*2021*年冬季仿真会议论文集 *S. Kim*、*B. Feng*、*K. Smith*、*S. Masoud*、*Z. Zheng*、*C. Szabo* 和 *M. Loper* 编

在排序与选择中实现“速率最优”

哈伦·阿夫吉
Barry L. Nelson
Andreas W¨achter

西北大学工业工程与管理科学系，美国伊利诺伊州埃文斯顿，邮编 60208

arXiv:2302.02254v1 [stat.CO] 4 Feb 2023

摘要

在 2004 年的开创性论文中，Glynn 和 Juneja 正式且精确地确立了用于从k 个仿真系统中选出最佳系统的速率最优错误选择概率重复仿真分配方案。对于独立正态分布输出，这种分配方案具有简单的形式，并以直观易懂的方式取决于真实均值和方差。当然，均值和（通常也包括）方差都是未知的，但速率最优分配为可实施的动态数据驱动策略提供了一个目标。本文比较了四种相关重复仿真分配策略的实证表现：Chen 和 Rzyhov 提出的 mCEI，以及我们新提出的 gCEI 策略，两者均收敛于 Glynn 和 Juneja 的分配方案；Peng 和 Fu 提出的 AOMAP，收敛于 OCBA 最优分配；以及 Russo 提出的 TTTS，其目标是错误选择后验概率的收敛速率。我们发现，这些策略在某些情形下表现出明显不同的行为。

1 引言

排序与选择（R&S）是解决随机仿真优化问题的基本方法之一。在排序与选择问题的经典版本中，目标是从有限数量（k）的系统中找出唯一的最优系统；每个系统的性能只能通过仿真输出进行估计。这里的“最优”是指性能期望值最大或最小。理想的排序与选择程序要么（a）分配有限的仿真预算，以最大限度地提高选出最优系统的可能性；要么（b）尽可能高效地分配仿真资源，直到达到预先设定的可能性。本文讨论的是表述（a）。

排序与选择文献中包含许多针对版本（a）的策略，这些策略通过从各系统中依次获取重复仿真，并随着输出数据不断增多而进行调整。这些策略通常采用贝叶斯方法或受贝叶斯方法启发，包括最优计算预算分配（OCBA，Chen 等人 2000年）、期望改进（EI，Jones 等人 1998年）、知识梯度（KG，Frazier 和 Powell 2007年）以及多臂老虎机（MAB，Jamieson 和 Nowak 2014年）的不同版本。本文中，分配指将固定的重复仿真预算分配给各个仿真系统的比例，而策略则是依次（通常以自适应方式）将单次重复仿真分配给各系统的一种算法。

Glynn 和 Juneja（2004）利用大偏差理论推导出了渐近最优的静态重复仿真分配方案。他们将分配给系统i 的重复次数表示为 $\alpha_i R$，其中R 是重复仿真的总预算，$\alpha_i > 0$，且 $\sum_{i=1}^{k} \alpha_i = 1$。该策略在这样的意义上是“最优”的：随着$R$增大，错误选择概率（PICS）以可能达到的最快指数速率衰减；错误选择是指选中$k-1$ 个劣势系统中的任意一个。不幸的是，最优分配取决于底层输出分布及其参数，而这些通常是未知的；简单的代入策略往往效果不佳。因此，人们提出了自适应策略的想法，

在样本量较小时积极追求最佳系统，并在极限情况下收敛到“最优”分配方案，例如 Glynn 和 Juneja（2004）提出的方案。

近年来，已有若干此类策略被提出。在某些条件下，Chen 和 Ryzhov（2019b）的修正完全期望改进（mCEI）策略的经验分配收敛于 Glynn 和 Juneja（2004）的速率最优分配。Peng 和 Fu（2017）的渐近最优短视分配策略（AOMAP）收敛于 OCBA 极限分配。最后，Russo（2020）的前二汤普森采样（TTTS）策略旨在使后验错误选择概率以最优速率收敛于 0；TTTS 是 Russo（2020）提出的三个“前二”策略之一。请注意，AOMAP 和 TTTS 并不 收敛于 Glynn 和 Juneja（2004）的最优分配，而是收敛于一些可以说更为理想的极限。

在此基础上，我们加入了新的*CEI* 梯度（gCEI）策略，该策略达到与 mCEI 相同的极限。随后，在支持这四种策略的假设下，我们对这四种策略的固定预算行为进行了经验评估：输出相互独立且服从正态分布，方差已知。与渐近行为相反，固定预算行为才是分析人员在实践中实际经历的情况。需要指出的是，在某种意义上，mCEI、gCEI、AOMAP 和 TTTS 都可能被专门设计来实现良好有限样本性能的策略“击败”，尤其是在系统数量k 非常大时；此类策略能够迅速完全排除明显的劣势系统，而 mCEI、gCEI、AOMAP 和 TTTS 则会让所有系统一直保留在竞争中，直到预算耗尽。尽管如此，它们仍是更复杂策略的构建模块，因此这种比较具有相关性。

本文其余部分安排如下：第 2 节简要回顾相关文献，第 3 节对排序与选择问题进行形式化表述。第 4 节介绍 AOMAP、mCEI 和 TTTS 策略，随后在第 5 节引入 gCEI 并概述其收敛性的证明。第 6 节给出这四种策略的经验性能。最后，第 7 节总结全文。

2 重要文献

期望改进准则最初由 Jones 等人（1998年）提出，用于确定性仿真的贝叶斯优化。Ryzhov（2016）将期望改进准则应用于具有独立正态观测的排序与选择问题，推导出该准则所隐含的渐近采样分配，并指出它与 Chen 等人（2000年）的最优计算预算分配相联系。由于期望改进无法达到指数收敛速率，Peng 和 Fu（2017）提出了一种称为 AOMAP 的期望改进变体。在已知方差情形下，AOMAP 的极限分配为 Chen 等人（2000年）提出的最优计算预算分配。Peng 和 Fu（2017）指出，可以对期望改进进行调整，以实现任意定义良好的极限分配，包括 Glynn 和 Juneja（2004）的分配；但这需要在每次迭代中基于代入估计求解极限分配。

由于 EI 最初是为确定性仿真创建的，因此无法直接考虑随机仿真输出中的不确定性。为纳入这种不确定性，Salemi 等人（2019）提出了完全期望改进（CEI）。对于具有独立正态观测的排序与选择问题，Chen 和 Ryzhov（2019b）提出了修正的完全期望改进策略。类似地，在更一般的抽样分布下，Chen 和 Ryzhov（2019a）提出了平衡最优大偏差策略，该策略评估近似的个体大偏差率函数，并对其进行迭代平衡。当方差已知，或方差未知但通过代入估计量持续更新时，这两种策略均能渐近地实现 Glynn 和 Juneja（2004）的最优分配。

最近，Russo（2020）提出了三种不同的贝叶斯策略，用于在包括仿真在内的随机决策问题中自适应地分配测量资源。在每次迭代中，这些策略都会利用输出参数（例如均值）的后验分布来确定前两名备选方案；然后随机选取其中一个进行测量（仿真）。选择概率是一个调节参数，不过 Russo（2020）发现 $1/2$ 具有稳健的经验性能。前两名概率抽样会找出后验最优概率最大的两个备选方案。类似地，前两名价值抽样会考虑每个系统的均值与其他系统中的最佳者之差的后验期望值。第三种方法是 TTTS；有关 Thompson 抽样的起源，请参见 Thompson（1933）。本文

采用选择概率为 $1/2$ 的 TTTS，并将在下文对此作完整介绍。对于已知方差的情形，Russo（2020）证明，当调节参数设为最优值或自适应地调整至最优值时，这些策略能够使真实最佳系统的后验错误选择概率以最佳指数速率收敛。

除了 AOMAP、mCEI 和 TTTS 之外，我们还提出了一种名为 gCEI 的新策略。该策略根据 CEI 关于从每个系统获得的重复次数的梯度来做出重复仿真分配决策，并将“重复次数”视为连续值。与 mCEI 一样，它实现了 Glynn 和 Juneja（2004）的最优分配。然而，gCEI 无需像 mCEI 那样直接强制平衡对最佳系统和劣势系统的仿真，即可达到这一极限。

3 预备知识

令 $\mathscr{S} = \{1,2,\ldots,k\}$ 表示系统集合。每个系统 $i \in \mathscr{S}$ 的均值 μ_i 未知。数值越大越好，而 $\mu_1 \le \mu_2 \le \ldots \le \mu_{k-1} < \mu_k$ 对我们来说是未知的。我们的目标是找到唯一的最佳系统 k。

从贝叶斯视角来看，每个系统i的均值都有一个先验分布 $\mu_i \sim \mathrm{N}(\bar{\mu}_i(0)\ 1/\theta_i(0))$，其中 $\bar{\mu}_i(0)$ 和 $\theta_i(0)$ 分别是先验均值和先验精度。先验均值 $\bar{\mu}_i(0)$ 表示对 μ_i 真实值的初始信念，而先验精度 $\theta_i(0)$ 则量化了对这一信念的信心。我们假设在该先验下，μ_i彼此独立。请注意，我们用 μ_i 表示系统的真实固定均值，用 $\bar{\mu}_i(t)$ 表示经过t次迭代后的后验均值，它是一个随机变量。

我们考虑一个具有固定仿真预算的有限时域问题：设R为有限时域的长度。在每次迭代$t = 0, 1,\ldots,R$时，我们通过仿真$x(t)$，获得一次重复仿真$Y(t+1)$；该仿真是一个独立同分布的 $\mathrm{N}(\mu_{x(t)}, \sigma^2_{x(t)})$ 随机变量，其中 $\sigma^2{}_i > 0$ 是系统 i 的随机仿真输出所固有的方差。本文假设已知 σ_i^2，且各系统的仿真彼此独立（不使用共同随机数）。

令 $\mathscr{F}^t$为由 $\{x(\tau)、Y(\tau+1)\}_{\tau=0}^{t-1}$生成的σ代数。根据 De Groot（1970）中的递归方法，系统 $i \in \mathscr{S}$在第 t </style}次迭代时的后验参数为

$$\bar{\mu}_i(t+1) = \begin{cases} \dfrac{\bar{\mu}_i(t)\theta_i(t) + Y(t+1)/\sigma_i^2}{\theta_i(t) + 1/\sigma_i^2} & \text{if } x(t) = i \text{ (i.e., if system } i \text{ is simulated at iteration } t) \\ \bar{\mu}_i(t) & \text{if } x(t) \neq i, \end{cases}$$

$$\theta_i(t+1) = \begin{cases} \theta_i(t) + 1/\sigma_i^2 & \text{if } x(t) = i \\ \theta_i(t) & \text{if } x(t) \neq i. \end{cases}$$

令 $r_i(t)$ 表示通过模拟系统 i 截至第 t次迭代所获得的重复次数，即 $r_i(t) = \sum_{\tau=0}^{t-1} \mathbb{I}_{\{x(\tau)=i\}}$ 其中 $\mathbb{I}_{\{\cdot\}}$ 为指示函数。我们采用非信息先验（即 $\theta_i(0) = 0$）。因此，我们有 $\bar{\mu}_i(t) = \bar{Y}_i(t)$，以及 $\theta_i(t) = r_i(t)/\sigma_i^2$其中

$$\bar{Y}_i(t) = \frac{1}{r_i(t)} \sum_{\tau=0}^{t-1} \mathbb{I}_{\{x(\tau)=i\}} Y(\tau+1)$$

是系统 i的样本均值。

令 $k(t)$ 表示第 t次迭代时的样本最优系统，$k(t) = \mathrm{argmax}_{i\in\mathscr{S}}\{\bar{\mu}_i(t)\} = \mathrm{argmax}_{i\in\mathscr{S}}\{\bar{Y}_i(t)\}$。我们将第 t 次迭代时的（频率学派）正确选择概率（PCS）定义为 $\mathbb{P}\{k(t) = k\}$；因此，错误选择概率（PICS）为 $\mathbb{P}\{k(t) \neq k\}$。这些量均相对于固定的真实均值而言。我们还可以定义系统 i 为或不为最佳系统的后验概率对应量，这与 TTTS 相关。

算法 1 给出了通用自适应策略。AOMAP、mCEI、gCEI 和 TTTS 在第 3 步中如何确定$x(t)$方面各不相同，以获得良好的有限R 和渐近$R \to \infty$ 性能。

算法 1 通用自适应策略

1：令$x(0) = i$，其中$i \in \mathscr{S}$。获取$Y(1)$并更新 $\mathscr{F}^1$。此外，令$t \leftarrow 1$。2：**当** $t < R$ **时执行**
3：决定模拟$x(t)$。4：通过模拟$x(t)$获取$Y(t+1)$，更新 $\mathscr{F}^{t+1} \leftarrow \mathscr{F}^t \cup \{x(t),\ Y(t+1)\}$ 以及
$t \leftarrow t+1$。5：**结束循环**6：返回$k(R) = \text{argmax}_{i\in\mathscr{S}}\{\bar{\mu}_i(R)\}$ 作为选出的最佳系统。

4 策略

本节总结近期文献中的三种不同策略。

4.1 AOMAP

对于我们的排序与选择问题，系统 i 在t 次迭代时的期望改进为

$$\text{EI}_i(t) = \mathbb{E}\left[\max\{\mu_i - \bar{\mu}_{k(t)}(t), 0\}|\mathscr{F}^t\right] = \sqrt{1/\theta_i(t)}\ f\left(\frac{\bar{\mu}_i(t) - \bar{\mu}_{k(t)}(t)}{\sqrt{1/\theta_i(t)}}\right)$$

其中，$f(z) = z\Phi(z) + \phi(z)$，其中 ϕ和 Φ分别为标准正态概率密度函数和累积分布函数。Ryzhov（ 2016）表明，当分配给劣势系统的预算趋于零时，期望改进并不能精确实现 OCBA 分配。在贝叶斯框架下，Peng 和 Fu（2017）提出了一种称为 AOMAP 的短视分配策略，作为期望改进的一种新变体，并证明当方差已知时，AOMAP 确实能够实现 OCBA 分配。采用此策略后，步骤 3 变为

$$x(t) = \text{argmax}_{i\in\mathscr{S}}\{\mathbb{E}\left[\max\{\mu_i - A_i(t), 0\}|\mathscr{F}^t\right]\} = \text{argmax}_{i\in\mathscr{S}}\left\{\sqrt{1/\theta_i(t)}\ f\left(\frac{\bar{\mu}_i(t) - A_i(t)}{\sqrt{1/\theta_i(t)}}\right)\right\}$$

其中，$A_i(t) = \bar{\mu}_{k(t)}(t)\,\mathbb{I}_{\{\bar{\mu}_i(t)\neq\bar{\mu}_{k(t)}(t)\}} + (\bar{\mu}_{k(t)}(t) + \xi_{k(t)}(t)\sigma_{k(t)})\,\mathbb{I}_{\{\bar{\mu}_i(t)=\bar{\mu}_{k(t)}(t)\}}$，且

$$\xi_{k(t)}(t) = \left(\sum_{i\in\mathscr{S}\setminus\{k(t)\}} \frac{\sigma_{k(t)}^2\sigma_i^2}{[\bar{\mu}_i(t) - \bar{\mu}_{k(t)}(t)]^4}\right)^{-1/4}.$$

注意，如果 $\xi_{k(t)}(t) = 0$，那么 $A_i(t) = \bar{\mu}_{k(t)}(t)$，因此期望变为 $\text{EI}_i(t)$。由于期望改进在分配给最佳系统时过于贪婪，因此附加项会调整分配给最佳系统的预算，使其随着迭代次数趋于无穷而不那么有利。这种调整使 AOMAP 能够实现 OCBA 极限分配。

4.2 mCEI 策略

由于期望改进无法完全捕捉随机仿真输出中的不确定性，Salemi 等人（2019）在高斯马尔可夫随机场框架下，针对离散仿真优化提出了 CEI。对于我们的排序与选择问题，迭代 t 时系统 $i \neq k(t)$ 的 CEI 为

$$\text{CEI}_i(t) = \mathbb{E}\left[\max\{\mu_i - \mu_{k(t)}, 0\}|\mathscr{F}^t\right] = \sqrt{1/\theta_i(t) + 1/\theta_{k(t)}(t)}\ f\left(\frac{\bar{\mu}_i(t) - \bar{\mu}_{k(t)}(t)}{\sqrt{1/\theta_i(t) + 1/\theta_{k(t)}(t)}}\right).$$

Chen 和 Ryzhov（2019b）提出了用于排序与选择的修改的条件期望改进策略 mCEI，这是对 Salemi 等人（2019）原始 CEI 策略的修改版本。在 mCEI 下，$x(t) = k(t)$ 当

$$\left(\frac{r_{k(t)}(t)}{\sigma_{k(t)}}\right)^2 < \sum_{i\in\mathscr{S}\setminus\{k(t)\}}\left(\frac{r_i(t)}{\sigma_i}\right)^2. \tag{1}$$

否则，$x(t) = c(t)$，其中 $c(t) = \text{argmax}^{i \neq k(t)} \text{CEI}_i(t)$。条件 (1) 在极限情况下保证了 Glynn 和 Juneja（2004）最优分配中的平衡条件。

4.3 TTTS 策略

由于 TTTS 不只是对算法 1 中的步骤 3 进行简单替换，因此我们将新的步骤 3 列为算法 2。通过将 β 调整至最优值，可以使 TTTS 获得渐近最优性能。不过，Russo（2020）通过简单地选择 $\beta = 1/2$，取得了良好的经验性能。

算法 2 TTTS 步骤 3

抽取 $\widehat{\mu}_i \sim \text{N}(\bar{\mu}_i(t)\ 1/\theta_i(t))$，其中 $i \in \mathscr{S}$，并令 $I \leftarrow \text{argmax}_{i \in \mathscr{S}} \widehat{\mu}_i$。 ⊲Thompson 抽样
抽取$B \sim$ 伯努利分布(β)。
如果$B = 1$ **那么**
  $x(t) = I.$
else
  重复
    抽取 $\widehat{\mu}_j \sim \text{N}(\bar{\mu}_j(t)\, 1/\theta_j(t))$，对于$j \in \mathscr{S}$，并设置$J \leftarrow \text{argmax}_{j \in \mathscr{S}} \widehat{\mu}_j$。
  直到$J \neq I$
  $x(t) = J.$
结束 if

5 gCEI策略

在确定性仿真的贝叶斯优化中，EI 已被证明是一种有效的搜索策略；CEI 将 EI 扩展到随机仿真；而 mCEI 则对 CEI 进行调整，使其在排序与选择中获得最优渐近性能，确保在极限情况下，对最佳系统和其余系统进行仿真的必要平衡得以实现；纯条件期望改进策略在下一次迭代中从不仿真当前样本最佳系统。

Salemi 等人（2019）中的 GMIA 等基于条件期望改进的仿真优化方法有一个特点：CEI 能识别有前景的解，但无法确定应在这些解上投入多少次重复仿真。gCEI 源于一项持续进行的研究，该研究旨在利用 CEI 的梯度来决定重复次数，方法是将重复次数视为连续变量。这里，我们仅用它来决定如何分配下一次单次 重复仿真，与其他策略一样。

为推导 CEI 梯度的表达式，首先注意到，f 对z的导数为$f'(z) = \Phi(z)$。然后，对于$i \neq k(t)$，

$$\frac{\partial}{\partial r_i(t)} \left(\frac{1}{\theta_i(t)} \right) = -\frac{\sigma_i^2}{(r_i(t))^2} \text{ and } \frac{\partial \bar{\mu}_i(t)}{\partial r_i(t)} = \frac{\partial \bar{Y}_i(t)}{\partial r_i(t)} = 0.$$

为简化符号，令 $\nu_i = 1/\theta_i(t) + 1/\theta_{k(t)}(t)$。于是我们有

$$\frac{\partial}{\partial r_i(t)} \left(\frac{\bar{\mu}_i(t) - \bar{\mu}_{k(t)}(t)}{\sqrt{\nu_i}} \right) = \frac{\left(\bar{Y}_i(t) - \bar{Y}_{k(t)}(t) \right)}{2\nu_i \sqrt{\nu_i}} \frac{\sigma_i^2}{(r_i(t))^2}.$$

因此，$\text{CEI}_i(t)$ 对 $r_i(t)$ 的导数为

$$\begin{aligned} &\frac{\partial \text{CEI}_i(t)}{\partial r_i(t)} \\ &= -\frac{\sigma_i^2}{2\sqrt{\nu_i}(r_i(t))^2} f\left(\frac{\bar{Y}_i(t) - \bar{Y}_{k(t)}(t)}{\sqrt{\nu_i}} \right) + \sqrt{\nu_i} \left[\frac{\left(\bar{Y}_i(t) - \bar{Y}_{k(t)}(t) \right)}{2\nu_i \sqrt{\nu_i}} \frac{\sigma_i^2}{(r_i(t))^2} \right] f'\left(\frac{\bar{Y}_i(t) - \bar{Y}_{k(t)}(t)}{\sqrt{\nu_i}} \right) \end{aligned}$$

$$
\begin{aligned}
&= -\frac{\sigma_i^2}{2\sqrt{v_i}(r_i(t))^2}\left[\frac{\bar{Y}_i(t)-\bar{Y}_{k(t)}(t)}{\sqrt{v_i}}\,\Phi\left(\frac{\bar{Y}_i(t)-\bar{Y}_{k(t)}(t)}{\sqrt{v_i}}\right)+\phi\left(\frac{\bar{Y}_i(t)-\bar{Y}_{k(t)}(t)}{\sqrt{v_i}}\right)\right.\\
&\quad\left.-\frac{\left(\bar{Y}_i(t)-\bar{Y}_{k(t)}(t)\right)}{\sqrt{v_i}}\,\Phi\left(\frac{\bar{Y}_i(t)-\bar{Y}_{k(t)}(t)}{\sqrt{v_i}}\right)\right]\\
&= -\frac{\sigma_i^2}{(r_i(t))^2}\,\frac{1}{2\sqrt{v_i}}\,\phi\left(\frac{\bar{Y}_i(t)-\bar{Y}_{k(t)}(t)}{\sqrt{v_i}}\right)\le 0.
\end{aligned}
$$

同理,

$$
\frac{\partial \mathrm{CEI}_i(t)}{\partial r_{k(t)}(t)} = -\frac{\sigma_{k(t)}^2}{(r_{k(t)}(t))^2}\,\frac{1}{2\sqrt{v_i}}\,\phi\left(\frac{\bar{Y}_i(t)-\bar{Y}_{k(t)}(t)}{\sqrt{v_i}}\right)\le 0
$$

而 $\partial \mathrm{CEI}_i(t)/\partial r_j(t) = 0$对于 $j \in \mathscr{S}\backslash\{i,\ k(t)\}$因为 $\partial(\theta_i(t))^{-1}/\partial r_j(t) = 0$。

由于在第t 次迭代时模拟样本最佳的$k(t)$会影响所有 CEI，因此模拟$k(t)$的总影响为 $\sum_{i\in\mathscr{S}\backslash\{k(t)\}}\partial$ $\mathrm{CEI}_i(t)/\partial r^{k(t)}(t)$。由于 CEI 值越低越好，因此在除$k(t)$之外的系统中，模拟

$$
g(t) = \mathrm{argmin}_{i\in\mathscr{S}\backslash\{k(t)\}}\frac{\partial \mathrm{CEI}_i(t)}{\partial r_i(t)}
$$

可能带来最大的改进。为了决定接下来模拟哪个系统，$k(t)$还是$g(t)$，我们提出以下条件:

$$
\sum_{i\in\mathscr{S}\backslash\{k(t)\}}\frac{\partial \mathrm{CEI}_i(t)}{\partial r_{k(t)}(t)} \overset{?}{\le} \min_{i\in\mathscr{S}\backslash\{k(t)\}}\frac{\partial \mathrm{CEI}_i(t)}{\partial r_i(t)} = \frac{\partial \mathrm{CEI}_{g(t)}(t)}{\partial r_{g(t)}(t)}.
$$

如果满足此条件，那么模拟$k(t)$的总影响可能大于模拟$g(t)$，因此我们优先模拟$k(t)$而非$g(t)$，即$x(t) = k(t)$。另一方面，如果不满足此条件，则优先模拟$g(t)$，即$x(t) = g(t)$。由此得到算法3中的 gCEI策略。

算法3gCEI策略

1：令$x(0) = i$，其中$i \in \mathscr{S}$。获取$Y(1)$并更新 $\mathscr{F}^1$。此外，令$t \leftarrow 1$。2：**当** $t < R$ **时**

3：**如果**

$$
\sum_{i\in\mathscr{S}\backslash\{k(t)\}}\partial \mathrm{CEI}_i(t)/\partial r_{k(t)}(t) \le \min_{i\in\mathscr{S}\backslash\{k(t)\}}\{\partial \mathrm{CEI}_i(t)/\partial r_i(t)\}
$$

then

4：x(t) = k(t)。5：否则 6：x(t) = g(t)，其中 g(t) = $\mathrm{argmin}_{i\in\mathscr{S}\backslash\{k(t)\}}\{\partial \mathrm{CEI}_i(t)/\partial$ $r_i(t)\}$。7：结束 if 8：通过模拟 x(t) 获取 Y (t + 1)，更新 $\mathscr{F}^{t+1} \leftarrow \mathscr{F}^t \cup \{x(t)、Y(t+1)\}$ 并将 t ←t + 1。9：结束 while 10：将 k(R) = $\mathrm{argmax}_{i=1,\ldots,k}\{\bar{\mu}_i(R)\}$ 作为选定的最佳系统返回。

这里简要说明 gCEI 收敛到 Glynn 和 Juneja（2004）速率最优分配的证明。首先，很容易证明，当$R \to \infty$时，gCEI策略将对所有系统进行无限次模拟。

仿照 Ryzhov（2016）中的分析，我们将真实均值代入估计值，考察 gCEI 的确定性行为。这意味着 $k(t) = k$ 对所有 t 均成立，并且

$$\frac{\partial \mathrm{CEI}_i(t)}{\partial r_i(t)} = -\frac{\sigma_i^2}{(r_i(t))^2}\ \frac{1}{2\sqrt{\nu_i}}\ \phi\left(\frac{\mu_i - \mu_k}{\sqrt{\nu_i}}\right) \text{ and } \frac{\partial \mathrm{CEI}_i(t)}{\partial r_k(t)} = -\frac{\sigma_k^2}{(r_k(t))^2}\ \frac{1}{2\sqrt{\nu_i}}\ \phi\left(\frac{\mu_i - \mu_k}{\sqrt{\nu_i}}\right).$$

考虑经验分配 $\{r_i(t)/t,\ i = 1,\ 2,\ ..,\ k\}$。我们知道它必定存在一个收敛子序列，$r_i(t)/t_{\ t} \xrightarrow{\to\infty} \alpha_i$ ；我们证明任何这样的子序列都必须收敛到速率最优分配（完整证明还包括证明该子序列的极限不为 0 ）。令 $\nu_i' = (\sigma_i^2/\alpha_i + \sigma_k^2/\alpha_k) = \lim_{t\to\infty} t\nu_i$。

首先考虑 第 3 步中不等式成立的子子序列。对于这样的迭代

$$\sum_{i\neq k} \frac{\sigma_k^2/(r_k(t))^2}{\sigma_j^2/(r_j(t))^2}\sqrt{\frac{\nu_j}{\nu_i}} \times \exp\left\{-\frac{1}{2}\left(\frac{(\mu_i - \mu_k)^2}{\nu_i} - \frac{(\mu_j - \mu_k)^2}{\nu_j}\right)\right\} \geq 1 \text{ for any } j \neq k. \tag{2}$$

然而，随着$t\ \to\infty$，指数项将趋于 0，或者 ∞ 除非

$$\frac{(\mu_i - \mu_k)^2}{\nu_i'} = \frac{(\mu_j - \mu_k)^2}{\nu_j'},\ i \neq j \neq k. \tag{3}$$

因此，要使方程（2）对任意j成立，方程（3）必须成立；这就是 Glynn 和 Juneja（2004）速率最优分配的两个条件中的第一个。因此，随着t 增加，不等式（2）变为（经过一些变换后）

$$\sum_{i\neq k} \frac{\sigma_k^2/(r_k(t))^2}{\sigma_j^2/(r_j(t))^2}\sqrt{\frac{1}{\nu_i}} \geq \sqrt{\frac{1}{\nu_j}} \text{ for any } j \neq k. \tag{4}$$

将两边相加　在$j = 1,2,\ldots,k-1$上求和，约去公因子，并令$t\ \to\infty$，得到　s

$$\sum_{j\neq k} \left(\frac{\sigma_k}{\alpha_k}\right)^2 \left(\frac{\alpha_j}{\sigma_j}\right)^2 \geq 1. \tag{5}$$

接下来考虑这样一个子子序列：在步骤3中选择了劣势系统$j\ \neq k$。这使不等式（4）反向，并且对每个$j\ \neq k$都必须成立。接着用类似的论证可知，（5）的左边必须为 ≤ 1。因此必须取等号，这就是 Glynn 和 Juneja（2004）的第二个条件。

6 经验性能

我们总共进行了16项实验，其中系统数量$k \in \{5,10,20,30\}$取了四个不同的值。对于每个k，我们设定 $\mu_i = cm_i$ ，其中m_i'是表1中给出的预缩放真实均值，c 是一个缩放常数，具体将在下文说明。在滑移配置和均值递增配置中，各系统的方差相等。在另外两种配置中，均值递增，而方差分别与预缩放均值成正比和成反比。

在每次实验中，我们先为每个系统分配 2 次重复仿真，然后再应用任何策略。为了构造合理的情形，我们对真实均值进行缩放，使得在 Glynn 和 Juneja（2004）的速率最优策略下，最佳系统与次佳系统之间的差异达到其估计差值的一个标准误之前，至少会消耗r_0 次重复仿真。具体而言，

$$\mu_k - \mu_{k-1} = c(m_k - m_{k-1}) = \sqrt{\frac{\sigma_{k-1}^2}{r_0\alpha_{k-1}^\star} + \frac{\sigma_k^2}{r_0\alpha_k^\star}}.$$

表1：实验的配置。

配置	预缩放真实均值	真实标准差
滑移	$m_i = -1$ 对于 $i \neq k$ 和 $m_k = 0$	$\sigma_i = 1$
递增均值	$m_i =$对数(i)	$\sigma_i = 1$
递增方差	$m_i = \log(i+1)$	$\sigma_i = \sqrt{m_i}$
递减方差	$m_i = \log(i+1)$	$\sigma_i = 1/\sqrt{m_i}$

因此，我们可以控制最佳系统与其他系统变得可区分的速度。为了求出满足方程的c，我们首先使用表1中的m_i，通过求解 Glynn 和 Juneja（2004）速率最优分配的表达式，计算$\alpha_i^\star$。常数c 不会改变最优分配，因为缩放所有 μ_i或所有 σ_i都不会产生任何影响。同时，我们希望总仿真预算R 足够大，以便观察各策略的收敛行为。我们设定$r_0 = 20k$，并令$R = 100k$。最后，我们将宏观重复次数M 设为 5000，以便在一系列取值范围内将 PICS 估计到小数点后两位。为了衡量各策略的性能，我们报告 $\widehat{\text{PICS}}(t) = \sum_{\tau=1}^{t} \mathbb{I}_{\{k(\tau)\neq k\}}/t$、 $\widehat{\alpha}_k(t) = r_k(t)/t$，以及每次迭代$t$时$\mu_k - \mu_{k(t)}$（即最优性差距）的均值和标准差。

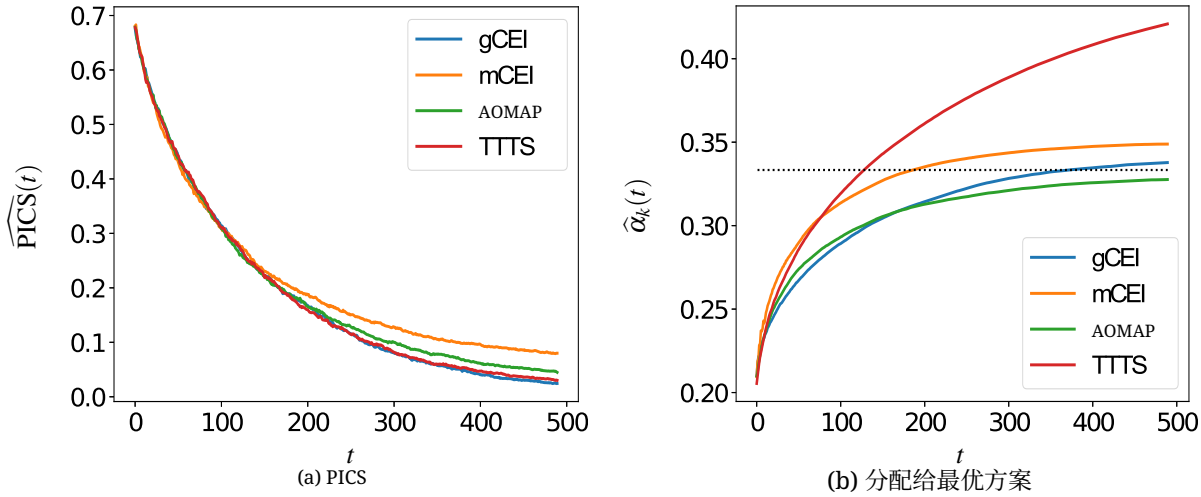
图1–2展示了滑移配置分别在$k = 5$ 和$k = 30$时的结果。为更仔细地观察尾部行为，图3展示了$k = 5$ 时预算较大的$R = 5000$ 对应的结果。我们总结主要观察结果如下：在每种策略下，PICS 都如预期收敛至零。不过，收敛行为并非总是相同，而且 gCEI 的表现似乎不逊于、甚至优于其他策略。对于最佳系统，gCEI 和 mCEI 都会先超调 Glynn 和 Juneja（2004）的渐近最优分配，但随后长期来看会如预期般收敛。与 mCEI 相比，gCEI 往往向最佳系统分配较少。另一方面，TTTS 向最佳系统分配的资源多得多；鉴于其源自多臂老虎机领域并以最小化遗憾为目标，这一点合情合理。请注意，AOMAP 和 TTTS 的极限分配并非Glynn 和 Juneja 的极限分配，因此我们不期望它们具有相同的分配。

对于滑移配置，平均最优性差距是 PICS 的一个缩放版本，因为无论选择哪个劣势系统作为最佳系统，差距都相同。最优性差距的标准差呈现出一种看似出乎意料的变化：先增大，再减小。这是因为在预算较小时，最佳系统难以辨别，而且所有劣势系统的均值都相同。随着最佳系统越来越容易辨认，变异性会逐渐增大至某个程度；随后，随着各策略对最佳系统身份的判断越来越确定，变异性便会下降。

为了了解动态策略相对于从一开始就采用 Glynn 和 Juneja（2004）的渐近最优分配时的表现，我们在滑移配置中比较 gCEI 与一种不切实际的策略：假设已知 Glynn 和 Juneja 的分配，并从第一次迭代开始就采用该分配，其中$k = 5$。更具体地说，在这种不切实际的策略下，每为一个劣势系统分配一次重复仿真，就应为最佳系统分配两次重复仿真。图4展示了这一比较结果，我们仅报告迭代次数为六的倍数时的结果。尽管 gCEI 在一段时间内向最佳系统分配了过多资源，但其表现仍优于这种不切实际的策略。这强调了速率最优分配针对的是大样本情形，而非小样本情形。

最后，图5展示了 $k = 10$ 时递增方差配置的结果。由于其他配置的结果与此非常相似，我们不再报告。根据我们的指标，所有策略在此处的表现都相近。唯一的差异体现在它们对最佳系统的分配上。与滑移配置类似，gCEI 分配给最佳系统的资源少于 mCEI。不过，与滑移配置不同，mCEI 和 gCEI 并未超出 Glynn 和 Juneja（2004）的渐近最优分配。

滑移配置当然不切实际，但它代表了一种最佳系统附近有许多接近的竞争者的情形。在这种情况下，gCEI似乎具有一定优势。当均值递增时，所有策略似乎都更容易控制PICS和最优性差距，因为劣势



(a) PICS　(b) 分配给最优方案

(c) 平均差距　(d) 差距标准差

图 1：当 $k = 5$、$R = 500$ 和 $M = 5000$ 时的滑移配置。(b) 中的虚线表示 Glynn & Juneja 提出的分配给最佳系统的最优分配。

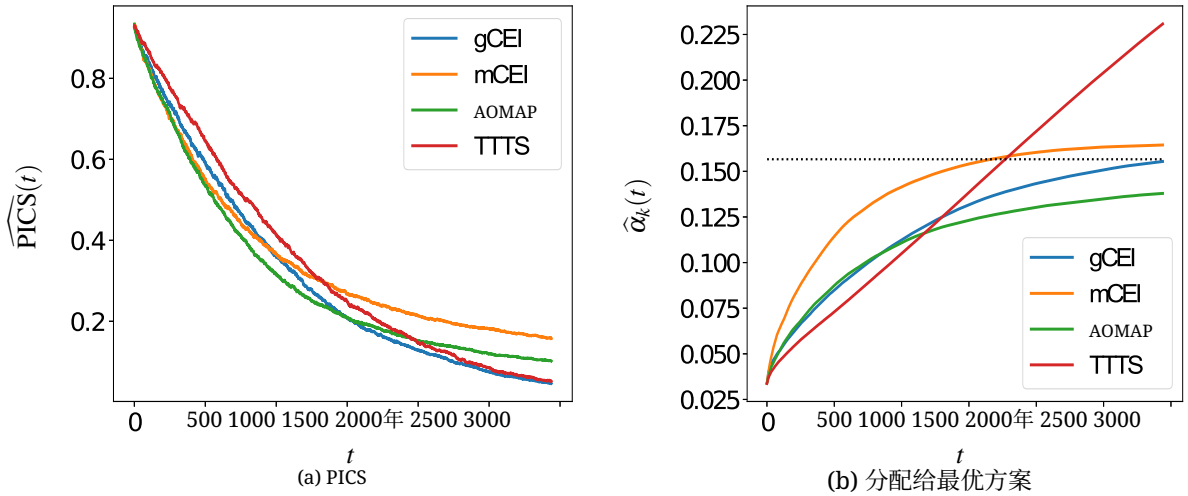
系统更容易识别；我们发现，在这些情况下，所有策略都倾向于将大多数重复仿真分配给排名前两位的系统。

7 结论

本文考察了三种近期提出的策略以及一种新策略，用于在固定预算排序与选择问题中为各系统分配重复仿真。所有策略都会随着获得更多仿真输出而进行自适应，并且随着预算增加，每种策略都能实现某种形式的最优分配；它们对“最优”的定义以及小样本行为各不相同。考察 PICS，以及终止时最优性差距的均值和标准差后，gCEI 的表现似乎不逊于 AOMAP、mCEI 和 TTTS，甚至更好。我们的比较没有考虑计算工作量（重复次数除外），也没有考虑能否在预先指定 PCS 的条件下停止；这些指标同样可以区分不同的排序与选择程序。

致谢

本研究部分由美国国家科学基金会资助，项目编号为 DMS-1854562。



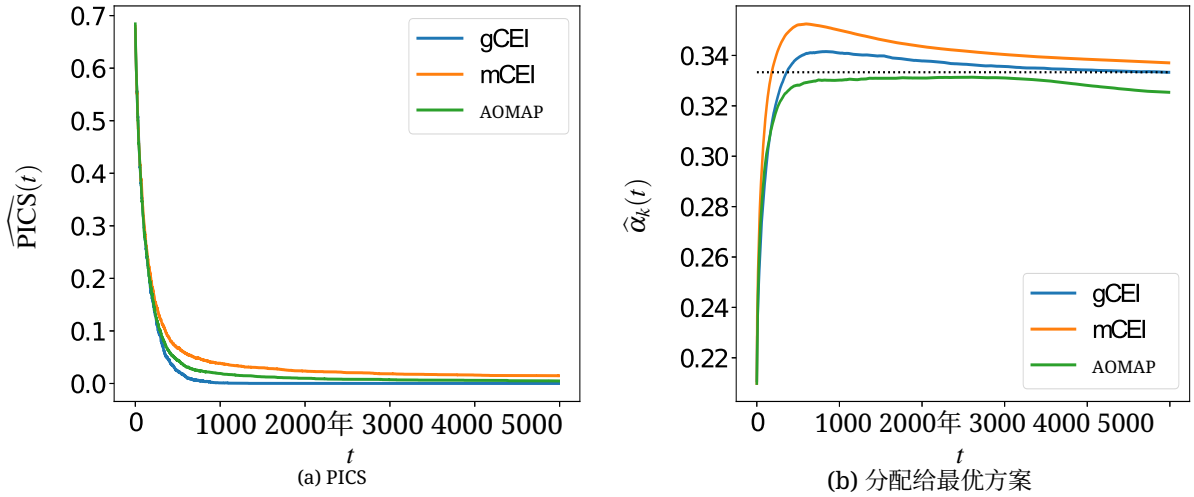
(a) PICS (b) 分配给最优方案

(c) 平均差距 (d) 差距标准差

图 2：$k = 30$、$R = 3000$ 和$M = 5000$下的滑移配置。(b) 中的虚线表示 Glynn & Juneja 提出的分配给最佳系统的最优分配。

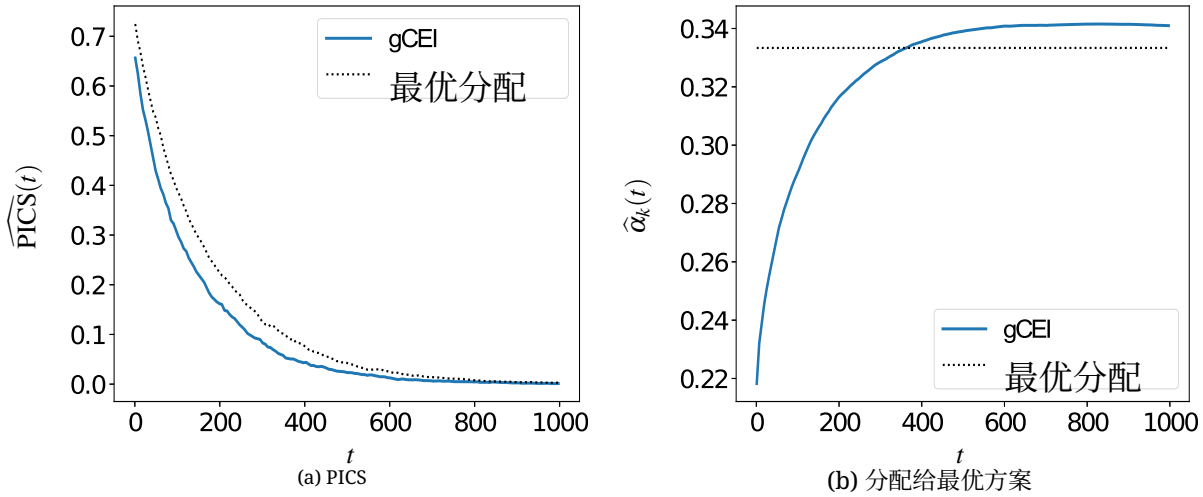
参考文献

Chen, C.-H., J. Lin, E. Y¨ucesan, and S. E. Chick. 2000年. “用于进一步提升序数优化效率的仿真预算分配”。离散事件动态系统10(3):251–270. Chen, Y., and I. O. Ryzhov. 2019a. “排序与选择中的最优大偏差平衡”。载于*2019*年冬季仿真会议论文集，由 N. Mustafee、K.-H. G. Bae、S. Lazarova-Molnar、M. Rabe、C. Szabo、P. Haas 和 Y.-J. Son 编，3368–3379。新泽西州皮斯卡塔韦：电气与电子工程师协会。Chen, Y., and I. O. Ryzhov. 2019b. “完全期望改进收敛至最优预算分配”。应用概率进展51(1):209–235. De Groot, M. H. 1970. 最优统计决策。纽约：约翰·威立父子出版公司。Frazier, P., and W. Powell. 2007年. “用于独立正态奖励离线学习的知识梯度策略”。载于*2007*年*IEEE*近似动态规划与强化学习国际研讨会，143–150。新泽西州皮斯卡塔韦：电气与电子工程师协会。Glynn, P., and S. Juneja. 2004. “序数优化的大偏差视角”。载于*2004*年冬季仿真会议论文集，由 R. G. Ingalls、M. D. Rossetti、J. S. Smith 和 B. A. Peters 编，577–585。新泽西州皮斯卡塔韦：电气与电子工程师协会。Jamieson, K., and R. Nowak. 2014. “固定置信度设置下多臂老虎机的最佳臂识别算法”。载于*2014*年第*48*届信息科学与系统年度会议（*CISS*），1–6。



(a) PICS　　(b) 分配给最优方案

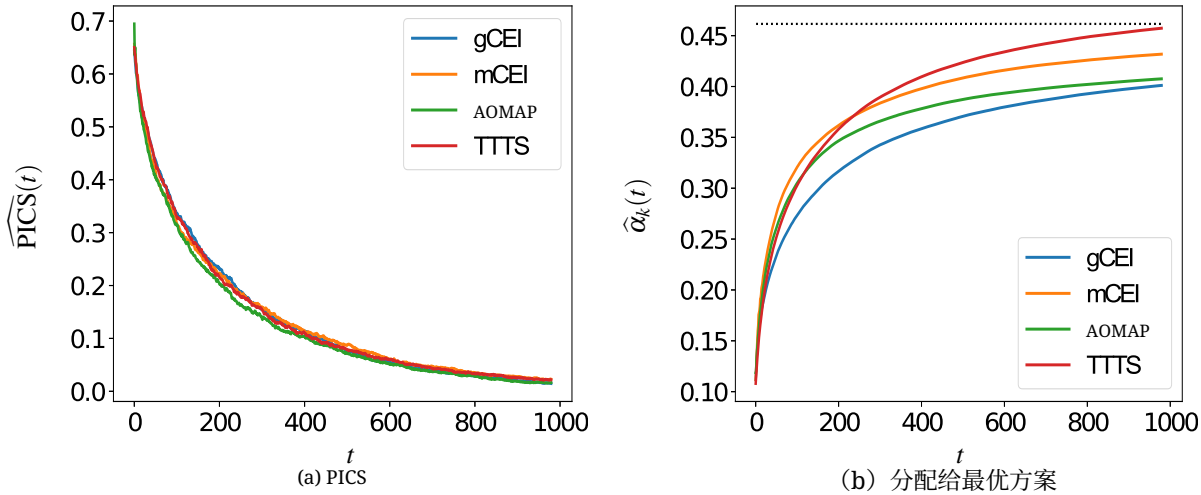
图3：包含 $k = 5$、$R = 5000$ 和$M = 5000$的滑移配置。(b) 中的虚线表示 Glynn & Juneja 提出的将分配给最优方案的最优分配。



(a) PICS　　(b) 分配给最优方案

图4：具有$k = 5$、$R = 1000$ 和$M = 5000$的滑移配置。(b)中的虚线表示 Glynn & Juneja 提出的分配给最佳系统的最优分配。

Jones, D. R., M. Schonlau, 和W. J. Welch。1998年。“昂贵黑箱函数的高效全局优化”。*Journal of Global Optimization* 13(4):455–492。Peng, Y., 和 M. C. Fu。2017。“具有渐近最优采样率的短视分配策略”。*IEEE Transactions on Automatic Control* 62(4):2041–2047。Russo, D. 2020。“用于最优臂识别的简单贝叶斯算法”。*Operations Research* 68(6):1625–1647。Ryzhov, I. O. 2016。“期望改进方法的收敛速率”。*Operations Research* 64(6):1515–1528。Salemi, P. L., E. Song, B. L. Nelson, 和 J. Staum。2019。“基于仿真的离散优化中的高斯马尔可夫随机场：框架与算法”。*Operations Research* 67(1):250–266。Thompson, W. R. 1933。“根据两组样本的证据，一个未知概率超过另一个未知概率的似然”。*Biometrika* 25(3/4):285–294。



(a) PICS

（b）分配给最优方案

（c）平均差距

（d）差距标准差

图5：递增方差配置，其中$k = 10$、$R = 1000$ 和$M = 5000$。图（b）中的点线表示 Glynn & Juneja 针对最佳系统的最优分配。

作者简介

哈伦·阿夫吉 是西北大学工业工程与管理科学系的博士生。他的研究兴趣包括仿真优化和仿真方法论。他的电子邮箱地址是 harun.avci@u.northwestern.edu。

巴里·L·纳尔逊 是西北大学工业工程与管理科学系的沃尔特·P·墨菲教授。他是美国运筹学与管理科学研究学会和美国工业工程师学会会士，并著有《随机仿真的基础与方法：入门课程》，由施普林格出版。他的电子邮件和网站地址分别为 nelsonb@northwestern.edu 和http://www.iems.northwestern.edu/~nelsonb。

¨安德烈亚斯·瓦赫特 是西北大学工业工程与管理科学系教授。他的研究主要聚焦于非线性优化数值算法的设计、分析、实现与应用。他因开发开源优化软件包 Ipopt 而获得 J·H·威尔金森数值软件奖，并且是工业与应用数学学会会士。他的电子邮件地址是andreas.waechter@northwestern.edu。