Sequential Sampling to Myopically Maximize the Expected Value of Information

Stephen E. Chick
INSEAD; Technology and Operations Management Area; Boulevard de Constance; F-77305 Fontainebleau, France, stephen.chick@insead.edu, http://faculty.insead.edu/chick/

Jürgen Branke
Institute AIFB; University of Karlsruhe (TH); D-76128 Karlsruhe, Germany, branke@aifb.uni-karlsruhe.de, http://www.aifb.uni-karlsruhe.de/~jbr

Christian Schmidt
Locom Software GmbH; Stumpfstr. 1, D-76131 Karlsruhe, Germany, christian.schmidt@locom.de, http://www.locom.com/de/

Statistical selection procedures are used to select the best of a finite set of alternatives, where “best” is defined in terms of each alternative’s unknown expected value, and the expected values are inferred through statistical sampling. One effective approach, which is based on a Bayesian probability model for the unknown mean performance of each alternative, allocates samples based on maximizing an approximation to the expected value of information (EVI) from those samples. The approximations include asymptotic and probabilistic approximations. This paper derives sampling allocations that avoid most of those approximations to the EVI, but entails sequential myopic sampling from a single alternative per stage of sampling. We demonstrate empirically that the benefits of reducing the number of approximations in the previous algorithms is typically outweighed by the deleterious effects of a sequential one-step myopic allocation when more than a few dozen samples are allocated. Theory clarifies the derivation of selection procedures that are based on the EVI.

Key words: 055 Decision analysis : Sequential; 793 Statistics, Bayesian; 902 Simulation : Statistical analysis

History: Submitted December 9, 2007; Revised November 28, 2008; Revised March 7, 2009

Selection procedures are used to select the best of a finite set of alternatives, where best is determined with respect to the largest mean, and the mean is inferred via statistical sampling (Bechhofer et al., 1995). Applications include stochastic simulation and drug testing. Because the unknown mean performance of each alternative is estimated with statistical samples, it is not possible to guarantee that the best alternative is selected with probability 1 in finite time. A natural goal is to minimize the expected loss from incorrect selections with a limited number of samples. The expected loss is usually measured by the probability of incorrect selection (PICS) or the expected opportunity cost (EOC) of potentially selecting an alternative that is not the best.

Table 1 cites some key references for several different approaches to this problem. The oldest and most extensive stream is the ranking and selection (R&S) stream, which has typically sought to provide an upper bound on the PICS, though some work has been done to bound the EOC.
Table 1  Selection problems can be approached with different assumptions and objectives.

<table>
<thead>
<tr>
<th>Objective</th>
<th>PICS</th>
<th>EOC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Large Deviation</td>
<td>Glynn and Juneja (2004)</td>
</tr>
<tr>
<td>Bayesian</td>
<td>OCBA</td>
<td>Chen (1996), Chen et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>EVI</td>
<td>Gupta and Miescke (1996), Chick and Inoue (2001)</td>
</tr>
<tr>
<td></td>
<td>Economic</td>
<td>-</td>
</tr>
</tbody>
</table>

An approach via large deviation theory allows for heavy-tailed samples, whereas most R&S work assumes normal or Bernoulli samples. Both of these frequentist approaches measure PICS with respect to the worst-case fraction of correct selections in repeated applications of a selection procedure to each selection problem in a specified class of problems. Bayesian approaches use an average loss criterion, rather than bounding worst-case losses. This is done by using Bayes’ rule to infer the distributions of the unknown means as data is observed. The optimal computing budget allocation (OCBA) allocates samples to improve an approximation to the expected loss, based on a thought experiment with posterior distributions of the unknown means. The expected value of information (EVI) approach, on the other hand, allocates samples to maximize the EVI that will be obtained from sampling, using predictive distributions of further samples. An economic approach that models the cost of sampling and a firm’s discount rate also uses ideas from the EVI approach with the EOC objective. Most Bayesian work to date assumes that samples have a normal distribution.

Chen et al. (2006) note an advantage of sequential procedures over two-stage procedures that dominated early work. Branke et al. (2007) compared several highly sequential sampling procedures that use R&S, OCBA or EVI allocations. A highly sequential procedure allocates one or a few samples per stage of sampling. Certain versions of the OCBA and EVI procedures were demonstrated to be particularly effective in numerical tests. EVI procedures can also serve as part of an economic approach to selection (Chick and Gans 2009).

This article assesses whether further improvements in the most effective of the existing EVI procedures can be made by eliminating certain approximations that appear in their derivation. The most effective EVI and OCBA procedures (Chick and Inoue 2001, Chen et al. 2000, He et al. 2007) use asymptotic approximations (in the number of samples), approximations to the distribution of the difference of two random variables with t distributions (e.g., Welch’s approximation for the so-called Behrens-Fisher problem), and approximations to bound the probability of certain events (e.g., Bonferroni’s or Slepian’s inequality). In particular, this article derives small-sample EVI
procedures for both the PICS and EOC selection criteria that avoid an asymptotic approximation, the Behrens-Fisher problem, and the need to use Bonferroni’s or Slepian’s inequality. Because those approximations are averted, one might expect that the new small-sample EVI procedures would perform better than the original large-sample counterparts.

We show that the new small-sample EVI procedures provide only a small improvement in special situations, such as when only a few samples can be observed or as a way to allocate the last few samples of other sequential procedures. When a medium to large number of samples are observed, the small sample approach is less effective – the penalty of myopically applying an optimal small-sample allocation sequentially outweighs the penalty from the approximations in the original large-sample counterparts. The theory below also clarifies the derivation of selection procedures that are based on the EVI. We assume that samples are independent and normally distributed with unknown means and variances. An appendix describes computational techniques that improve the numerical stability of the new EVI procedures when they are pushed to extreme levels of correct selection. Preliminary numerical results appeared in Chick et al. (2007). Frazier et al. (2008) used dynamic programming to independently develop a similar allocation, called the knowledge gradient, for the case of normally distributed samples with known variances.

1. Problem Formulation and Selection Procedures

We first formalize the problem and describe the assumptions, then describe measures of evidence for correct selection that will be useful for deriving the new procedures.

1.1. Assumptions and Notation

The best of $k$ alternatives is to be identified, where ‘best’ denotes the largest mean. Let $X_{ij}$ be a random variable whose realization $x_{ij}$ is the output of the $j$th sample from alternative $i$, for $i = 1, \ldots, k$ and $j = 1, 2, \ldots$. Let $w_i$ and $\sigma_i^2$ be the unknown mean and variance of alternative $i$, and 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, alternative $[k]$, is to be identified by sampling. We assume that

$$\{X_{ij} : j = 1, 2, \ldots\} \sim \text{Normal}(w_i, \sigma_i^2), \text{ for } i = 1, \ldots, k.$$ 

Vectors are written in boldface, such as $\mathbf{w} = (w_1, \ldots, w_k)$ and $\mathbf{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_k^2)$.

A problem instance (configuration) is denoted by $\chi = (\mathbf{w}, \mathbf{\sigma}^2)$. 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 $\hat{\sigma}_i^2 = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2/(n_i - 1)$ be the sample variance. Let $\bar{x}_{(1)} \leq \bar{x}_{(2)} \leq \ldots \leq \bar{x}_{(k)}$ be the ordering of the sample means based on all samples seen so far. The quantities $n_i$, $\bar{x}_i$, $\hat{\sigma}_i^2$ and $(i)$ are updated as more samples are observed.
Each selection procedure generates estimates $\hat{w}_i$ of $w_i$, for $i = 1, \ldots, k$. For the procedures studied here, $\hat{w}_i = \bar{x}_i$, and a correct selection occurs when the selected alternative, alternative $\mathcal{D}$, is the best alternative, $[k]$. Here, alternative $\mathcal{D} = (k)$ is selected as best.

The objective of a selection procedure can be defined in terms of a loss function. The PICS objective can be defined as an expectation of the zero-one loss function, $L_{0-1}(\mathcal{D}, w) = 1 \{ w_\mathcal{D} \neq w[k] \}$, where the indicator function $1 \{ \cdot \}$ equals 1 if its argument is true, and is 0 otherwise. The EOC objective can be defined as an expectation of $L_{LL}(\mathcal{D}, w) = w[k] - w_\mathcal{D}$, the difference between the best and the selected alternative (which is zero if the best is correctly selected).

We will need to describe random variables with Student $t$ distributions. If $T_\nu$ is a random variable with standard $t$ distribution with $\nu$ degrees of freedom, we denote (as do Bernardo and Smith 1994) the distribution of $\mu + \frac{1}{\sqrt{\nu}} T_\nu$ by $St(\mu, \kappa, \nu)$. If $\nu > 2$, then the variance is $\kappa^{-1}\nu/(\nu - 2)$. As $\kappa \to \infty$, $St(\mu, \kappa, \nu)$ converges in distribution to a point mass at $\mu$. Denote the cumulative distribution function (cdf) of the standard $t$ distribution ($\mu = 0, \kappa = 1$) by $\Phi_\nu()$ and probability density function (pdf) by $\phi_\nu()$. The standard EOC function $\Psi_\nu[s]$ gives the EOC when an alternative with known mean $s$ is selected in preference to a single alternative whose unknown mean has a $St(0,1,\nu)$ distribution,

$$\Psi_\nu[s] = \int_{u=s}^{\infty} (u - s) \phi_\nu(u) du = \frac{\nu + s^2}{\nu - 1} \phi_\nu(s) - s \Phi_\nu(-s).$$

Assuming a noninformative prior distribution for the unknown mean and variance, the posterior marginal distribution for the unknown mean $W_i$, given $n_i > 2$ samples, is $St(\bar{x}_i, n_i/\hat{\sigma}^2_i, \nu_i)$, where $\nu_i = n_i - 1$ (e.g., de Groot 1970). We assume a noninformative prior distribution for the rest of the paper. Branke et al. (2007) describe how to adapt the analysis to conjugate prior distributions.

1.2. Evidence for Correct Selection

There are several frequentist measures that describe a procedure’s ability to identify the actual best alternative. They are defined as expected values of the loss function, the expectation taken over the samples $X_{ij}$ (the realizations $x_{ij}$ determine the selected alternative, $\mathcal{D}$), conditional on the problem instance $\chi = (w, \sigma^2)$. The frequentist PICS is denoted by $\text{PICS}_{IZ}(\chi)$

$$\text{PICS}_{IZ}(\chi) \overset{\text{def}}{=} \mathbb{E}[L_{0-1}(\mathcal{D}, w) \mid \chi] = \Pr(w_\mathcal{D} < w[k] \mid \chi).$$

The frequentist opportunity cost of incorrect selections is $\text{EOC}_{IZ}(\chi) \overset{\text{def}}{=} \mathbb{E}[L_{LL}(\mathcal{D}, w) \mid \chi]$.

---

1 $LL$ stands for linear loss, an alternative name for opportunity cost, and avoids naming conflicts with the OCBA.

2 The subscript $IZ$ denotes indifference zone, a frequentist approach to selection procedures.
Bayesian measures of performance of sampling procedures are defined with respect to the posterior distribution of the unknown means, given all data $\mathcal{E} = \{(x_{i1}, x_{i2}, \ldots, x_{in_i})$, for $i = 1, 2, \ldots, k\}$ seen so far. The Bayesian measures of interest here are

$$
PICS_{\text{Bayes}} \overset{\text{def}}{=} E[\mathcal{L}_{D-1}(\mathcal{D}, \mathcal{W}) | \mathcal{E}] = \Pr \left( W_D < \max_{i \not\in \mathcal{D}} W_i | \mathcal{E} \right),$$

$$EOC_{\text{Bayes}} \overset{\text{def}}{=} E[\mathcal{L}_{LL}(\mathcal{D}, \mathcal{W}) | \mathcal{E}] = E \left[ \max_{i=1,2,\ldots,k} W_i - W_D | \mathcal{E} \right].$$

We define $PCS_{\text{Bayes}} = 1 - PICS_{\text{Bayes}}$.

Approximations in the form of bounds on the above losses are useful below to derive sampling allocations and to define stopping rules. We define

$$d_{jk}^r = d_{(j)(k)} \lambda^{-1/2}_{jk}$$

with $d_{(j)(k)} = \bar{x}(k) - \tilde{x}(j)$ and $\lambda^{-1}_{jk} = \frac{\hat{\sigma}^2_{(j)}}{n(j)} + \frac{\hat{\sigma}^2_{(k)}}{n(k)}$.

$$\nu_{(j)(k)} = \frac{[\hat{\sigma}^2_{(j)}/n(j) + \hat{\sigma}^2_{(k)}/n(k)]^2}{[\hat{\sigma}^2_{(j)}/n(j)]^2/(n(j) - 1) + [\hat{\sigma}^2_{(k)}/n(k)]^2/(n(k) - 1)},$$

where $\nu_{(j)(k)}$ comes from Welch’s approximation for the difference $W_{(k)} - W_{(j)}$ of two $t$ random variables (Law and Kelton 2000, p. 559). Slepian’s inequality and the Bonferroni inequality (e.g., Kim and Nelson 2006) imply that the posterior evidence that alternative $(k)$ is best satisfies

$$PCS_{\text{Bayes}} \geq \prod_{j:(j)\neq(k)} \Pr \left( W_{(k)} > W_{(j)} | \mathcal{E} \right) \approx \prod_{j:(j)\neq(k)} \Phi_{\nu_{(j)(k)}}(d_{jk}^r) \overset{\text{def}}{=} PCS_{\text{Step}},$$

$$PCS_{\text{Bayes}} \geq 1 - \sum_{j:(j)\neq(k)} \Pr \left( W_{(k)} < W_{(j)} | \mathcal{E} \right) \approx 1 - \sum_{j:j\neq(k)} \Phi_{\nu_{(j)(k)}}(-d_{jk}^r) \overset{\text{def}}{=} PCS_{\text{Bonf}},$$

$$EOC_{\text{Bayes}} \leq \sum_{j:(j)\neq(k)} \int_{-\infty}^{\infty} w f_{(j)(k)}(w) dw \approx \sum_{j:j\neq(k)} \lambda^{-1/2}_{jk} \Psi_{\nu_{(j)(k)}}[d_{jk}^r] \overset{\text{def}}{=} EOC_{\text{Bonf}},$$

where $f_{(j)(k)}(\cdot)$ is the posterior pdf for the difference $W_{(j)} - W_{(k)}$ given $\mathcal{E}$. That difference has approximately a $St \left(-d_{(j)(k)}, \lambda_{jk}, \nu_{(j)(k)}\right)$ distribution.

2. Sequential EVI Procedures

This section formalizes the problem of maximizing the EVI in a given stage of a selection procedure, identifies approximations in previous EVI work, and indicates how new procedures below can avoid most of those approximations.

2.1. Setup for Optimizing the EVI

The EVI approach allocates a finite number of samples at each stage of a sequential sampling procedure, either to minimize the expected loss once those samples are observed or to balance the expected reduction in loss from those samples with the expected cost of sampling. Minimizing the expected loss following sampling is equivalent to maximizing the EVI from those samples. Here, we
focus on allocating \( \tau > 0 \) samples to the \( k \) alternatives in order to maximize the EVI with respect to the PICS and EOC objectives in a single additional stage of sampling.

We need additional notation to account for the random output that will be seen during the next stage of sampling. Let \( Y = (Y_1, Y_2, \ldots, Y_k) \) denote the data to be collected in the next stage of sampling, with \( \tau_i \geq 0 \) samples to be observed for alternatives \( i = 1, 2, \ldots, k \). Let \( y = (y_1, y_2, \ldots, y_k) \) be the realization of \( Y \), let \( \bar{y}_i \) be the sample average of the output of alternative \( i \) in that additional stage if \( \tau_i > 0 \) and set \( \bar{y}_i = 0 \) if \( \tau_i = 0 \). Then the overall sample mean for alternative \( i \) is

\[
\bar{z}_i = \frac{n_i \bar{x}_i + \tau_i \bar{y}_i}{n_i + \tau_i}.
\]

Before \( Y \) is observed, the sample average that will arise after sampling, \( Z_i = (n_i \bar{x}_i + \tau_i \bar{Y}_i)/(n_i + \tau_i) \), is random. Using the predictive distribution for \( Y \), one can show (de Groot 1970) that

\[
Z_i \sim \mathcal{N} \left( \bar{x}_i, n_i(n_i + \tau_i)/(\tau_i \sigma^2_i), n_i - 1 \right).
\]

The following standardized statistics for the difference \( Z_{(j)} - Z_{(j)} \) will be useful below.

\[
d^2_{(j)} = \lambda_{(j)}^{-1/2} d_{(j)} \quad \text{where} \quad \lambda_{(j)}^{-1} = \left( \frac{\tau_{(k)} \hat{\sigma}^2_{(k)}}{n_{(k)}(n_{(k)} + \tau_{(k)})} + \frac{\tau_{(j)} \hat{\sigma}^2_{(j)}}{n_{(j)}(n_{(j)} + \tau_{(j)})} \right)
\]

Finally, let \( \mathcal{D}(Y) = \arg \max_{i=1,\ldots,k} \{Z_i\} \) be the alternative with the highest overall sample mean.

We can now define the problem of allocating \( \tau > 0 \) samples to \( k \) alternatives in order to maximize the EVI of a single additional stage of sampling:

\[
\min_{\tau_1, \tau_2, \ldots, \tau_k \geq 0} \mathbb{E}[\mathcal{L}(\mathcal{D}(Y), W)] \\
\text{such that} \quad \tau = \sum_{i=1}^k \tau_i
\]

where \( \mathcal{L} = \mathcal{L}_{0-1} \) for the PICS objective and \( \mathcal{L} = \mathcal{L}_{LL} \) for the EOC objective. Minimizing the predicted expected loss in (3) is equivalent to maximizing the EVI.

### 2.2. Asymptotic EVI Allocations for PICS and EOC

Problem 3 does not have a nice closed form solution. This section presents the asymptotic EVI allocations for PICS and EOC objectives that Chick and Inoue (2001) used for Procedures 0-1 and LL, respectively, identifies the approximations that were used to derive those allocations, and fills in details for the derivation of Procedure 0-1 (Chick and Inoue 2001, outline of proof of Theorem 2).

We start with the 0-1 loss function and note that \( \mathbb{E}[\mathcal{L}_{0-1}(\mathcal{D}(Y), W)] = \mathbb{E}[\mathbb{E}[\mathcal{L}_{0-1}(\mathcal{D}(Y), W) \mid Y]] \).

The inner expectation can be more easily analyzed with a modified loss function. We subtract \( \mathcal{L}_{0-1}((k), W) \) from the objective function in (3) to obtain an auxiliary loss function,

\[
\mathcal{L}^*_0((i), W) = \mathcal{L}_{0-1}((i), W) - \mathcal{L}_{0-1}((k), W)
\]
\[
\begin{align*}
\lambda_{ik}^{-1} &= \frac{\hat{\sigma}^2_{(k)}}{n_{(k)} + \tau_{(k)}} + \frac{\hat{\sigma}^2_{(i)}}{n_{(i)} + \tau_{(i)}}.
\end{align*}
\]

Note that \( \Pr \left( V_{(i)} > 0 \mid z_{(i)} - z_{(k)} \right) \approx \Phi_{\hat{\nu}_{(i)}}(\hat{\lambda}_{ik}^{1/2}(z_{(i)} - z_{(k)})) \).

The distributions of differences like \( U_{(i)} = Z_{(i)} - Z_{(k)} \) are needed to assess the probability that alternative \( (i) \) is selected after \( Y \) is observed. The distribution of \( U_{(i)} \) is approximately
\[
U_{(i)} \sim \mathcal{St} \left( -d_{(i)(k)}, \lambda_{(i)(k)}, \nu_{(i)(k)} \right),
\]
where the degrees of freedom \( \nu_{(i)(k)} \) is estimated with Welch’s approximation if needed.

Suppose that the order statistics reflect the order of the \( E[W_j] \) before sampling (so that \( \bar{x}_{(1)} = \cdots = \bar{x}_{(k)} \)). Alternative \( (i) \) is selected after sampling when \( Z_{(i)} > Z_{(j)} \) for all \( j \neq (i) \). Further, the conditional distribution of \( W \) given \( Y \) only depends on \( Y \) through \( Z \). (5) therefore implies
\[
E \left[ \mathcal{L}^*_{0-1} (\mathfrak{D}(Y), W) \right] = \sum_{i: i \neq k} \Pr \left( W_{(i)} \leq W_{(k)}, Z_{(i)} > Z_j \text{ for all } j \neq (i) \right) \\
- \Pr \left( W_{(i)} > W_{(k)}, Z_{(i)} > Z_j \text{ for all } j \neq (i) \right).
\]

We now analyze (7) for the special case of a pairwise comparison between alternatives \( (i) \) and \( (k) \).
\[
E \left[ \mathcal{L}^*_{0-1} (\mathfrak{D}(Y), W) \right] = \Pr \left( Z_{(i)} > Z_{(k)} \right) \left( 1 - 2\Pr \left( W_{(i)} > W_{(k)} \mid Z_{(i)} > Z_{(k)} \right) \right) \\
\approx E \left[ \mathbf{1} \left\{ U_{(i)} > 0 \right\} \left( 1 - 2\Phi_{\hat{\nu}_{(i)}}(\hat{\lambda}_{ik}^{1/2} V_{(i)}) \right) \right].
\]
The approximation in (8) comes from the Welch approximations above for \( U_{(i)} \) and for \( V_{(i)} \).

We now add the \( \mathcal{L}_{0-1} ((k), W) \) that was subtracted earlier and take expectations,
\[
E[\mathcal{L}_{0-1} (\mathfrak{D}(Y), W)] = E[\mathcal{L}_{0-1} ((k), W)] + E[\mathcal{L}^*_{0-1} (\mathfrak{D}(Y), W)]
\]
\[
\sum_{(i)\neq (k)} \Phi_{\nu((i)k)}(-\lambda_{ik}^{1/2} d_{(i)(k)}) - \Phi_{\nu((i)k)}(-\lambda_{ik}^{1/2} d_{(i)(k)})
\]

(10)

Substituting (10) for the objective function in (3), solving with Lagrange multipliers for the constraint \(\sum \tau_i = \tau\), and finally noting that \(\lambda_{ik} \rightarrow \lambda_{ik}\) as \(\tau \rightarrow \infty\) (the second asymptotic approximation for this analysis), leads to the following allocation of samples for each alternative:

\[
\tau_{(i)} = \frac{(\tau + \sum_{j=1}^{k} n_j)(\hat{\sigma}_{(i)}^2 + \hat{\gamma}_j)^{1/2}}{\sum_{j=1}^{k} (\hat{\sigma}_{(i)j}^2 + \hat{\gamma}_j)^{1/2}} - n_{(i)}
\]

(11)

where \(\gamma_i = \{\lambda_{ik} d_{ik}^2 \phi_{\nu((i)k)}(d_{ik}^2) \text{ for } (i) \neq (k)\}
\sum_{j:j \neq k} \gamma_{(j)} \text{ for } (i) = (k)\).

Procedure 0-1 allocates \(\tau\) samples to the \(k\) alternatives at each stage of a sequential sampling procedure with (11), after an initial stage when \(n_0 > 2\) samples are taken from each alternative.

Procedure 0-1 is the asymptotic EVI procedure for the PICS objective. A full description of Procedure 0-1, including how to insure that each \(\tau_i \geq 0\) when \(\tau\) is small, is found in Appendix A.

Procedure LL is the EOC objective analog of Procedure 0-1. Procedure LL also allocates with (11), but with \(\gamma_{(i)}\) given by

\[
\gamma_{(i)} = \left\{\begin{array}{ll}
\lambda_{ik}^{1/2} \nu_{(i)(k)}^{(d_{ik}^2) + (d_{ik}^2)} \phi_{\nu((i)k)}(d_{ik}^2) / \nu_{(i)(k)} - 1 \sum_{j:j \neq k} \gamma_{(j)} & \text{for } (i) \neq (k) \\
\sum_{j:j \neq k} \gamma_{(j)} & \text{for } (i) = (k)
\end{array}\right.
\]

(13)

The full derivation of LL is in Chick and Inoue (2001). In order to facilitate the derivation of its
small-sample counterpart below, we note a few parallels with the above development. The auxiliary objective function for EOC that is analogous to (1) is

$$L_{LL}^\ast ((i), w) = L_{LL} ((i), w) - L_{LL} ((k), w)$$

$$= \begin{cases} w(k) - w(i) & \text{if } (i) \neq (k) \text{ and } (i) \text{ is best} \\ 0 & \text{if } (i) = (k) \text{ and } (k) \text{ is best.} \end{cases}$$

$$E[L_{LL}^\ast ((i), W) | Y] = z(k) - z(i)$$  \hspace{1cm} (14)

For a pairwise comparison between $(i)$ and $(k)$, the expected loss is approximately (using Welch’s approximation, Chick and Inoue 2001)

$$E[L_{LL} ((i), W) | Y] = E[L_{LL} ((k), W)] + E[L_{LL}^\ast ((i), W) | Y]$$

$$= \lambda_{ik}^{-1/2} \psi_{(i)(k)} \left[ \lambda_{ik}^{1/2} d_{(i)(k)}(k) \right] \lambda_{ik}^{-1/2} \psi_{(i)(k)} \left[ \lambda_{ik}^{1/2} d_{(i)(k)}(k) \right]$$  \hspace{1cm} (15)

We note that the EVI is the expected improvement in the expected loss that comes from sampling, and so the EVI of such a pairwise comparison is approximately $\lambda_{ik}^{-1/2} \lambda_{ik}^{1/2} d_{(i)(k)}(k)$. The Bonferroni approximation and Welch approximation are used in the derivation of (13) for the same reasons they are used in 0-1, as is the asymptotic approximation $\lambda_{ik} \rightarrow \lambda_{ik}$ as $\tau \rightarrow \infty$. There is no need in the EOC analysis to approximate $Pr(W_{(i)} > W_{(k)} | Z(i) > Z(k))$ by 1 because there is no probability statement in the right hand side of (14), unlike (5).

2.3. Adapting the Analysis to New Small-Sample EVI Procedures

The new EVI procedures are motivated by the fact that Welch’s approximation can be avoided, for a given stage of the procedure, if only one alternative gets all the samples. In particular, Welch’s approximation is only needed in (1) if both $\tau_{(i)} > 0$ and $\tau_{(k)} > 0$. If samples are taken for alternative $(k)$ but not for alternative $(i)$, then the correct degrees of freedom for $U_{(i)}$ is $\nu_{(i)(k)} = n(k) - 1$ because $Z(i) \rightarrow \bar{x}(i)$ when $\tau(i) \rightarrow 0$. If samples are taken only for alternative $(i)$ then $\nu_{(i)(k)} = n(i) - 1$.

We therefore specialize the analysis of Section 2.2 to procedures that allocate all samples to a single alternative in a given stage. We call them small-sample EVI procedures. The EOC objective is handled before the PICS objective.

When $\tau(j) > 0$ for one alternative but all of the other $\tau(i) = 0$, all but one of the $Z(i)$ are distributed as a point mass at $\bar{x}(i)$. The one alternative with $\tau(j) > 0$ has $Z(j) \sim \text{St}(\bar{x}(j), n(j)(n(j) + \tau(j))/(\tau(j)\bar{x}(j)), n(j) - 1)$ distribution, so that

$$E[L_{LL}^\ast ((j), W) | Y] = \begin{cases} \bar{x}(k) - Z(j) & \text{if } (j) \neq (k), Z(j) > \bar{x}(k) \text{ (alternative } j \text{ is new best)} \\ 0 & \text{if } (j) \neq (k), Z(j) \leq \bar{x}(k) \text{ (alternative } k \text{ remains best)} \\ 0 & \text{if } (j) = (k), Z(j) \geq \bar{x}(k) \text{ (old best remains best)} \\ Z(k) - \bar{x}(k-1) & \text{if } (j) = (k), Z(j) < \bar{x}(k-1) \text{ (old best no longer best)} \end{cases}$$  \hspace{1cm} (16)
shows that the EVI, with respect to the EOC objective, of allocating all $\tau$ samples to a single alternative can be described by a comparison between the alternative that gets the samples and the current mean of the best of the other alternatives. The EVI of $\tau(j) > 0$ samples for alternative $(j)$ and none for the others is therefore exactly and non-asymptotically

$$EVI_{LL,(j)} = \begin{cases} \lambda_{\{j\}}^{-1/2} \Psi n_{(j)}^{-1} \left[ d_{\{ik\}}^* \right] & \text{if } (j) \neq (k) \\ \lambda_{\{k\}}^{-1/2} \Psi n_{(j)}^{-1} \left[ d_{\{ik\}}^* \right] & \text{if } (j) = (k) \end{cases},$$

where $\lambda_{\{jk\}}$ and $d_{\{ik\}}^*$ are as in (2). The Bonferroni bound is eliminated because only one of two alternatives can be selected as best if only one alternative is simulated–either the alternative sampled becomes the best, or the best of the other $k-1$ alternatives becomes the new best if the value output is low enough to lower that alternative in the rank order. We note that $x_{\{k\}} + EVI_{LL,(j)}$ is the expected reward when only taking $\tau(j)$ samples for alternative $(j)$, then selecting the alternative with the best sample mean.

justifies the allocation of $LL_1$, the following new small-sample EVI procedure for the EOC objective. The names of the new procedures are distinguished from their asymptotic counterparts by the subscript $1$ (one alternative gets all samples in a given stage of sampling).

**Procedure $LL_1$** allocates all $\tau$ samples in each stage after the first to the alternative $(j)$ that maximizes $EVI_{LL,(j)}$. In the first stage, $n_0 > 2$ samples are taken from each alternative.

**Procedure 0-1** allocates all $\tau$ samples in each stage after the first to the alternative $(i)$ that maximizes the EVI with respect to the 0-1 loss function. That EVI is approximately

$$EVI_{0-1,(i)} = \begin{cases} \Phi_{n_{\{i\}}^{-1}}(-d_{\{ik\}}^*) & \text{if } (i) \neq (k) \\ \Phi_{n_{\{k\}}^{-1}}(-d_{\{k-1,k\}}^*) & \text{if } (i) = (k) \end{cases}.$$  

Note that $EVI_{0-1,(i)}$ appears in the subtrahend of the summand on the right hand side of (10), as anticipated for the EVI of sampling. The values of $EVI_{LL,(i)}$ and $EVI_{0-1,(i)}$ both avoid the Bonferroni and Welch approximations and one asymptotic allocation by allocating samples to only one alternative. The value $EVI_{0-1,(i)}$ does not avert approximating $\Pr(W_{(i)} > W_{(k)} | Z_{(i)} > Z_{(k)})$ by 1. With that last approximation, we can still simplify the objective function to (10).

Appendix C adapts these small-sample allocations for normally distributed samples with known variances, as well as how they relate to the work of Gupta and Miescke (1996).

### 3. Empirical Results

This section summarizes the main qualitative conclusions of a numerical study that was designed to answer the following questions.

1. Are the new small-sample procedures more effective than their asymptotic counterparts?
2. Does the difference in the effectiveness between the new and asymptotic procedures depend on the type of stopping rule? Stopping rules are described below in Section 3.1.

3. If there are differences in performance, does it make sense to switch from one allocation to another, to create an improved hybrid procedure?

Appendix D describes techniques to improve the numerical stability for EVI\(_{L,i}\) and EVI\(_{0-1,i}\) relative to using built-in libraries for those functions, when a large number of samples are observed.

### 3.1. Summary of Procedures and Their Evaluation

We assessed the asymptotic procedures (LL and 0-1) and the new procedures (LL\(_1\) and 0-1\(_1\)) by running an initial stage of sampling with \(n_0\) samples per alternative, then observing \(\tau\) samples per stage. We also tested the “equal” allocation that allocates samples in round robin fashion. We used default values of \(n_0 = 6\) and \(\tau = 1\). The value \(n_0 = 6\) yielded good results in Branke et al. (2007) for LL and 0-1, as did smaller values of \(\tau\).

For a given configuration, sampling continued sequentially until a stopping rule was satisfied. Stopping rules for sequential procedures may depend on Bayesian measures [Branke et al., 2007], which are functions of the data \(E\) observed so far. Stopping rules tested include:

1. Sampling budget rule (\(S\)): Continue sampling if \(\sum_{i=1}^{k} n_i < B\) for some specified budget \(B\).

2. PCS\(_{\text{Slep}}\) rule: Continue sampling if PCS\(_{\text{Slep}} < 1 - \alpha^*\) for a specified probability \(1 - \alpha^* \in [1/k, 1)\) and given \(\delta^* \geq 0\).

3. EOC\(_{\text{Bonf}}\) rule: Continue sampling if EOC\(_{\text{Bonf}} > \beta^*\), for a specified EOC target \(\beta^* > 0\).

An adaptive stopping rule allows the total number of samples to depend on (adapt to) the observed values of the samples seen so far. For example, PCS\(_{\text{Slep}}\) and EOC\(_{\text{Bonf}}\) are adaptive stopping rules.

Procedures are named with the stopping rule in parentheses, as in 0-1(\(S\)) and LL\(_1\)(EOC\(_{\text{Bonf}}\)).

The efficiency of the procedures is measured below in terms of the average number of samples \(E[N]\) needed to reach a given frequentist PICS or EOC, including the initial stage of sampling. For the \(S\) stopping rule, the number of replications is exactly \(E[N] = kn_0 + B\). For the adaptive stopping rules (PCS\(_{\text{Slep}}\) and EOC\(_{\text{Bonf}}\)) we estimate \(E[N]\) as a sample average.

All results reported are averages of 100,000 applications of the selection procedure to a given configuration. To sharpen the contrasts between different procedures, common random numbers (CRN) were used to generate common random problem instances, and to synchronize the samples across procedures. Samples were independent and normally distributed within each procedure.

### 3.2. Summary of Configurations

Appendix B provides additional information about the details for how the configurations were structured and parameterized for the experiments. In summary, we applied each procedure to
several different configurations to assess their performance in a variety of conditions. We varied the number of alternatives, \( k = 2, 5, 10, 20 \).

For the vector of means of the configurations, we tested

- the slippage configuration (SC) where the best alternative has a mean that is exactly \( \delta \) better than each of the \( k - 1 \) other alternatives, for several \( \delta \);
- the monotone decreasing means (MDM) configuration where the means from one alternative to the next is \( \delta \) for several \( \delta \) (with \( w_i = i\delta \), for \( i = 1, 2, \ldots, k \));
- random problem instances (RPI) where configurations are sampled with a random mean from an independent normal distribution.

The variances were assumed to follow a variety of patterns for the SC and MDM configuration, including common equal variances; better alternatives have larger variances; better alternatives have smaller variances. For the RPI, variances were selected randomly and independently for each alternative using an inverted gamma distribution. The structure of the test problems that we checked is therefore similar to the test set in Branke et al. (2007), but is not as comprehensive.

### 3.3. Summary of Results

Claims in this section that are not supported with graphical or numerical data are supported by Appendix B of the online companion.

#### 3.3.1. New small-sample procedures versus their asymptotic counterparts.

In a first comparison between the procedures, we assume throughout this section that a fixed total sampling budget is available (stopping rule \( S \)). All four allocations (\( LL_1(S), LL(S), 0-1_1(S), 0-1(S) \)) were essentially equally efficient when there were \( k = 2 \) alternatives, whether the variances of each were the same or somewhat different for each type of configuration (SC, which is the same as MDM when \( k = 2 \), and RPI).

When \( k = 5, 10, 20 \), the new allocations (\( LL_1(S), 0-1_1(S) \)) both performed slightly better than their asymptotic counterparts (\( LL(S), 0-1_1(S) \)) for the SC. For the MDM configuration, the new allocations were only better than the asymptotic allocations for a relatively small number of additional samples (from one to a several dozen, depending on the specific spacing of the means and relative size of the variances). For example, Figure 1 shows the new allocations beating \( LL(S) \) over the range from 60 to 110 samples, when \( LL(S) \) becomes best among procedures with the \( S \) stopping rule. This is consistent with intuition, since the small sample procedures benefit from the removal of some approximations, while the disadvantage of being myopic increases with the number of samples. For a medium to larger number of samples, the asymptotic allocations were better than...
the new allocations for the MDM configurations. For the RPI configurations, the $LL_1(S)$, $LL(S)$, $0-1_1(S)$ allocations performed relatively similarly, and $0-1(S)$ was the worst of the four procedures for that allocation. As in Branke et al. (2007), $LL(S)$ consistently outperformed $0-1(S)$, which consistently outperformed Equal($S$).

For a given total sampling budget, then, the new allocations perform better than, or at least as good as, the asymptotic allocations when the slippage configurations is used or if problem configurations are randomly sampled from the RPI setup above. In the MDM configuration, the new procedures were worse than the asymptotic procedures once a medium to large number of samples was observed.

3.3.2. The change in effectiveness and the stopping rule. The relative merits of the new allocations versus the old allocations change somewhat when the adaptive stopping rules, PCS$_{Slep}$ and EOC$_{Bonf}$, are used. Figure 2 exemplifies this point. With adaptive stopping rules, the new small-sample allocations are no longer more efficient than the asymptotic allocations, except when $k = 2$ or when fewer than a few dozen samples were collected (depending on the parameters of the SC, MDM or RPI configuration). In cases when the new allocations were better, they were only slightly better than the asymptotic allocations.

The benefit of using $LL$ rather than $LL_1$, with an adaptive stopping rule, increases with $k$.

3.3.3. Hybrid procedures. A hybrid procedure can switch back and forth between two or more allocations at different stages of sampling. Because the new EVI procedures can perform more efficiently than their asymptotic EVI counterparts when only a few samples are taken, we assessed several hybrid procedures.
Figure 2 Influence of the stopping rule on the \( LL \) and \( LL_1 \) allocations (SC, \( k = 5 \), equal variances).

- \( LL_1(S < X)LL \) takes \( n_0 \) samples from each alternative in an initialization stage, then allocates \( X \) samples with the \( LL_1 \) allocation, then the remaining samples with the \( LL \) allocation.

- \( LL(S < X)LL_1 \) takes \( n_0 \) samples from each alternative, then allocates \( X \) samples with the \( LL \) allocation, then the remaining samples with the \( LL_1 \) allocation.

- A procedure that switches back and forth between \( LL \) and \( LL_1 \) at regular intervals.

The tests demonstrated two main observations. One, the efficiency is slightly improved for a small number of samples (typically 5-20 in our tests) after first switching from \( LL \) to \( LL_1 \), at which point the efficiency curve becomes somewhat less effective than would be the case if the \( LL \) allocation were retained. Figure 3 demonstrates this point. Two, the efficiency curve tends to rapidly realign itself to the efficiency curve of the second allocation after switching from one allocation to the other. There was no significant improvement to variations where the allocation was switched back and forth multiple times. Thus, there may be a slight improvement when switching from \( LL \) to \( LL_1 \) for the last dozen samples.

4. Discussion and Conclusion

The choice of the selection procedure and its parameters can have a tremendous effect on the effort spent to select the best alternative. Branke et al. (2007) showed that certain EVI procedures (\( LL \)) and the OCBA procedures are highly effective, especially with adaptive stopping rules.

The new small-sample EVI procedures that are derived in this paper avoid Bonferroni’s and Welch’s approximations as well as an asymptotic approximation. This paper therefore provides a potential net benefit relative to previous EVI derivations which assumed several asymptotic and probabilistic approximations. The small-sample procedures suffer from a myopic allocation that
presumes that a selection will be made after one allocation. Repeatedly applying that allocation greedily in order to obtain a sequential procedure results in a sub-optimality. That sub-optimality is more significant in numerical tests if the number of samples allocated before selection is large (i.e., if very low levels of $PICS_{IZ}$ are desired). An area of potential future research is to use other optimization techniques besides the above small-sample approach and the asymptotic normality approximation for EVI to see if further efficiency improvements can be obtained.

Our empirical results show that the small-sample procedures are competitive if either the number of additional samples allocated is very small, a fixed budget is used as stopping rule (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 simulation replications are very long. In most other settings, the new allocations are either somewhat less effective or much less effective than the asymptotic allocations. When the new procedures are better, they are only mildly better.

The poorer performance of the new procedures, relative to the asymptotic procedures, is not due to coordinate optimization. Coordinate optimization can lead to suboptimality relative to algorithms with better gradient information in some contexts (e.g., Markov Chain Monte Carlo can mix more slowly with coordinate-based Gibbs samplers than with other samplers in some cases [Gilks et al. 1996]). But in the examples here, all EVI procedures allocated 1 sample per stage. They each depended on coordinate steps in the same way.

The asymptotic allocations are designed to sample for long run behavior, and therefore perform well when more stringent levels of evidence are required. Still, if the asymptotic $LL$ allocation is used for the bulk of the sampling, there is a small efficiency improvement that is associated with switching to $LL_1$ to allocate the last few samples.
Acknowledgments
We appreciate the constructive feedback of the editors and referees that helped us to improve the paper.

Appendix A: Full Versions of Sequential EVI Procedures
This appendix gives complete pseudo-code for the EVI procedures in the paper. Procedures 0-1 and LL are sequential variations of the asymptotic EVI procedures in Chick and Inoue (2001) that iteratively improve PCSBonf and EOCBonf, respectively. Procedures 0-1 and LL1 are their new small-sample counterparts.

Procedure 0-1.
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 \( X_1, \ldots, X_{n_0} \) independently, and initialize the number of samples \( n_i \leftarrow n_0 \) so far for each alternative, \( i = 1, \ldots, k \).
3. Determine the sample statistics \( \bar{x}_i \) and \( \hat{\sigma}_i^2 \), and the order statistics, so that \( \bar{x}_1 \leq \cdots \leq \bar{x}_k \).
4. WHILE stopping rule not satisfied DO another stage:
   (a) Initialize the set of alternatives considered for additional samples, \( S \leftarrow \{1, \ldots, k\} \).
   (b) For each \( (i) \) in \( S \backslash \{k\} \): If \( (k) \) then set \( \lambda_{ik} \leftarrow \hat{\sigma}_{i(k)}^2 / n_{(i)} + \hat{\sigma}_{(k)}^2 / n_{(k)} \), and set \( \nu_{(i)(k)} \) with Welch’s approximation. If \( (k) \notin S \) then set \( \lambda_{ik} \leftarrow n_{(i)} / \hat{\sigma}_{(k)}^2 \) and \( \nu_{(i)(k)} \leftarrow n_{(i)} - 1 \).
   (c) Tentatively allocate a total of \( \tau \) samples to alternatives \( (i) \in S \) (set \( \tau_{(j)} = 0 \) for \( (j) \notin S \)):
   \[
   \tau_{(i)} \leftarrow \frac{\tau + \sum_{j \in S} n_{(j)}(\hat{\sigma}_{(i)}^2 / \gamma_{(i)})^{\frac{2}{3}}}{\sum_{j \in S} (\hat{\sigma}_{(j)}^2 / \gamma_{(j)})^{\frac{2}{3}}} - n_{(i)},
   \]
   where \( \gamma_{(i)} \leftarrow \left\{ \begin{array}{ll}
   \lambda_{ik} d_{ik}^* \Phi_{\nu_{(i)(k)}}(d_{ik}^*) & \text{for } (i) \neq (k) \\
   \sum_{(j) \in S \backslash \{(k)\}} \gamma_{(j)} & \text{for } (i) = (k).
   \end{array} \right. \)
   (d) If any \( \tau_i < 0 \) then fix the nonnegativity constraint violation: remove \( (i) \) from \( S \) for each \( (i) \) such that \( \tau_{(i)} \leq 0 \), and go to Step 4b. Otherwise, round the \( \tau_i \) so that \( \sum_{i=1}^k \tau_i = \tau \) and go to Step 4c.
   (e) Observe \( \tau_i \) additional samples for alternative \( i \), for \( i = 1, \ldots, k \). Update sample statistics \( n_i \leftarrow n_i + \tau_i \); \( \bar{x}_i, \hat{\sigma}_i^2 \), and the order statistics, so that \( \bar{x}_1 \leq \cdots \leq \bar{x}_k \).
5. Select the alternative with the best estimated mean, \( \mathcal{D} = (k) \).

Procedure LL. Same as Procedure 0-1, except in Step 4e set
\[
\gamma_{(i)} \leftarrow \left\{ \begin{array}{ll}
   \lambda_{ik}^{1/2} \Psi_{\nu_{(i)(k)}}^* [d_{ik}^*]^{\frac{2}{3}} & \text{for } (i) \neq (k) \\
   \sum_{(j) \in S \backslash \{(k)\}} \gamma_{(j)} & \text{for } (i) = (k).
   \end{array} \right. \]

Procedure LL1. Same as Procedure 0-1, except replace Steps 4a-4d by:
- For each \( i \in \{1, 2, \ldots, k\} \), see if allocating to \( (i) \) is best:
  - Tentatively set \( \tau_{(i)} \leftarrow \tau \) and \( \tau_\ell \leftarrow 0 \) for all \( \ell \neq (i) \); set \( \lambda_{(i)}^\ell, d_{\ell(k)}^* \) with \( \square \) for all \( j \).
  - Compute the EVI of the samples with respect to the opportunity cost loss function,
    \[
    \text{EVI}_{LL, (i)} = \left\{ \begin{array}{ll}
   \lambda_{(i)}^{1/2} \Phi_{\nu_{(i)(-1)}}^{-1} [d_{(k)}^*] & \text{if } (i) \neq (k) \\
   \lambda_{(k)-1}^{1/2} \Phi_{\nu_{(k)(-1)}}^{-1} [d_{(k-1,k)}^*] & \text{if } (i) = (k).
   \end{array} \right. \]
- Set \( \tau_{(i)} \leftarrow \tau \) for the alternative that maximizes EVI_{LL, (i)}, and \( \tau_\ell \leftarrow 0 \) for the others.

Procedure 0-1. Same as Procedure LL1, except the EVI is for the 0-1 loss function, as approximated by
\[
\text{EVI}_{0-1, (i)} = \left\{ \begin{array}{ll}
   \Phi_{\nu_{(i)(-1)}}^{-1} (-d_{(i)}^*) & \text{if } (i) \neq (k) \\
   \Phi_{\nu_{(k)(-1)}}^{-1} (-d_{(k-1,k)}^*) & \text{if } (i) = (k).
   \end{array} \right. \]
References


Appendix B: More Empirical Results

This section provides additional descriptions of the various configurations that were tested in the numerical study, as well as graphs that describe the output for some of those configurations. Tests were run under the assumption that there were \( k = 2, 5, 10 \) or 20 alternatives.

Note that none of the expected value of information procedures yet have a provable frequentist performance guarantee. On the other hand, the EOC\(_{\text{Bonf}}\) and PCS\(_{\text{Slep}}\) stopping rules provide Bayesian guarantees for the posterior expectation of the opportunity costs and probability of incorrect selections, to the extent that the Welch approximation approximates those probabilities well. For a more detailed comparison of some frequentist ranking and selection work with Bayesian OCBA and EVI procedures, please see Inoue et al. (1999) and Branke et al. (2007).

B.1. Monotone decreasing means

In a monotone decreasing means (MDM) configuration the means are equally spaced with distance \( \delta \). We assume independent outputs for alternatives \( i = 1, \ldots, k \) that are normally distributed,

\[
X_{ij} \sim \text{Normal} \left( -(i-1)\delta, 2\rho^2/(1+\rho) \right). 
\]

The parameter \( \rho \) controls the variances. Values of \( \rho < 1 \) mean that better alternatives have a smaller variance. We tested values of \( \delta = 0.25, 0.5, 1.0 \) and 2.0 with values of \( \rho = 0.707, 1 \) and 1.414 (related to \( \sqrt{2} \)).

Figure 4 compares the asymptotic EVI procedures LL and 0-1 to the small-sample counterparts LL\(_1\) and 0-1\(_1\), together with an efficient adaptive stopping rule, EOC\(_{\text{Bonf}}\). The small-sample procedures slightly outperform the original versions if only a few samples are taken (corresponding to relatively high levels of empirical PICS\(_{12}\)). This is the scenario that the small-sample procedures have been designed for, and the removal of the Bonferroni and Welch approximations pays off.

The situation changes completely with more stringent levels of PICS\(_{12}\), as shown in Figure 5. In that case, the asymptotic procedures outperform the small-sample procedures by a wide margin, and the difference grows with increasing \( \mathbb{E}[N] \). Still, the small-sample procedures are much better than the naïve Equal allocation.

Figure 1 shows the output for the same configuration, but with a fixed sampling budget (\( S \) stopping rule) as opposed to an adaptive stopping rule. Figure 6 zooms in on Figure 1 for the case of a small number of replications. All allocations become less efficient with the fixed sampling budget. That observation is consistent with the fact that several other allocations also perform less effectively with a fixed sampling budget (\( S \)) than with the (EOC\(_{\text{Bonf}}\)) stopping rule (Branke et al. 2007). The small-sample allocations do well as long as there are only a few samples allocated with the \( S \) stopping rule. Then LL allocation outperforms
the new allocations once a medium number of samples have been allocated. Procedure 0-1(S) is the worst procedure for this experiment as it uses the most approximations in its derivation. The two small-sample procedures perform similarly and somewhere between $LL(S)$ and $0-1(S)$ for large $E[N]$.

The influence of the stopping rule is even more apparent in Figure 7 which compares three different stopping rules for $LL$ and $LL_1$. For $LL$, the EOC$_{Bonf}$ stopping rule is clearly the most efficient, followed
Figure 6  PICS\textsubscript{IZ} efficiency for the original and new allocations with \(S\) stopping rule (MDM, \(k = 10\), \(\delta = 0.5\), \(\rho = 1\)).

by PCS\textsubscript{Slep} and \(S\). The influence is rather large: For example, to reach a PICS\textsubscript{IZ} of 0.005, \(LL\) requires approximately 115, 125, or 172 samples, depending on stopping rule – the \(S\) stopping rule is much worse than the other two. For \(LL_1\), the ranking is the same for higher acceptable PICS\textsubscript{IZ}, but the influence is smaller. For the above example of a PICS\textsubscript{IZ} of 0.005, the required numbers of samples are approximately 162, 190, and 195 for the three stopping rules. For very low values of PICS\textsubscript{IZ}, the stopping rule with a fixed budget becomes even better than the PCS\textsubscript{Slep} stopping rule (rightmost two lines); the corresponding efficiency line shows less curvature.

Figure 8 shows the influence of the number of alternatives, \(k\), on the relative performance of \(LL(\text{EOC}_{\text{Bonf}})\) and \(LL_1(\text{EOC}_{\text{Bonf}})\). For larger \(k\) (e.g., \(k = 20\) in the figure), the gap between \(LL(\text{EOC}_{\text{Bonf}})\) and \(LL_1(\text{EOC}_{\text{Bonf}})\) is larger when a low level of PICS\textsubscript{IZ} is desired (e.g., \(LL_1(\text{EOC}_{\text{Bonf}})\) requires approximately 67% more samples than \(LL(\text{EOC}_{\text{Bonf}})\) to reach a PICS\textsubscript{IZ} of 0.003). The relative performance of the small-sample procedure improves with decreasing \(k\), until it is basically equivalent for \(k = 2\).

Figure 8 focuses on the EOC\textsubscript{Bonf} stopping rule. For the \(S\) stopping rule (not shown), all procedures become less efficient, as shown previously. For \(k = 2\), \(LL(S)\) and \(LL_1(S)\) perform very similarly. For \(k = 20\), \(LL_1(S)\) requires 32% more samples than \(LL(S)\) to reach a PICS\textsubscript{IZ} level of 0.003, compared to 67% in combination with EOC\textsubscript{Bonf}. The \(LL_1\) allocation suffers somewhat less, relative to \(LL\), with increasing \(k\) when the \(S\) stopping rule is used, as compared to the EOC\textsubscript{Bonf} stopping rule.

We varied \(\rho\) to see if the qualitative observations above would change if the sample variance of the alternatives were different. We did not observe any perceptible trends that deviated from the qualitative observations above.
Figure 7  Influence of the stopping rule on the $LL$ and $LL_1$ allocations (MDM, $k = 10$, $\delta = 0.5$, $\rho = 1$).

B.2. Slippage configuration

In a slippage configuration (SC) the means of all alternatives except the best are tied for second best. We use the parameters $\delta$ and $\rho$ to describe the configurations of the independent outputs with $\text{Normal}(w_i, \sigma_i^2)$ distribution,

$$X_{1j} \sim \text{Normal}(0, \sigma_1^2)$$
$$X_{ij} \sim \text{Normal}(-\delta, \sigma_1^2/\rho) \quad \text{for } i = 2, \ldots, k$$

All alternatives have the same variance if $\rho = 1$. The best alternative has the largest variance if $\rho > 1$. We set $\sigma_1^2 = 2\rho/(1 + \rho)$ so that $\text{Var}[X_{1j} - X_{ij}]$ is constant for all $\rho > 0$. We tested values of $\delta = 0.25, 0.5, 1.0$ and $2.0$ with values of $\rho = 0.707, 1$ and $1.414$ (related to $\sqrt{2}$).

The MDM example above with $k = 2$ is actually also a SC. More SC results with $k = 5$ are shown in Figure 2. In combination with the EOC\textsubscript{Bonf} stopping rule, $LL$ is again more efficient than $LL_1$ (when $k > 2$). In contrast to the MDM tests, $LL_1(S)$ outperforms $LL(S)$ for $k > 2$ in each of our tests (up to $k = 20$).

As for the MDM experiments, changing $\rho$ did not have any discernible effect on the qualitative observations above. Also similar to the MDM experiments, the $LL_1$ allocation suffers somewhat less, relative to $LL$, with increasing $k$ when the $S$ stopping rule is used, as compared to the EOC\textsubscript{Bonf} stopping rule. See Figure 9 for comparison of a typical magnitude of the effect for the $S$ stopping rule, as compared with a typical effect for the EOC\textsubscript{Bonf} stopping rule in Figure 8 above.

B.3. Random problem instances

Random problem instances (RPI) are more realistic in the sense that problems faced in practice typically are not the SC or MDM configuration. The RPI experiment here samples configurations $\chi$ from the normal-inverse gamma family. If $S \sim \text{InvGamma}(\alpha, \beta)$, then $E[S] = \beta/(\alpha - 1)$ and $S^{-1} \sim \text{Gamma}(\alpha, \beta)$ with $E[S^{-1}] = \frac{\beta}{\alpha - 1}$. 

In this experiment, we used $\alpha = 10$ and $\beta = 10$ for all $S$. We tested a range of $\delta = 0, 0.5, 1.0, 1.5, 2.0$ with $\rho = 0.707, 1, 1.414$. We also tested $S^{-1}$ and $\chi$ with $\alpha = 10$ and $\beta = 10$. We used the MDM stopping rule with $k = 5$.
Figure 8  Influence of the number of alternatives, $k$, on $LL(EOC_{Bonf})$ and $LL_1(EOC_{Bonf})$ (MDM, $\delta = 0.5$, $\rho = 1$).

Figure 9  Influence of the number of alternatives, $k$, on $LL(S)$ and $LL_1(S)$ (SC, $\delta = 0.5$, $\rho = 1$).

$\alpha^{-1}$ and $\text{Var}[S^{-1}] = \alpha^{-2}$. A random $\chi$ was generated by sampling the $\sigma_i^2$ independently, then sampling the $W_i$ conditionally independent, given $\sigma_i^2$,

$$p(\sigma_i^2) \sim \text{InvGamma}(\alpha, \beta)$$

$$p(W_i | \sigma_i^2) \sim \text{Normal}(\mu_0, \sigma_i^2/\eta).$$

Increasing $\eta$ makes the means more similar. We set $\beta = \alpha - 1 > 0$ to standardize the mean of the variances to be 1. Increasing $\alpha$ reduces the variability in the variances. We tested values of $\alpha = 2.5$ and 100 with values
of $\eta = 0.5$ and 1. The noninformative prior distributions that can be used in the derivation of EVI procedures correspond to $\eta \to 0$, so there is a mismatch in the sampling distribution of $\chi$ and the prior distributions assumed by the EVI.

A typical result for the RPI problem instances is shown in Figure 10. It is consistent with the previous observations – the asymptotic variants are better in combination with the EOC\textsubscript{Bonf} stopping rule, while the small-sample procedures are at least competitive in combination with the $S$ stopping rule or for a small number of additional samples.

It is interesting to compare the true frequentist performance to the desired target performance when adaptive stopping rules are used. Because an adaptive stopping criterion depends on the samples drawn, the obtained (frequentist) performance is also influenced by the sampling rule. A typical plot is shown in Figure 11. The new procedure leads to a slightly more conservative behavior than the asymptotic procedure, meaning that it over-delivers slightly more.

**B.4. Hybrid procedures**

The preceding results show that the small sample procedures perform better when only a small number of samples is allocated, while the asymptotic procedures perform better for a larger number of samples that are sequentially allocated one after another. These results are consistent with intuition, given the derivation of the procedures. The asymptotic procedures allocate in order to solve a large-sample asymptotic approximation to the expected value of information, while the small sample procedures greedily maximize the outcome assuming only one additional sample.

A natural question is whether it is possible to combine the procedures. In particular, if a large number of samples is allocated, is it better to first allocate samples with the asymptotic procedure, then switch to the small sample procedure to allocate the last few remaining samples.
Figure 11  Frequentist \( \text{EOC}_{IZ} \) depends on target EOC \( \beta^* \) (RPI, \( k = 5 \), \( \eta = 1 \), \( \alpha = 100 \)).

Figure 3 and Figure 12 show the effect of switching from one allocation to another. The procedures in those figures allocate one sample at a time. The notation \( LL \) denotes the usual \( LL(S) \) procedure. The notation \( LL_1(S < 200) \) denotes the use of \( LL_1 \) to allocate until 200 samples have been observed, and \( LL \) thereafter. In Figure 3 the results obtained with \( LL \) are significantly better than the results obtained with \( LL_1 \). Switching from \( LL \) to \( LL_1 \) after 200 samples (lower pair of curves) is beneficial for at least several additional samples (approximately 20 in this case). Switching from \( LL_1 \) to \( LL \) after 200 samples (upper pair of curves) temporarily decreases performance. Continuing with \( LL \) for a longer term (not shown here) returns \( LL \) to a state that dominates \( LL_1 \). The plot in the first panel of Figure 3 is somewhat more jagged than the other plots, due to the scaling of the ordinate and the abscissa.

Similar observations hold when switching at the beginning of the run, at a time when \( LL_1 \) is still better than \( LL \), as in Figure 12.

Overall, this suggests that there is a small benefit to allocate for the bulk of the stages with \( LL \), and to switch to \( LL_1 \) for the last few samples. It does not appear to be effective to start with \( LL_1 \) then later switch to \( LL \), hoping that the initially better performance could be sustained by \( LL \) in the long run.

**Appendix C: Small-Sample EVI Procedures for Known Variances**

Straightforward algebra that follows the logic of the derivation in Section 2 below shows that the analog of (17) for the case of a known variances \( \sigma_i \) for \( i = 1, 2, \ldots, k \) is

\[
\text{EVI}^{\text{known } \sigma_i}_{LL,(i)} = \begin{cases} 
\lambda_{(i,k)}^{-1/2} \Psi \left[ d_{(i,k)}^{(i,k)} \right] & \text{if } (i) \neq (k) \\
\lambda_{(k-1,k)}^{-1/2} \Psi \left[ d_{(k-1,k)}^{(i,k)} \right] & \text{if } (i) = (k),
\end{cases}
\]

(20)

if \( \sigma_j^2 \) replaces \( \hat{\sigma}_j^2 \) in the definition of \( \lambda_{(i,k)} \) for \( j \in \{i, k\} \), and where \( \phi(s), \Phi(s) \) and \( \Psi[s] = \phi(s) - s(1 - \Phi(s)) \) are the pdf, cdf and standard EOC loss functions, respectively, for the standard normal distribution.
Figure 12 Effect of switching from $LL$ to $LL_1$ or vice versa, after 50 samples (RPI, $k = 5$, $\eta = 0.5$, $\alpha = 100$).

Similarly, if $\sigma_{(j)}^2$ replaces $\hat{\sigma}_{(j)}^2$ in the definition of $\lambda_{(i,k)}$ for $j \in \{i,k\}$ then

$$
\text{EVI}_{0-1,(i)}^{\text{known } \sigma_i} = \begin{cases} 
\Phi(-d_{(ik)}') & \text{if } (i) \neq (k) \\
\Phi(-d_{(k-1,k)}') & \text{if } (i) = (k)
\end{cases}
$$

(21)

is the expected value of information approximation for the 0-1 loss function when the variances are all known.

One can readily verify that $\Psi[-s] = s + \Psi[s]$ and $\Psi_\nu[-s] = s + \Psi_\nu[s]$. The expected reward if samples are allocated only to alternative $(i)$ and to no other alternative, for the known and unknown variance cases, is therefore

$$
\begin{align*}
\text{E}[\max_j W_j | \text{known } \sigma_i] &= \bar{x}_{(k)} + \text{EVI}_{LL,(i)}^{\text{known } \sigma_i} \\
\text{E}[\max_j W_j | \text{unknown } \sigma_i] &= \bar{x}_{(k)} + \text{EVI}_{LL,(i)}.
\end{align*}
$$

(22)

Intuitively, (22) means that the expected reward of a one-step look ahead procedure equals the expected reward if one were to stop now plus the expected value of information in the one-step of sampling, as measured by the expected opportunity cost. This generalizes the observation of the greedy one-step look ahead algorithm of Gupta and Miescke (1996, Lemma 4), which assumes a known common variance for each alternative. Their $\sigma_i$ is our $\frac{\tau_{(i)}^2}{n_{(i)}(n_{(i)} + \tau_{(i)})}$.

We can assess the optimal one-step allocation of $\tau$ samples to one alternative when $k = 2$ and the variances are known. Straightforward algebra shows that the maximizer of $\text{EVI}_{LL,(i)}^{\text{known } \sigma_i}$ is the maximizer of

$$
\frac{\sigma_{(i)}^2 \tau}{n_{(i)}(n_{(i)} + \tau)}.
$$

(23)

If we formally let $\tau \downarrow 0$, this allocation becomes the allocation that adds samples to minimize the difference between $(n_{(1)} + \tau_{(1)})/\sigma_{(1)}$ and $(n_{(2)} + \tau_{(2)})/\sigma_{(2)}$. To see this, take the derivative of (23) with respect to $\tau$, send $\tau$ to 0, and select the alternative that maximizes the result. The implication is that samples should be added so that the $\sigma_i/n_i$ for the two alternatives stay as equal as possible (subject to the integer sampling constraint). The same result, interestingly, is obtained for the 0-1 loss function for $k = 2$ alternatives with
known variances as well. In both cases, with \( k = 2 \) and equal variances, the optimal allocation samples equally often from both alternatives, for both loss functions that are considered above.

This last observation, for the case of \( k = 2 \) alternatives with known variances, matches the optimal allocation for this special case as formulated in Gupta and Miescke (1994, comments after Theorems 1 and 2) for the EVI approach, and in Chen et al. (2000) Remark 1, page 259, which handles the case of unknown variances by assuming a known variance approximation) for the OCBA approach. This greedy one-step procedure is optimal for a one-stage procedure but may be suboptimal when applied sequentially. Gupta and Miescke (1996) proposed using such a one-stage lookahead, but left the case of unknown variances for further research. The present work handles the case of unknown variances.

### Appendix D: Implementation Issues

The implementation that generated the analysis and graphs in this paper was written in C++. It used the Gnu Scientific Library (gsl) for calculating cdfs, the Mersenne twister random number generator (Matsumoto and Nishimura 1998, with 2002 revised seeding), and FILIB++ (Lerch et al. 2001) for interval arithmetic.

This section describes some numerical analysis techniques that were used in order to improve the performance of some of the procedures in this paper. Some of what follows appeared in the online companion of Branke et al. (2007), which discussed numerical stability techniques for the OCBA allocations and certain stopping rules that are used by EVI and OCBA procedures. Some of those techniques are required to improve the stability of the new EVI allocations. Some additional numerical stability issues are also raised here.

In order to calculate PCS\(_{\text{Step}}\) with high numerical accuracy, we did not calculate it directly with the formula in Section 1.2. We calculated PCS\(_{\text{Step}}\) indirectly via 

\[ 1 - \text{PCS}_{\text{Step}} = -(\exp(\sum_{j} \log(1 - \Phi_{n}(u_{(j)}(-d_{jk})))) - 1). \]

This transformation is useful because \( \log(1 - x) = \log(1 + (-x)) \) and \( \exp(x) - 1 = \expm1(x) \) from the C runtime library have increased accuracy for \( x \) near 0, and \( \Phi_{n}(t) \) is more stable for \( t \to -\infty \).

Numerical stability problems may arise in implementations of the \( LL_{1} \) and 0-1 allocations, even with double-precision floating point arithmetic, as the total number of samples gets quite large (i.e., for low values of \( a^{*} \) or EOC bounds). To better distinguish (i) which alternative should receive samples in a given stage and (ii) when to stop sampling, we improved numerical stability in several ways for \( LL_{1} \) and 0-1.

We evaluate the alternative that maximizes log EVI. If the numerical stability does not suffice to calculate \( \log \Phi_{n}(t) \) and \( \log \Psi_{n}(t) \) (underflow error), we considered two general approaches. In the first approach, we use interval arithmetic based on bounds for \( \log \Phi_{n}(t) \) and \( \log \Psi_{n}(t) \) that are based on the following property of the cdf of a \( t \)-distribution (Evans et al. 1993),

\[
\Phi_{n}(t) = \begin{cases} 
\frac{1}{2} \text{beta}\left(\frac{\nu}{2}, \frac{1 - t}{2}\right) & \text{if } t \leq 0, \\
1 - \frac{1}{2} \text{beta}\left(\frac{\nu}{2}, \frac{1 - t}{2}\right) & \text{if } t > 0,
\end{cases}
\]

where \( \beta\left(a, b, x\right) = \beta(a, b)^{-1} \int_{0}^{x} u^{a-1}(1 - u)^{b-1} du \) is the incomplete beta function, and \( \beta(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b) \) denotes the beta function. A lower bound for \( \log \Phi_{n}(-t) \) for \( t > 0 \) can be derived as follows.

If \( f(u) = u^{-1}(1 - u)^{-1}, \) then \( f(0) = 0, f'(u) \geq 0 \) and \( f''(u) > 0 \) for \( a = \frac{\nu}{2} > 1, b = \frac{1}{2} \) and all \( u \in [0, 1]. \) So the area below \( f(u) \) over \([0, x]\) is always larger than the area below the tangent at \((x, f(x)).\)

\[
\log(\Phi_{n}(-t)) \geq \frac{x}{2} \log\frac{\nu}{\nu + x} + \frac{1}{2} \log(1 - \frac{\nu}{\nu + x}) - \log\left(\frac{(\nu - 1)(1 - \frac{\nu}{\nu + x})}{2}\right) - \log 2
\]

(25)
For the upper bound, recall (1). As \( \Psi_\nu(t) > 0 \) for all \( t > 0 \) we obtain \( \Phi_\nu(-t) < \frac{1}{t} \nu \frac{t^\nu}{\nu-1} \phi_\nu(t) \), so
\[
\log \Phi_\nu(-t) < \log \frac{1}{t} \nu \frac{t^\nu}{\nu-1} + \log \phi_\nu(t).
\]

The bounds in (25) and (26) help for \( \Phi_\nu \) but cannot directly be used for bounds on \( \Psi_\nu(t) \). But the numerical stability can be increased by bringing the log inside the calculation of \( \log \Psi_\nu(t) \):
\[
\log \Psi_\nu(t) = \log \frac{t^\nu}{\nu-1} + \log \phi_\nu(t) + \log \left[ 1 - \frac{t}{(\nu-1)} \exp(\log \Phi_\nu(-t) - \log \phi_\nu(t)) \right]
\]
Equation (26) can be used to determine an upper bound for \( \log \Psi_\nu(t) \).

Note that \( \log \phi_\nu(t) \) can be calculated by using the logarithm of the Gamma-function,
\[
\log \phi_\nu(t) = \log \Gamma \left( \frac{\nu+1}{2} \right) - \log \Gamma \left( \frac{\nu}{2} \right) - \frac{\nu}{2} \log(\nu\pi) - \frac{\nu+1}{2} \log \left( 1 + \frac{t^2}{\nu} \right).
\]

Collisions, due to the EVI estimates not having numerically unique interval bounds, occurred somewhat regularly for the new small-sample procedures. If there was no clearly defined best and the EVI was not numerically different from 0 for any alternative (with interval arithmetic), then we repeatedly doubled \( \tau \) for purposes of calculating the EVI, until at least one alternative was numerically greater than 0. The “winner” then received 1 sample. Usually, doubling at most 3 times \( (\tau = 8) \) was sufficient to select a winner. If there was no clearly defined best because two or more alternatives whose EVI had overlapping intervals but the intervals did not contain 0, then we allocated \( \tau = 1 \) sample to the alternative with the highest upper bound for the interval.

The second way to address numerical stability in calculating \( \log \text{EVI} \) was to simply use the above upper bounds. Since using the complex interval arithmetic increased CPU time by 50%, while not leading to significant enhancements, we ran the experiments reported in this paper with the simple upper bound variant.

The slight bend to the right that can be observed in efficiency plots for low values of \( \alpha^* \) or EOC bounds may suggest a potential inefficiency due to another numerical issue that we have not yet identified. Increasing \( n_0 \) can help to reduce the bending for low values of \( \alpha^* \) or EOC.

We also attempted other numerical methods to improve stability. In particular, the EVI for Procedure 0-1, can be determined by integrating (7) for \( k > 2 \), or with (8) for \( k = 2 \). We implemented both naive Monte Carlo and importance sampling methods to estimate the EVI in (8) but found the estimators to be too inefficient and noisy to be useful. An alternative with \( k = 2 \) is one-dimensional quadrature for
\[
\text{EVI}_{0-1,(t)} = E \left[ \mathbf{1} \{ U > 0 \} \left( 2 \Phi_\nu(\lambda^{1/2}_{ik} U) - 1 \right) \right]
= \int_0^\infty \phi_{\nu(t)(k)} \left( y + \lambda^{1/2}_{ik} d_{(i)(k)} \right) \left( 2 \Phi_\nu(\lambda^{1/2}_{ik} y^{1/2} \lambda^{1/2}_{ik}) - 1 \right) dy.
\]

Quadrature behaved ‘better’ when transformed to the unit interval, \( \int_0^\infty f(y)dy = \int_0^1 f((1-t)/t^2 dt, \) but the CPU time to compute the allocation was an order of magnitude slower than for 0-1, and accuracy was still poor if the differences in means was large.

In summary, we tested a number of techniques in order to improve the numerical stability of each procedure. Some, but not all, of those numerical techniques required additional computational overhead in order to
compute the allocation. Monte Carlo estimates and complicated quadrature, the techniques that required the most additional overhead, did not necessarily help more. Interval arithmetic worked fine, but again the minimal improvements did not seem to justify the computational overhead. Definitely useful were the transformations that allowed us to use functions from the C runtime library with higher precision in the critical value space.

The $LL$ and 0-1 sampling allocations did not experience numerical stability problems with collisions. Those allocations did not require the above numerical stability techniques.

Any procedure that uses the $PCS_{Slep}$ stopping rule can benefit from the technique to improve the accuracy of estimating $PCS_{Slep}$ with high numerical accuracy.

References

Beyond references in the main paper, the online companion also refers to the following works.


