see also Sect. 10.4 of Law [41] for an extensive presentation of the problem and procedure). However, the number of additional samples for the second stage is computed based on a least favorable configuration, resulting in very conservative allocations, so that the required computational cost is much higher than actually needed.

This problem can be easily handled by changing the correct selection definition in *OCBA*. Specifically, $PCS = \mathrm{P}(W_{(i)} \geq W_{(j)}, \forall\, (i) \in S_r, (j) \notin S_r)$ where S_r is the optimal subset. Chen et al. [10] show that the allocation rule is similar to (3.6), that is $\frac{n_{(i)}}{n_{(j)}} = \left(\frac{\sigma_{(i)}/(\bar{x}_{(i)} - q)}{\sigma_{(j)}/(\bar{x}_{(j)} - q)} \right)^2$, where q is a value between $\bar{x}_{(r)}$ and $\bar{x}_{(r+1)}$. A suggested value of q, which asymptotically maximizes both $\mathrm{P}(W_{(r)} \leq q)$ and $\mathrm{P}(W_{(r+1)} \geq q)$, is given by $q = \frac{\widehat{\sigma}_{(r+1)}\bar{x}_{(r)} + \widehat{\sigma}_{(r)}\bar{x}_{(r+1)}}{\widehat{\sigma}_{(r)} + \widehat{\sigma}_{(r+1)}}$.

Handling Optimization with Multiple Performance Measures

OCBA has also been extended to tackle other simulation-based optimization problems. The needs of optimization problems with multiple performance measures become more evident. We can categorize these problems according to whether there are any constraints and whether the constraints are stochastic or deterministic.

In the case of multi-objective optimization where no performance measures are constrained, Lee et al. [43, 44] consider the problem of finding the non-dominated Pareto set where the evidences for correct selection used are type I and type II errors.

There are cases where the secondary stochastic performance measures act as constraints. In this case, the simulation budget allocation can be allocated based on the optimality only, feasibility only, or both. Lee et al. [45] propose an *OCBA* approach that maximizes a lower bound of *PCS*. The procedure is applicable for both the independent case and the case with correlated performance measures. In some situations, the decision makers are only interested in differentiating the feasible alternatives from the infeasible ones, which is called feasibility determination problem as addressed by Szechtman and Yücesan [55]. In this case, there is no need to select the best alternative.

It is also possible to consider the descriptive complexity preference. In this case, an alternative that is simpler, i.e., having smaller descriptive complexity is

preferred compared to a complex one if they have similar performance. The decision makers therefore want to select the top-r simplest alternatives of which performance measures are above certain desired level [37, 60].

Other Recent Developments

There are several other related works along the lines of the *OCBA* research. This section gives some examples, which is by no means an exhaustive list. The *OCBA* rule presented in Sect. 3.3.2 assume that the variances are known. In the case of unknown variances, Chen et al. [13] proposed the *OCBA* algorithm based on t distribution. It is found that the differences between the results obtained based on the t distribution model and the normal distribution model is not significant. Instead of finding the alternative with the best mean, Trailovic and Pao [57] develop an *OCBA* approach for finding an alternative with minimum variance. Unlike the independence assumption of simulation samples required in this book, Fu et al. [27] extend the *OCBA* to problems in which the simulation outputs between alternatives are correlated. Chen et al. [9] study of the benefit of dynamic allocation. Glynn and Juneja [28] extend the *OCBA* to problems in which the simulation output is no longer normally distributed by utilizing large deviation theory. Blanchet et al. [3] further extend the work to heavy-tailed distributions, also utilizing large deviation theory.

Brantley et al. [5, 6] enhance *OCBA* efficiency by incorporating information from across the domain into a regression equation. Morrice et al. [46, 47] further extend the concepts to a method for selecting the best alternative based on a transient mean performance measure.

Generalized *OCBA* Notions

The different extensions presented earlier indicate that there exists a consistent notion where an optimization model is used to determine the best allocation scheme to maximize a certain desired quality of the outcome given a fixed budget. The *OCBA* notion has been generalized for different purposes well beyond selecting the best alternative as presented in previous subsections. The main idea is that it is possible to replace the objective function *PCS* with other objectives. In addition, the budget to be allocated is not necessarily in terms of computer time or simulation replications. We give three examples in this section.

The Cross-Entropy (CE) method introduced by Rubinstein [51] belongs to a class of global optimization algorithms called estimation of distribution algorithms, which work with a probability distribution over the solution space. In every iteration of CE, we will first generate a population of solutions from a probability density function (p.d.f.) with a certain parameter. After these generated solutions are simulated, the parameters of the distribution are updated by minimizing the Kullback–Leibler (KL) divergence (or the cross entropy) between the parameterized

p.d.f. and the target optimal p.d.f.. The CE method has shown to be promising in solving difficult global optimization problems, but its main focus has been on deterministic optimization problems. For the stochastic setting, He et al. [34] develop the *OCBA–CE* procedure that integrates the objectives of minimizing the KL divergence from a parameterized distribution that generates the candidate solutions in the CE method with that of minimizing the total computing budget per iteration. Numerical testing indicates that the *OCBA–CE* is promising, resulting in substantial computational efficiency gains over the CE method with equal allocation.

The second example is the work by Shortle et al. [54], which uses the notion of *OCBA* to the problem of estimating a rare-event probability using splitting simulation. Multi-level splitting is an effective variance reduction technique. The basic idea is to create separate copies (splits) of the simulation whenever it gets close to the rare event. Each level is smaller and much easier to simulate than the original. Note that this problem is not an optimization problem as the decision maker does not need to select the best alternative. The problem of determining the number of splits is formulated as an optimal computing budget allocation problem. In this context, the objective is to minimize the variance of the rare-event probability estimator. The budget is the total computation time, which needs to be allocated to different levels.

OCBA can also be extended to problems without simulation or optimization. For example, Wong et al. [59] propose an *OCBA* approach for Data Envelopment Analysis (DEA), which is a mathematical programming approach by Charnes et al. [7] for measuring efficiency for decision-making units with multiple inputs and multiple outputs. The idea is to compare different decision-making units in terms of how many inputs they have used in achieving the outputs. In this case, the objective is to minimize the expected mean square error for the prediction of the efficiency score in DEA. The budget refers to the total budget for data allocation. Therefore, the budget allocation problem is to determine how the data should be collected, i.e., finding the optimal number of data points allocated for different unknown variables to maximize the predicted efficiency score.

3.4 Expected Value of Information (*EVI*)

The *EVI* approach is based on a Bayesian decision theoretic approach rather than a frequentist statistical approach. It is Bayesian in the sense that it presumes that uncertainty about all unknown parameters, such as the unknown means of each alternative, be described with probability distributions. It is decision theoretic in the sense that sampling allocation decisions, as well as decisions to select a given alternative as best, are based on maximizing an expected reward (or equivalently, minimizing an expected loss).

There are several potential loss functions of interest in this framework. One loss function is the $0-1$ loss function: a loss of 1 is incurred if the true best alternative is not selected as best, and a loss of 0 is incurred if the best alternative is best. With

the $0-1$ loss function, the expected loss is the *PCS* in (3.1). Another loss function of interest is the opportunity cost. The opportunity cost is 0 if the best alternative is selected as best. Otherwise, the opportunity cost is the difference between the performance of the best alternative and the performance of the alternative selected as best. The expected opportunity cost (*EOC*), averaged over sampling decisions and realizations of the unknown true performance, has particular significance when the outputs of simulated alternatives are linear measures of financial (profit) outcomes for the alternatives they simulate. Minimizing the *EOC* in that case is equivalent to maximizing the expected profit. The *EOC* is therefore particularly interesting when using a selection procedure to select an alternative with the greatest financial benefit.

Independent of whether the loss function is the $0-1$ loss function, the opportunity cost, or some other loss function, the *EVI* approach allocates samples in each stage of sampling with the goal of reducing the expected loss obtained after the samples are observed. This problem can be solved analytically in a few special cases (e.g., normal distributed output with known sampling variances, $k=2$ alternatives); otherwise analytically motivated approximations that have attractive theoretical properties can be assembled to provide good heuristics. Some of these are described below.

We highlight three *EVI* procedures that focus on information with respect to the *EOC* loss function. Section 3.4.1 presents the *LL* procedure, which allocates samples to multiple alternatives in each stage. Section 3.4.2 focuses on a simplification where only one additional sample to a single alternative can be collected in each stage. Section 3.4.5 presents the multi-step valuation of information where the proposed procedure that looks at the value of multiple stages of sampling in order to further improve effectiveness of an allocation in a given stage.

3.4.1 Linear Loss (LL)

Chick and Inoue [16] proposed a procedure to determine the number of samples to minimize the expected opportunity cost (*EOC*).

To find the expected value of sampling an alternative in the next stage of sampling, we introduce some additional notations to account for the random output that will be observed. Let Y_{ij} be a random variable where y_{ij} is the output of the jth sample from alternative i, $j=1, 2, \ldots$, that will be observed in the next stage of sampling. Let $\bar{y}_i$ be the sample average of the output in the additional stage based on the additional τ_i samples, $\bar{y}_i = \sum_{j=1}^{\tau_i} y_{ij}/\tau_i$. As defined in Sect. 3.2.1, $\bar{x}_i$ is the sample mean based on n_i, the number of samples from alternative i so far. The overall sample mean for alternative i, denoted as z_i is therefore

$$z_i = \frac{n_i \bar{x}_i + \tau_i \bar{y}_i}{n_i + \tau_i}. \tag{3.10}$$

Before the sampling is done, $Z_i = \frac{n_i \bar{x}_i + \tau_i \overline{Y}_i}{n_i + \tau_i}$ is random. Thus, Z_i is the posterior mean of the unknown mean for alternative i given that ε is the information seen so far and the fact that τ_i samples from alternative i will be observed but have not yet been observed. The distribution of Z_i depends on the value of τ_i.

Minimization of *EOC*

Conceptually, at each stage of sampling, the *LL* procedure seeks to allocate samples to several alternatives so that the expected *EOC* that will be obtained will be minimized after the next stage of sampling. That so-called predictive *EOC* depends on the samples taken in the next stage as they will determine the $\{Z_i\}$, which in turn determine which alternative will be selected.

Assessing that predictive *EOC* is analytically intractable when $k > 2$ or when variances are unknown, even when samples are normally distributed. The *LL* procedure attempts to minimize an upper bound on that predictive *EOC*. In particular, we minimize (for asymptotically large τ) the term $EOC_{bnd}(\tau_1, \tau_2, \ldots, \tau_k)$, where

$$EOC_{bnd} = \sum_{i:(i)\neq(k)} \mathrm{E}\left[\left(W_{(i)} - W_{(k)}\right)^+\right] - \mathrm{E}\left[\left(Z_{(i)} - Z_{(k)}\right)^+ \middle| \mathscr{E}\right]. \tag{3.11}$$

depends on the $\tau_{(i)}$ via the $Z_{(i)}$, is an upper bound on the predictive *EOC*, where (k) is the alternative with the highest mean given the data ε observed so far, and $(x)^+ = \max(0, x)$.

Conceptually, the difference in the summand in the equation for EOC_{bnd} is the expected opportunity cost of selecting with no additional information, $\mathrm{E}[(W_{(i)} - W_{(k)})^+|\varepsilon]$, minus the expected value of information of sampling, $\mathrm{E}[(Z_{(i)} - Z_{(k)})^+|\varepsilon]$, in a pairwise comparison between alternatives (i) and (k). When no additional samples are taken, the term $\mathrm{E}[(Z_{(i)} - Z_{(k)})^+|\varepsilon]$ is 0. When an infinite number of samples are taken for both (i) and (k), the difference in the summand is 0 (an infinite number of samples gives perfect information).

At each stage of sampling, ε is updated, and the goal of the *LL* procedure is to find an allocation that minimizes EOC_{bnd}.

$$\min_{\tau_1, \tau_2, \ldots, \tau_k} EOC_{bnd} \; s.t. \sum_{i=1}^{k} \tau_i = \tau, \; \tau_i \geq 0. \tag{3.12}$$

A solution to the problem in (3.12) for the case of known variances is not known other than by searching on a lattice, as the τ_i are non-negative integers. We therefore derive an asymptotically optimal solution to (3.12) which assumes that τ is very large and allows for real-valued τ_i. When sampling variances are also unknown, an additional approximation is needed to account for the fact that closed form solutions for the distribution of differences like $W_{(k)} - W_{(j)}$ are not available (the Behrens–Fisher problem). Thus, we use the so-called Welch approximation to describe those

differences. To do so, we define

$$d_{jk}^{*} = \frac{d_{(j)(k)}}{\widehat{\lambda}_{jk}^{-1/2}}, \tag{3.13}$$

where $d_{(j)(k)} = \overline{x}_{(k)} - \overline{x}_{(j)}$ and $\widehat{\lambda}_{jk}^{-1/2} = \sqrt{\frac{\widehat{\sigma}_{(j)}^2}{n_{(j)}} + \frac{\widehat{\sigma}_{(k)}^2}{n_{(k)}}}$. We also describe the standardized statistics for the difference $Z_{(k)} - Z_{(j)}$ with

$$d_{\{jk\}}^{*} = \frac{d_{(j)(k)}}{\widehat{\lambda}_{\{jk\}}^{-1/2}}, \tag{3.14}$$

where $\widehat{\lambda}_{\{jk\}}^{-1/2} = \sqrt{\frac{\tau_{(j)}\widehat{\sigma}_{(j)}^2}{n_{(j)}\left(n_{(j)}+\tau_{(j)}\right)} + \frac{\tau_{(k)}\widehat{\sigma}_{(k)}^2}{n_{(k)}\left(n_{(k)}+\tau_{(k)}\right)}}$. The notation $d_{\{jk\}}^{*}$ differs slightly from the standardized statistics for the difference $W_{(k)} - W_{(j)}$ that is d_{jk}^{*}.

We denote the Student t distribution by $St(\mu, \kappa, \nu)$, the cumulative distribution function of the standard Student t distribution ($\mu = 0,\ \kappa = 1$) by $\Phi_\nu(\cdot)$, and the probability density function by $\phi_\nu(\cdot)$. The posterior marginal distribution for the unknown mean W_i has a Student t-distribution, $St\left(\overline{x}_i,\ n_i/\sigma_i^2,\ \nu_i\right)$. The standard *EOC* function $\Psi_\nu[m]$ gives the *EOC* when an alternative with known mean m is selected in preference to a single alternative whose unknown mean has a $St(0, 1, \nu)$ distribution,

$$\Psi_\nu[m] = \frac{\nu + m^2}{\nu - 1}\phi_\nu(m) - m\Phi_\nu(-m). \tag{3.15}$$

Welch's approximation for the degrees of freedom of $W_{[k]} - W_{[j]}$ is

$$\nu_{(j)(k)} = \frac{\left[\widehat{\sigma}_{(j)}^2/n_{(j)} + \widehat{\sigma}_{(k)}^2/n_{(k)}\right]^2}{\left[\widehat{\sigma}_{(j)}^2/n_{(j)}\right]^2/\left(n_{(j)}-1\right) + \left[\widehat{\sigma}_{(k)}^2/n_{(k)}\right]^2/\left(n_{(k)}-1\right)}. \tag{3.16}$$

With this notation, and with the Welch approximation for the differences, we can approximate EOC_{bnd} by

$$EOC_{bnd} \approx \sum_{j:(j)\neq(k)} \widehat{\lambda}_{jk}^{-1/2}\Psi_{\nu_{(j)(k)}}\left[d_{jk}^{*}\right] - \widehat{\lambda}_{\{jk\}}^{-1/2}\Psi_{\nu_{(j)(k)}}\left[d_{\{jk\}}^{*}\right]. \tag{3.17}$$

When all of the $\tau_i = 0$, then $\widehat{\lambda}_{\{jk\}}^{-1/2}\Psi_{\nu_{(j)(k)}}\left[d_{\{jk\}}^{*}\right]$ becomes 0. We define

$$EOC_{Bonf} \triangleq EOC_{bnd}(0,0,\ldots,0) = \sum_{j:(j)\neq(k)} \widehat{\lambda}_{jk}^{-1/2}\Psi_{\nu_{(j)(k)}}\left[d_{jk}^{*}\right]. \tag{3.18}$$

EOC_{Bonf} is an upper bound on the posterior EOC given ε (without considering further sampling). Closed form solutions for the EOC are only available in special cases, so the LL procedure is derived by myopically minimizing an upper bound on the EOC at each stage of sampling. The upper bound is essentially a Bonferroni-type bound that emerges from considering the EOC in a comparison between the current best alternative, alternative (k), relative to each of the $k-1$ alternatives. See Chick and Inoue [16] and Branke et al. [4] for further details.

Allocation Rules

Given that there are τ samples to be allocated to k alternatives, the allocation rule for LL is given by

$$\tau_{(i)} = \frac{\left(\tau + \sum_{j=1}^{k} n_j\right)\left(\widehat{\sigma}^2_{(i)}\gamma_{(i)}\right)^{1/2}}{\sum_{j=1}^{k}\left(\widehat{\sigma}^2_{j}\gamma_{j}\right)^{1/2}} - n_{(i)}, \tag{3.19}$$

where

$$\gamma_{(i)} = \begin{cases} \widehat{\lambda}_{ik}^{1/2}\frac{\upsilon_{(i)(k)}+\left(d^*_{ik}\right)^2}{\upsilon_{(i)(k)}-1}\phi_{\upsilon_{(i)(k)}}\left(d^*_{ik}\right), & \text{for } (i)\neq(k) \\ \sum\limits_{j:(j)\neq(k)}\gamma_{(j)}, & \text{for } (i)=(k). \end{cases} \tag{3.20}$$

This formula asymptotically minimizes (3.17) when τ is arbitrarily large (so that all of the $\tau_{(i)}$ are nonnegative). If τ is not sufficiently large, then (3.19) might prescribe a nonpositive number of samples for some $\tau_{(i)}$. If that is the case, then a better approximation to the distribution of the posterior mean $Z_{(i)}$ should be used. Fortunately, one is available: there will be no change in the posterior mean following samples if no new samples are observed. The formulas in (3.19) and (3.20) can therefore be adapted to the case of small τ by making use of that observation. The following steps which check nonnegativity of the $\tau_{(i)}$ implement the necessary computations.

a. Initialize the set of alternatives considered for additional samples, $\mathscr{S} \leftarrow \{1,\ldots,k\}$.
b. For each (i) in $\mathscr{S}\backslash\{(k)\}$: If $(k)\in\mathscr{S}$ then set $\widehat{\lambda}_{ik}^{-1} \leftarrow \widehat{\sigma}^2_{(i)}/n_{(i)} + \widehat{\sigma}^2_{(k)}/n_{(k)}$, and set $\upsilon_{(i)(k)}$ with Welch's approximation. If $(k)\notin\mathscr{S}$ then set $\widehat{\lambda}_{ik} \leftarrow n_{(i)}/\widehat{\sigma}^2_{(i)}$ and $\upsilon_{(i)(k)} \leftarrow n_{(i)} - 1$.
c. Tentatively allocate a total of τ samples to alternatives $(i)\in\mathscr{S}$ $\left(\text{set } \tau_{(j)} \leftarrow 0 \text{ for } (j)\notin\mathscr{S}\right)$:

$$\tau_{(i)} \leftarrow \frac{\left(\tau + \sum_{j \in \mathscr{S}} n_j\right)\left(\widehat{\sigma}^2_{(i)}\gamma_{(i)}\right)^{\frac{1}{2}}}{\sum_{j \in \mathscr{S}}\left(\widehat{\sigma}^2_{(j)}\gamma_j\right)^{\frac{1}{2}}} - n_{(i)},$$

$$\text{where } \gamma_{(i)} \leftarrow \begin{cases} \lambda_{ik}^{1/2} \frac{\upsilon_{(i)(k)} + \left(d^*_{ik}\right)^2}{\upsilon_{(i)(k)} - 1} \phi_{\upsilon_{(i)(k)}}\left(d^*_{ik}\right) & \text{for } (i) \neq (k) \\ \sum\limits_{(j) \in \mathscr{S} \setminus \{(k)\}} \gamma_{(j)} & \text{for } (i) = (k). \end{cases}$$

d. If any $\tau_{(i)} < 0$ then fix the nonnegativity constraint violation: remove (i) from $\mathscr{S}$ for each (i) such that $\tau_{(i)} \leq 0$, and go to Step 4b. Otherwise, round the τ_i so that $\sum_{i=1}^{k} \tau_i = \tau$ (the allocation is determined).

Chick and Inoue [16] derive this algorithm, and generalize it to minimize the CPU time (rather than the number of samples) if the CPU time per sample differs from one alternative to the next. The sequential algorithm is the same as that in Sect. 3.2.2 except that we use the allocation rule in (3.19).

3.4.2 Small-Sample EVI Allocation Rule (LL_1)

The somewhat cumbersome check for nonnegativity of the allocation in the *LL* allocation rule above can be avoided if all samples in a given stage are allocated to a single alternative (the alternative simulated in a given stage can still change from one stage to the next). The LL_1 allocation rule does allocate all samples to one alternative in a given stage, and therefore simplifies the computation of the optimal allocation. This has been done for the simulation context by Chick et al. [18] and independently by Frazier et al. [22]. The latter paper used the term knowledge gradient to describe the idea of one-stage lookahead for the value of sampling from one alternative at each stage of sampling. The LL_1 allocation rule has therefore also been referred to as *KG* or KG_1 in the literature.

Small-Sample *EVI*

As there is only one alternative to be sampled, the small-sample *EVI* procedures avoid the asymptotic approximation, the use of Bonferroni's inequality and the Welch approximation which were employed in the previous *EVI* procedure.

In the small-sample *EVI* that seeks to minimize the posterior *EOC*, the alternative to be sampled is the one with highest $EVI_{LL,(i)}$ where

$$EVI_{LL,(i)} = \begin{cases} \lambda_{\{ik\}}^{-1/2} \Psi_{n_{(i)}-1} \left[d^*_{\{ik\}} \right] & \text{if } (i) \neq (k) \\ \lambda_{\{k-1,k\}}^{-1/2} \Psi_{n_{(k)}-1} \left[d^*_{\{k-1,k\}} \right] & \text{if } (i) = (k) \end{cases}. \tag{3.21}$$

Note that $\bar{x}_{(k)} + EVI_{LL,\,(i)}$ is the expected reward when only taking $\tau_{(i)}$ samples for alternative (i), and then selecting the alternative with the best sample mean. Thus, $EVI_{LL,(i)}$ relates to the expected value of information when comparing the alternative selected for sampling and the best of the other $k-1$ alternatives.

Sequential Algorithm

The following is the procedure for LL_1. It is similar to that for LL except in step 4 where the allocation rule of LL_1 states that only one alternative with the highest EVI will be sampled:

a. Set $\tau_{(i)} \leftarrow \tau$ for the alternative that maximizes $EVI_{LL,(i)}$, and $\tau_\ell \leftarrow 0$ for the others.

3.4.3 Stopping Rules

We now introduce some notation for some stopping rules and more formally describe them. They were originally proposed in the context of *EVI* procedures, but are equally applicable to both *OCBA* and *EVI* contexts.

The deterministic sampling rule, which was used in the *OCBA* algorithm presented in Sect. 3.3.3, continues sampling until a predefined sampling budget has been exhausted. That is, sampling continues if and only if $\sum_{i=1}^{k} n_i < \beta$ for some user-specified β. This is denoted here as the S stopping rule.

There are several adaptive stopping rules that may be used in either the *OCBA* or *EVI* selection procedures—although they were introduced in the context of the *EVI* approach. The EOC_{Bonf} stopping rule continues sampling until the posterior *EOC* is sufficiently small. In particular, sampling continues until an upper bound, EOC_{Bonf} in (3.18), drops below a user-specified threshold $\varepsilon^* > 0$. The threshold ε^* can be chosen to be the highest acceptable expected opportunity cost associated with a possibly incorrect selection. (Someone comfortable with the indifference zone approach might select ε^* to be the indifference zone parameter times the maximal acceptable probability of incorrect selection [19]).

The PCS_{Slep} stopping rule continues sampling until a user-specified lower bound is exceeded by a lower bound on the posterior *PCS*, where the lower bound is due to Slepian. That is, sampling continues until $PCS_{Slep} = \prod_{(i)\neq(k)} \Phi_{v_{(i)(k)}} \left(d^*_{jk} \right)$ is at least as great as a user-specified threshold for the posterior *PCS*, $1-\alpha^*$, where $\alpha^* > 0$ is a user-specified acceptable level of probability of incorrect selection.

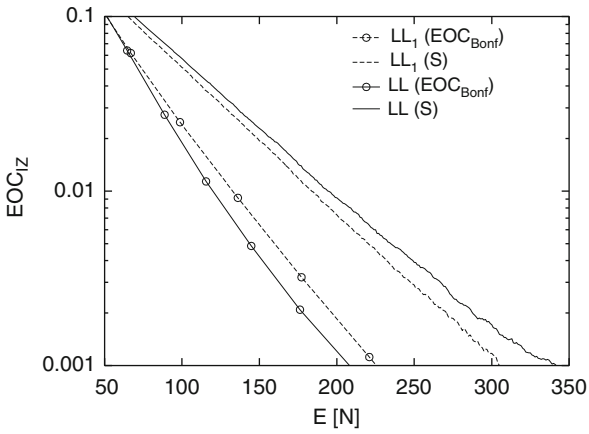

Fig. 3.4 EOC_{IZ} efficiency for LL and LL_1 in a slippage configuration with five alternatives

3.4.4 Numerical Results for LL and LL_1

Branke et al. [4] and Chick et al. [18] provide extensive numerical results for procedures derived from the *OCBA* and *EVI* approaches. This includes procedures based on the *OCBA*, *LL*, and LL_1 allocation rules. They also assess *EVI*-based allocation rules that focus on improving the posterior *PCS*, rather than the posterior *EOC*, and which are named the $0-1$ allocation rule and the $0-1_1$ allocation rule. These allocation rules seek to improve the value of information for the $0-1$ loss function, whose expected value is the posterior *PCS*, when allocating samples to multiple alternatives and to one alternative, respectively, per stage of sampling. The papers also assess how different stopping rules assess the efficiency of the sampling procedures. This section recalls some of the key results from those papers.

In making empirical assessments of selection procedures, allocation rules and stopping rules were combined into a procedure, and were repeatedly applied to different classes of selection problems. Curves like those in Fig. 3.4 were plotted for each selection procedure and for each selection problem. The x-axis of the curve gives the expected total number of samples when the procedure stops, $E[N]$. The y-axis plots the empirical evidence for correct selection. In Fig. 3.4, this evidence is the empirical expected opportunity cost EOC_{IZ}, the average, over repeated applications of the procedure to a selection problem, of the true opportunity cost, $w_{[k]} - w_{\mathfrak{D}}$, when alternative $\mathfrak{D}$ is selected as best. In Fig. 3.5, this evidence is the empirical fraction $PICS_{IZ} = 1 - PCS_{IZ}$, where PCS_{IZ} is the probability of correctly selecting the true best alternative, the fraction estimated over repeated applications of the procedure to a selection problem.

Figure 3.4 presents a representative graph that supports a claim that was systematically found to be true in numerical experiments: the adaptive stopping rules, EOC_{Bonf} and PCS_{Slep}, perform much more effectively than the deterministic stopping rule. This claim is supported in the graph because the curves with the EOC_{Bonf} stopping rule are lower (have a lower loss for any given expected number

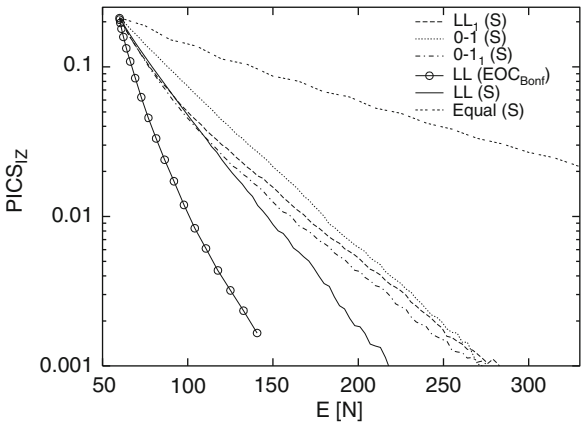

Fig. 3.5 $PICS_{IZ}$ efficiency for LL and LL_1 in a monotonically decreasing means configuration with ten alternatives (means of each alternative evenly spaced)

of samples) than for the (deterministic) S stopping rule. In Fig. 3.4, there are five alternatives in a so-called slippage configuration (the mean of the best alternative is exactly the same value better than the means of the other alternatives).

In Fig. 3.4, we also observe that for a given stopping rule, that the LL allocation rule was at least as good as the LL_1 allocation rule. This was not systematically true. Both of these allocations were equally good for $k = 2$ alternatives. The LL_1 allocation was sometimes better than the LL allocation when there were a relatively small number of total samples (say, 50–120 depending on the problem structure) or when the true means of all competing systems were close to the true mean of the best alternative. As k increases, LL tends to improve in performance relative to LL_1.

In Fig. 3.5, we see that the LL allocation rule can perform even better than the $0-1$ allocation rule for PCS-based figures of merit, even though the latter is specifically designed to improve PCS. This is apparently due to the additional approximations that the $0-1$ loss would seem to impose to obtain analytical results for an easily computable allocation.

Overall, the empirical results showed that the small-sample procedures (LL_1) are competitive if either the number of additional samples allocated is very small, a fixed budget stopping rule is used (as opposed to an adaptive stopping rule such as EOC_{Bonf}), or the number of alternatives is small. This may be the case when alternatives are costly to sample, as when sampling time is very long. In most other settings, the LL performed better. As a general rule, the $OCBA$, the LL and the LL_1 allocations were found to be superior to the equal allocation, many times substantially so.

In terms of overall robustness, Branke et al. [4] and Chick et al. [18] found that the LL allocation or $OCBA_{LL}$ allocation rule (an allocation rule based on the EOC loss function instead of the PCS-based loss function used for the $OCBA$ allocation rule from Sect. 3.3 above; see [33]), with the EOC_{Bonf} stopping rule, proved to be an effective combination in a procedure that works best over a broad class of selection problem structures.

3.4.5 Economics of Selection Procedures (ESP)

The *OCBA*, *LL*, and LL_1 allocations above use criteria to allocate samples during a given stage of sampling that essentially are myopic and greedy: they try to maximize (up to approximations) the evidence for correct selection by the end of the stage. *ESP* procedures account for the value of the ability to continue to sample further after that stage. Thus, it better accounts for the information that a sequential selection procedure can provide.

The *ESP* procedures have been developed assuming a linear loss reward function, as the economic benefit of implementing the alternative selected as best is used to drive both the allocation of samples, and the decision of when to stop sampling. In making the decision for when to stop sampling, samples are assumed to come at a cost. There are at least two ways in which samples might come at a cost. One way is that additional simulation causes delays in implementation. In large scale business decisions, that may cause a discounting penalty due to delays in decisions. Another way that costs may be incurred is from the marginal cost of samples: computer time and resources cost money. Chick and Gans [15] explored the case of discounted sampling, either with or without additional marginal costs per sample. Chick and Frazier [14] handled the case of marginal costs of simulation without discounting.

This section describes the latter case: where simulation samples have costs but there is no additional penalty due to discounting. Thus, to implement them, one ideally needs to estimate the financial cost of sampling (e.g., the cost of run times on a bank of servers) and to have simulation output that expresses financial value. Alternatively, one might use the resulting selection procedure in the absence of those financial interpretations, and use a notional value of the cost to sample in order to determine how many simulation samples to run, and in which order. A reduction of the notional value of the cost to sample would increase the number of samples, if more samples were desired before selecting an alternative.

By its nature, both allocation rules and stopping rules are derived with this approach. The focus on economic criteria to determine sampling plans and when to stop sampling led to the choice of name Economics of Sampling Procedure (*ESP*).

Maximizing Expected Reward

Similar to the previous procedures, we want to maximize the expected reward. However, we go beyond a one-step lookahead for the value of sampling and seek to choose a sequence of alternatives to sample from so that the stream of costs and terminal reward together maximize the expected net reward from the start to the time of selecting an alternative as best.

To describe how to do so, it is useful to introduce some new concepts. A selection policy π is a dynamic method of choosing at each stage t whether to sample an alternative or to stop and select the alternative. The policy, at stage t, can use all information obtained up until stage t. Let $T \in \{t = 0, 1, 2, \ldots\}$ be the stage when

the decision maker decides to stop and select the best alternative to implement. For $t < T$, let $i(t)$ be the index of the alternative to be simulated at stage t and $I(T)$ be the index of the selected alternative. Then the selection policy $\pi = (i(\cdot), T, I(T))$ is the choice of a sequence of alternatives to sample from, the stopping time, and the selected alternative.

Let X_i be the random variable of the unknown reward of alternative i and c_i is the cost per sample of alternative i. Note that c_i and the output X_i are in terms of monetary values. The sampling selection problem is the problem of maximizing expected value of the cost of sequential sampling plus the reward of implementing the alternative selected as best defined by

$$\sup_{\pi} V^{\pi} = \mathrm{E}_{\pi}\left[\sum_{t=0}^{T-1} - c_{i(t)} + X_{I(T),\, T+1} \middle| \mathscr{E}\right]. \tag{3.22}$$

It can be shown that, under certain technical conditions, a policy π maximizes V^{π} if it minimizes the sum of the expected total sampling cost ($\mathrm{E}_{\pi}[\sum_{t=0}^{T-1} - c_{i(t)} | \varepsilon]$) and the expected opportunity cost when an alternative is selected ($\mathrm{E}_{\pi}[\mathscr{L}_{LL}(I(T), \boldsymbol{W}) | \varepsilon]$).

Finding a policy π that achieves the maximum in (3.22) is challenging to solve optimally in general, but Chick and Frazier [14] provided a solution, to an asymptotic approximation, for the special case of comparing one alternative with an unknown mean to a given standard alternative whose mean reward is known to be m. From that special case, they handle the case of $k > 1$ alternatives by using the following heuristic: At each stage, each alternative is assessed to see if it is worth performing more simulations if one were to compare that alternative (with unknown mean) to the mean of the best other alternative (presuming its mean to be known, with m set to the current estimate of the mean of the best other alternative).

Optimal Stopping Problem for the Special Case of $k = 1$ Alternative

Consider comparing one alternative with a known mean reward, m. The sampling selection problem in (3.22) then becomes an optimal stopping problem. Let $c, \bar{x}_t$ and n_t be the cost per sample, the sample mean, and the number of samples seen so far up to time t for the one alternative. The Bellman's recursion for this problem is

$$V^*(m, \bar{x}_t, n_t) = \max\left\{m, -c + \mathrm{E}\left[V^*\left(m, \frac{n_t \bar{x}_t + X_{i,t+1}}{n_t + 1}, n_t + 1\right) \middle| \bar{x}_t, n_t\right], \bar{x}_t\right\}. \tag{3.23}$$

The first maximand in Bellman's recursion indicates selecting the known alternative. The second maximand includes the cost of observing one sample from the alternative with the unknown mean and thenhaving the option to select

an alternative or to continue sampling, given the information from that sample. The third maximand is associated with stopping to sample and to implement the alternative with the unknown mean.

Chick and Frazier [14] show how to convert the discrete time dynamic program in (3.23) into a continuous time problem. That continuous time problem is a free boundary problem for a heat equation, and whose solution gives an asymptotic approximation to the expected value of the option to continue sampling to learn more (a non-myopic, multistage version of the one-stage value of information calculations used in the LL and LL_1 procedures), and upper and lower optimal stopping boundaries. If the statistics of the alternative with the unknown mean are below the lower stopping boundary, it would be optimal to stop sampling and to select the known alternative for reward m. If the statistics of the alternative with an unknown mean are above the upper boundary, the alternative with the unknown mean should be selected as best with no further sampling. If the statistics of the alternative with the unknown mean are between the upper and lower stopping boundaries, then one additional sample should be taken for that alternative, and the process should be repeated.

Solving this problem for practical use in a procedure could require computing the solution of a partial differential equation with a free boundary. Fortunately, Chick and Frazier [14] provide a numerical approximation to the solution to the free boundary problem which has shown to be useful in problems and which does not require heavy mathematical machinery for implementations. In particular, they show that the upper and lower optimal stopping boundaries are given by

$$m \pm c_i^{1/3} \sigma_i^{2/3} b\left(\sigma_i^{2/3} / c_i^{2/3} n_i\right), \tag{3.24}$$

for some function $b(s) \geq 0$ for $s \geq 0$, and that a useful approximation to $b(s)$ is

$$\tilde{b}(s) = \begin{cases} 0.233 s^2 & \text{if } 0 \leq s \leq 1, \\ 0.00537 s^4 - 0.06906 s^3 + 0.3167 s^2 - 0.02326 s & \text{if } 1 < s \leq 3, \\ 0.705 s^{1/2} \ln(s) & \text{if } 3 < s \leq 40, \\ 0.642 \left[s(2\ln(s))^{1.4} - \ln(32\pi) \right]^{1/2} & \text{if } 40 < s. \end{cases} \tag{3.25}$$

The region that is between the upper and lower stopping boundaries in (3.24) for $n_i > 0$ is called the continuation region (because it determines where sampling should continue).

ESP Allocation Rule and Stopping Rule

The approach for the optimal stopping problem can be extended in a heuristic way to provide a new allocation rule and a new stopping rule when there are $k > 1$ alternatives.

The stopping rule is motivated and presented first. From the previous subsection, a given alternative warrants continued simulation if it is between the upper and lower stopping boundaries in (3.24). Adapted to the case of $k > 1$ alternatives, each with unknown mean, one can substitute the value of m in (3.24) with the mean of the alternative with the best estimated mean (specifically, the highest posterior mean of the other alternatives). This motivates what we will call the *ESP_b stopping rule*: sampling continues as long as there is at least one alternative that would merit additional sampling if it were considered in comparison with the best of the other alternatives (at least up to asymptotic approximations). More formally, sampling continues as long as there is at least one alternative i such that

$$c_i^{1/3}\sigma_i^{2/3} b\left(\sigma_i^{2/3}/\left(c_i^{2/3} n_{i,\,t}\right)\right) > \widehat{\Delta}_{i,\,t}, \tag{3.26}$$

where $\widehat{\Delta}_{i,\,t} = \left|\bar{x}_{i,t} - \max_{j\neq i}\bar{x}_{j,t}\right|$ is the difference in expected value between the alternative i and the best of other alternatives (including the known mean standard m) conditional on information through stage t, and $n_{i,\,t}$ is the number of samples from alternative i through stage t. When the sampling variance σ_i^2 is unknown, it is appropriate in implementations to substitute the sample variance $\widehat{\sigma}_i^2$ for it in (3.26), as well as in (3.27) below [14].

The *ESP_b allocation rule* allocates one sample to the alternative that is 'furthest inside' the continuation region, in a standardized coordinate alternative that is natural to consider in this application [14]. In particular, one should sample from the alternative that is the solution to

$$\arg\max_i b\left(\sigma_i^{2/3}/\left(c_i^{2/3} n_{i,\,t}\right)\right) - \widehat{\Delta}_{i,\,t}/\left(c_i^{1/3}\sigma_i^{2/3}\right). \tag{3.27}$$

The term $c_i^{1/3}\sigma_i^{2/3}$ in (3.26) is the cube root of the product of the sampling costs and sampling variance, which is inversely proportional to the sampling efficiency [32].

Chick and Frazier [14] provide another allocation rule that is based more directly on the diffusion approximation for the expected reward of continuing to sample. But that other allocation rule requires access to the full solution of the free boundary problem, and is therefore less attractive from an implementation standpoint.

They also provide numerical results that illustrate the performance of a selection procedure with various allocation rules and stopping rules. The *LL* allocation rule was not tested with the ESP_b stopping rule, because *ESP* is based on approximations that presume one sample is taken at a time. Table 3.4, adapted from Chick and Frazier [14], shows the expected total penalty for not knowing the mean rewards $\mathrm{E}[cT + OC]$ for five different combinations between allocation rules and stopping rules calculated using Monte Carlo Simulation with 10^6 samples for $k = 2, 5, 10, 20$ and 10^5 samples for $k = 100$. Here, the sampling cost was assumed to be $c = 1$, T is the number of samples observed when an alternative is selected as best, and OC is the realized opportunity cost of selecting that alternative.

Table 3.4 Expected total penalty for not knowing the mean rewards $E[cT + OC]$ for several allocation and stopping rules

		k				
Allocation rule	Stopping rule	2	5	10	20	100
LL	EOC_{Bonf}	320	577	821	1,095	2,168
Equal allocation	EOC_{Bonf}	321	629	1,040	1,815	8,425
LL_1	EOC_{Bonf}	318	546	728	916	1,577
LL_1	ESP_b	231	506	694	875	1,516
ESP_b	ESP_b	233	505	700	856	1,308

In Table 3.4, the stopping rule appears to be more influential than the allocation rule. Specifically, the ESP_b stopping rule performs better than the EOC_{Bonf} stopping rule for each given allocation rule. Performance with the ESP_b stopping rule improves even upon the performance from the EOC_{Bonf} stopping rule with these allocations in all experiments run so far (the ESP_b stopping rule has been subject to less testing to date than has the EOC_{Bonf} stopping rule). This appears to be because the ESP_b stopping rule considers the benefit of multiple future stages of sampling when considering when to stop, whereas the EOC_{Bonf} stopping rule only looks one step ahead.

The ESP_b stopping rule works very well with the LL_1 and ESP_b allocation rules, and either of those options can be recommended for the selection problem.

3.4.6 Other Variants of EVI

Apart from ESP_b, Chick and Gans [15] propose another economics of selection procedure where there is a positive discount cost. That algorithm, which has somewhat different stopping boundaries due to the discounting, also has a stopping boundary which is approximated with an easy to compute function. That stopping boundary was shown to be related to the optimal stopping boundary for the Bayesian bandit problem when samples from each bandit are independent Gaussian whose means are unknown (and are inferred through sampling).

In addition, Chick and Gans [15] provide a discussion on whether the decision maker should develop a simulation platform. They proposed a view of simulation as an option to learn more about alternatives before making a selection, and a mechanism to quantify the value of that learning. That value could be compared with the time and financial costs of developing the simulation model in the first place: responding to the question "To simulate, or not to simulate?" when a Bayesian prior distribution can be used to quantify uncertainty about the potential financial benefit of various simulated alternatives, at least in some contexts. The approach is therefore different from most R&S work in simulation which assumes that a simulation platform has been built and is available.

There are also several other one-step lookahead policies to myopically maximize *EVI*. Frazier et al. [23] consider the case where the prior beliefs about the rewards are correlated and Ryzhov et al. [52] apply it to propose a new type of online learning policy. Frazier et al. [24] consider the case with both correlations in the prior belief and correlation in sampling which is achieved through common random numbers. The use of common random number is also considered in the work by Chick and Inoue [17], which minimizes the expected opportunity cost in two stages.

3.5 Conclusion

This chapter discussed some Bayesian approaches to the problem of selecting the best from a small- to medium-sized set of alternatives, where best is the largest mean and the means are to be estimated through sampling. The samples are presumed to be independent, as are the values of the unknown means. The approaches reviewed include *OCBA* and *EVI*, which have variations that are among the most effective known today. The variations are with different allocation rules and stopping rules. The *OCBA* algorithm is very effective in cases where *PCS* is of particular interest. To date, it appears that the most effective procedures in a very broad range of tests are based either on the *LL* or $OCBA_{LL}$ allocation rules in conjunction with the EOC_{Bonf} stopping rule if an adaptive stopping rule is allowed, or with a fixed budget stopping rule if that is required. It appears that the LL_1 and ESP_b allocation rule are even more effective when used in conjunction the ESP_b stopping rule, in cases where sampling costs can be estimated and the output of the simulation has a financial impact. An analysis of tweaking a notional cost of sampling when such conditions do not hold is yet to be done, but it would appear that such an approach could be more broadly applicable. More testing on that would be useful. The papers cited above also explain how a number of other practicalities might be addressed, such as the question of including common random numbers. The codes of the algorithms presented in this chapter can be obtained from the authors.

In cases where the number of alternatives is so large that it is not possible to simulate all alternatives at least a few times, some other techniques may be required. The simulation optimization literature addresses this problem. For example, see Chap. 2 and [25, 35, 56] for reviews. In such cases, one might integrate an R&S procedure with a search algorithm [42, 53]. He et al. [34] extend the *OCBA* notion to the cross-entropy method for combinatorial problems. Or one might use a response surface technique so that information observed for one alternative can be informative for the mean values of other alternatives. Initial work on this approach that is related to the *EVI* approach above includes the use of Kriging models for the unknown means of the alternatives when samples are either independent [22, 23] or correlated such as with common random numbers [24]. Brantley et al. [5, 6] develop a new *OCBA* procedure under the use of a regression equation and partitioning of the decision domain.

Acknowledgements This work has been supported in part by National Science Foundation under Award CMMI-1233376, Department of Energy under Award DE-SC0002223, NIH under Grant 1R21DK088368-01, National Science Council of Taiwan under Award NSC-100-2218-E-002-027-MY3, and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning.

References

1. R. E. Bechhofer, T. J. Santner, and D. M. Goldsman. *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*. John Wiley & Sons, New York, 1995.
2. J. O. Berger and J. Deely. A Bayesian approach to ranking and selection of related means with alternatives to analysis-of-variance methodology. *Journal of the American Statistical Association*, 83(402):364–373, 1988.
3. J. Blanchet, J. Liu, and B. Zwart. Large deviations perspective on ordinal optimization of heavy-tailed systems. In S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. Fowler, editors, *Proceedings of the 2008 Winter Simulation Conference*, 489–494, 2008.
4. J. Branke, S. E. Chick, and C. Schmidt. Selecting a selection procedure. *Management Science*, 53:1916–1932, 2007.
5. M. W. Brantley, L. H. Lee, C. H. Chen, and A. Chen. Efficient simulation budget allocation with regression. *IIE Transactions*, 45(3):291–308, 2013.
6. M. W. Brantley, L. H. Lee, C.-H. Chen, and J. Xu. An efficient simulation budget allocation method incorporating regression for partitioned domains. *Automatica*, 50(5):1391–1400, 2014.
7. A. Charnes, W. W. Cooper, and E. Rhodes. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2:429–444, 1978.
8. C. H. Chen and L. H. Lee. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. World Scientific Publishing Co, 2011.
9. C. H. Chen, D. He, and M. C. Fu. Efficient dynamic simulation allocation in ordinal optimization. *IEEE Transactions on Automatic Control*, 51(12):2005–2009, 2006.
10. C. H. Chen, D. He, M. C. Fu, and L. H. Lee. Efficient simulation budget allocation for selecting an optimal subset. *INFORMS Journal of Computing*, 20:579–595, 2008.
11. C. H. Chen, V. Kumar, and Y.C. Luo. Motion planning of walking robots in environments with uncertainty. *Journal of Robotic Systems*, 16(10):527–545, 1999.
12. C. H. Chen, J. Lin, E. Yücesan, and S. E. Chick. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems: Theory and Applications*, 10:251–270, 2000.
13. C. H. Chen, E. Yücesan, L. Dai, and H. C. Chen. Efficient computation of optimal budget allocation for discrete event simulation experiment. *IIE Transactions*, 42(1):60–70, 2010.
14. S. E. Chick, and P. Frazier. Sequential sampling with economics of selection procedures. *Management Science*, 58(3):550–569, 2012.
15. S. E. Chick, and N. Gans. Economic analysis of simulation selection problems. *Management Science*, 55(3):421–437, 2009.
16. S. E. Chick and K. Inoue. New two-stage and sequential procedures for selecting the best simulated system. *Operations Research*, 49(5):732–743, 2001.
17. S. E. Chick and K. Inoue. New procedures to select the best simulated system using common random numbers. *Management Science*, 47(8):1133–1149, 2001.
18. S. E. Chick, J. Branke, and C. Schmidt. Sequential sampling to myopically maximize the expected value of information. *INFORMS Journal of Computing*, 22:71–80, 2010.
19. S. E. Chick and Y. Z. Wu. Selection procedures with frequentist expected opportunity cost bounds. *Operations Research*, 53(5):867–878, 2005.
20. M. H. de Groot. *Optimal Statistical Decisions*. McGraw-Hill, New York, 1970.

21. E. J. Dudewicz and S.R. Dalal. Allocation of observations in ranking and selection with unequal variances. *The Indian Journal of Statistics,* 37B, 1:28–78, 1975.
22. P. Frazier, W. B. Powell, and S. Dayanik. A knowledge-gradient policy for sequential information collection. *SIAM Journal on Control and Optimization*, 47(5):2410–2439, 2008.
23. P. Frazier, W. B. Powell, and S. Dayanik. The knowledge-gradient policy for correlated normal beliefs. *INFORMS Journal of Computing*, 21:599–613, 2009.
24. P. Frazier, J. Xie, and S. E. Chick. Bayesian optimization via simulation with correlated sampling and correlated prior beliefs. In S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, M. Fu, editors, *Proceedings of the 2011 Winter Simulation Conference*, 3979–3911, 2011.
25. M. C. Fu, C. H. Chen, and L. Shi. Some topics for simulation optimization. In S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. Fowler, editors, *Proceedings the 2008 Winter Simulation Conference*, 27–38, 2008.
26. M. C. Fu, F. W. Glover, and J. April. Simulation optimization: A review, new developments, and applications. In M. E. Kuhl, N. M. Steiger, F. B. Armstrong, J. A. Joines, editors, *Proceedings the 2005 Winter Simulation Conference*, 83–95, 2005.
27. M. C. Fu, J. Q. Hu, C. H. Chen, and X. Xiong. Simulation allocation for determining the best design in the presence of correlated sampling. *INFORMS Journal of Computing*, 19:101–111, 2007.
28. P. Glynn and S. Juneja. A large deviations perspective on ordinal optimization. In R. G. Ingalls, M. D. Rosetti, J. S. Smith, B. A. Peters, editors, *Proceedings of the 2004 Winter Simulation Conference,* 577–585, 2004.
29. D. Goldsman and B. L. Nelson. Comparing systems via simulation. In J. Banks, editor, *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*. John Wiley & Sons, New York, 273–306, 1998.
30. S. S. Gupta and K. J. Miescke. Bayesian look ahead one stage sampling allocations for selecting the largest normal mean. *Statistical Papers*, 35:169–177, 1994.
31. S. S. Gupta and K. J. Miescke. Bayesian look ahead one-stage sampling allocations for selecting the best population. *Journal of Statistical Planning and Inference*, 54(2):229–244, 1996.
32. J. M. Hammersley and D.C. Hanscomb. *Monte Carlo Methods*. Methuen, London, 1964.
33. D. He, S. E. Chick, and C. H. Chen. The opportunity cost and OCBA selection procedures in ordinal optimization. *IEEE Transactions on Systems, Man, and Cybernetics—Part C (Applications and Reviews)*, 37(4):951–961, 2007.
34. D. He, L. H. Lee, C. H. Chen, and M. C. Fu, and S. Wasserkrug. Simulation optimization using the cross-entropy method with optimal computing budget allocation. *ACM Transactions on Modeling and Computer Simulation*, 20(1):Article 4, 2010.
35. L. J. Hong and B. L. Nelson. A brief introduction to optimization via simulation. In M. D. Rosetti, R. R. Hill, B. Johansson, A. Dunkin, R. Ingalls, editors, *Proceedings of the 2009 Winter Simulation Conference*, 75–85, 2009.
36. B. W. Hsieh, C. H. Chen, and S.C. Chang. Scheduling semiconductor wafer fabrication by using ordinal optimization-based simulation. *IEEE Transactions on Robotics and Automation*, 17(5):599–608, 2001.
37. Q. S. Jia, E. Zhou, and C. H. Chen. Efficient computing budget allocation for finding simplest good designs. *IIE Transactions*, 45(7):736–750, 2013.
38. S.-H. Kim and B. L. Nelson. A fully sequential procedure for indifference-zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation*, 11:251–273, 2001.
39. S.-H. Kim. and B. L. Nelson. Selecting the best system. In S. G. Henderson, B. L. Nelson, editors, *Handbooks in Operations Research and Management Science: Simulation*. Elsevier Science, Oxford, 501–534, 2006.
40. L. W. Koenig and A. M. Law. A procedure for selecting a subset of size m containing the one best of k independent normal populations, with applications to simulation. *Communications in Statistics* B14, 3:719–734, 1985.
41. A. M. Law. *Simulation Modeling and Analysis*, 5th Edition. McGraw-Hill, New York, 2013.

42. L. H. Lee, E. P. Chew, and P. Manikam. A general framework on the simulation-based optimization under fixed computing budget. *European Journal of Operational Research*, 174:1828–1841, 2006.
43. L. H. Lee, E. P. Chew, S. Teng, and D. Goldsman. Optimal computing budget allocation for multi-objective simulation models. In R. G. Ingalls, M. D. Rosetti, J. S. Smith, B. A. Peters, editors, *Proceedings of the 2004 Winter Simulation Conference*, 586–594, 2004.
44. L. H. Lee, E. P. Chew, S. Teng, and D. Goldsman. Finding the non-dominated Pareto set for multi-objective simulation models. *IIE Transactions*, 42:656–674, 2010.
45. L. H. Lee, N. A. Pujowidianto, L. W. Li, C. H. Chen, and C. M. Yap. Approximate simulation budget allocation for selecting the best system in the presence of stochastic constraints. *IEEE Transactions on Automatic Control*, 57(11):2940–2945, 2012.
46. D. J. Morrice, M. W. Brantley, and C. H. Chen. An efficient ranking and selection procedure for a linear transient mean performance measure. In S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. Fowler, editors, *Proceedings of the 2008 Winter Simulation Conference*, 290–296, 2008.
47. D. J. Morrice, M. W. Brantley, and C. H. Chen. A transient means ranking and selection procedure with sequential sampling constraints. In M. D. Rosetti, R. R. Hill, B. Johansson, A. Dunkin, R. Ingalls, editors, *Proceedings of the 2009 Winter Simulation Conference*, 590–600, 2009.
48. B. L. Nelson, J. Swann, D. Goldsman, and W. M. Song. Simple procedures for selecting the best simulated system when the number of alternatives is large. *Operations Research*, 49(6):950–963, 2001.
49. W. B. Powell and I. O. Ryzhov. *Optimal Learning*. Wiley, 2012.
50. Y. Rinott. On two-stage selection procedures and related probability inequalities. *Communications in Statistics* A7:799–811, 1978.
51. R. Rubinstein. The cross-entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability*, 1(2): 127–190, 1999.
52. I. O. Ryzhov, W. B. Powell, and P. Frazier. The knowledge-gradient algorithm for a general class of online learning problems. *Operations Research*, 60(1):180–195, 2012.
53. L. Shi and C. H. Chen. A new algorithm for stochastic discrete resource allocation optimization. *Discrete Event Dynamic Systems: Theory and Applications*, 10:271–294, 2000.
54. J. F. Shortle, C. H. Chen, B. Crain, A. Brodsky, and D. Brod. Optimal splitting for rare-event simulation. *IIE Transactions*, 44(5):352–367, 2012.
55. R. Szechtman and E. Yücesan. A new perspective on feasibility determination. In S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. Fowler, editors, *Proceedings of the 2008 Winter Simulation Conference*, 273–280, 2008.
56. E. Tekin and I. Sabuncuoglu. Simulation optimization: A comprehensive review on theory and applications. *IIE Transactions*, 36:1067–1081, 2004.
57. L. Trailovic and L. Y. Pao. Computing budget allocation for efficient ranking and selection of variances with application to target tracking algorithms. *IEEE Transactions on Automatic Control*, 49:58–67, 2004.
58. R. Waeber, P. Frazier, and S. G. Henderson. Performance measures for ranking and selection procedures. In B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, E. Yücesan, editors, *Proceedings of the 2010 Winter Simulation Conference*, 1235–1245, 2010.
59. W. P. Wong, W. Jaruphongsa, and L. H. Lee. Budget allocation for effective data collection in predicting an accurate DEA efficiency score. *IEEE Transactions on Automatic Control*, 56(6):1235–1245, 2011.
60. S. Yan, E. Zhou, and C. H. Chen. Efficient selection of a set of good enough designs with complexity preference. *IEEE Transactions on Automation Science and Engineering*, 9(3):596–606, 2012.