高效学习以选择排名前-m的上下文相关设计

Gongbo Zhang、Sihua Chen、Kuihua Huang 和 Yijie Peng

arXiv:2305.04086v2 [stat.ML] 9 Jun 2023

摘要——**我们研究依赖上下文的决策中的一个仿真优化问题，其目标是为所有上下文确定排名前-m 的设计。在贝叶斯框架下，我们将最优动态抽样决策表述为随机动态规划问题，并提出一种序贯抽样策略，以高效学习每种设计在每个上下文下的性能。我们推导出渐近最优抽样比例，以达到最坏情形错误选择概率的最优大偏差率。我们证明了所提出的抽样策略是一致的，并表明其渐近抽样比例是渐近最优的。数值实验表明，所提出的方法提高了选择排名前-m的上下文相关设计的效率。**

给实践者的说明——**给定设计的性能可能因上下文而异，通过完善可用的上下文信息，可以实现更好的决策。我们研究一个依赖上下文的排序与选择问题，在该问题中，最优选择可取决于决策前获得的上下文信息。我们提出一种动态抽样方案，以高效学习并选择所有上下文中排名前-m 的设计。数值实验（包括蜜罐欺骗博弈问题和医疗资源分配问题）表明，所提出的抽样方案显著提高了依赖上下文的排序与选择排名前-m设计的效率。**

索引词—**仿真优化；依赖上下文的决策；top-m 选择；动态抽样；渐近最优性。**

I. 引言

仿真IMULATION是分析和优化复杂随机系统的有力工具。近年来，人们开始关注这样一类情形：复杂系统可以进行仿真，并根据可用的上下文信息，在个体层面作出多个决策。我们考虑一个仿真优化问题：对于所有已实现的上下文，从有限的 k 个设计 ($1 \leq m < k,\ m,\ k \in \mathbb{Z}^+$) 中确定均值最高（最低）的前 m 个设计。每个设计在每种上下文下的平均表现都是未知的，只能通过蒙特卡洛仿真进行估计且仿真预算有限。每个设计的表现取决于上下文，排名前 m 的设计也依赖上下文。例如，在新冠疫情期间急诊科的医疗资源分配中，良好的分配设计可能会随着疫情演变所引发的急诊患者流变化而改变 [1], [2]。在营销和金融等其他领域，企业可以根据用户的位置、人口统计特征和兴趣提供个性化定价、广告、促销活动及网页布局，从而借助上下文信息增加收入并提升用户满意度 [3], [4]。

对于某个上下文和 $m = 1$，该问题可归结为在单个上下文下选择最佳设计，这在仿真文献中通常称为排序与选择（R&S）问题。R&S 程序在有限设计集合中高效分配仿真预算，以学习最佳设计。正确选择概率（PCS）是评估 R&S 中抽样程序统计效率的常用指标。经典问题已推广到选择前 m 个设计 [5]–[8]，这为决策者提供了更多选择，使他们能够融入个人偏好并适应突发事件，从而做出更具韧性的决策。与 R&S 中的前 m 个选择相比，我们问题中的前 m 个设计取决于决策制定之前已实现的上下文。在本文中，我们旨在为所有上下文正确识别前 m 个设计。在医疗资源分配的例子中，危重患者在某一分配设计下的等待时间取决于急诊患者流，而急诊患者流在新冠疫情期间不断变化。例如，德尔塔变种比阿尔法变种传染性更强，导致对医疗服务的需求增加。一些分配设计在紧急情况下可能变得不可用，例如设备故障和人员重新部署，从而使具有最短等待时间的最佳分配设计变得不可行。针对每种急诊患者流情形准备前 m 个分配设计，将为医院管理提供更灵活的决策支持。我们的抽样程序旨在在所有设计-情境对之间高效分配有限的仿真预算。对于每个上下文，任何前 m 个设计的错误选择都可能导致我们的目标失败。为衡量抽样程序在整个上下文空间中的统计效率，我们使用最坏情形 PCS（PCS_W）作为指标，其中“最坏情形”指 PCS 最低的上下文。

采样效率是排序与选择中的核心问题，因为仿真通常计算成本高昂。在我们的问题中，

Gongbo Zhang 和 Yijie Peng 任职于中国北京大学光华管理学院管理科学与信息系统系，邮编 100871（电子邮箱：gongbozhang@pku.edu.cn；pengy-ijie@pku.edu.cn）。Sihua Chen 任职于中国南昌江西财经大学信息管理学院，邮编 330013（电子邮箱：doriancsh@foxmail.com）。Kuihua Huang 任职于中国长沙国防科技大学系统工程学院，邮编 410073（电子邮箱：khhuang@nudt.edu.cn）。通讯作者：Sihua Chen 和 Yijie Peng。论文修订日期：2021年12月25日；2022年1月12日；2022年10月11日；2023年1月4日；2023年6月9日

可能存在大量候选设计和上下文，并且在抽样程序中需要考虑所有设计-情境对。因此，学习效率比经典排序与选择问题中的效率更为重要。为了提高设计-情境对的学习效率，我们的研究目标是将仿真预算高效地分配给各个设计-情境对，从而最大化选择正确概率$_W$ 。我们提出了一种动态采样策略，以高效地学习每个设计-情境对的性能。为刻画该抽样策略的渐近最优性，我们通过优化最坏情形错误选择概率（错误选择概率$_W$，即 $1-$选择正确概率$_W$）的大偏差率，推导出渐近最优抽样比例。所提出的抽样程序具有一致性，即随着仿真预算趋于无穷，最终能够选出所有上下文下排名前m的设计；并且证明，在正态抽样分布下，其渐近抽样比例是渐近最优的。我们进行了严格分析，从渐近角度刻画了所提出的抽样程序中各设计-情境对的抽样策略。总而言之，我们的工作有三方面贡献。

- 据我们所知，本文首次研究了排序与选择问题中前-m （特别是 $m \neq 1$）个依赖上下文的设计的选择，并在随机控制框架下提出了一种动态采样策略。仿真预算需要分配给所有上下文中的设计，以便对其进行学习，这带来了一个重大挑战，使该策略有别于经典的排序与选择方法。

- 我们推导了前-m个上下文相关选择问题的渐近最优性条件，将 [8] 和 [9]中的结果加以推广。具体而言，当 $m \neq 1$ 且$m \neq k-1$时，所推导的渐近最优性条件仅是实现渐近最优抽样比例的必要条件。与 [9]相比，这使得刻画抽样策略的渐近最优性面临重大挑战。此外，所推导的渐近最优性条件揭示了在某一上下文内采样与在不同上下文间采样之间的权衡，这使我们的分析有别于 [8]中的分析。

- 我们设计了一种高效的动态采样策略，称为AOAmc。该策略基于价值函数近似（VFA）的单步前瞻方法。我们证明了所提出策略的一致性和渐近最优性。在这些证明中，我们的理论贡献包括应对刻画渐近最优性所面临的挑战，以及平衡不同设计和上下文之间的抽样。这些挑战使我们的证明有别于 [8]中的证明。我们通过涉及合成示例以及两个实际应用——蜜罐欺骗博弈问题和医疗资源分配问题——的数值实验，展示了该抽样策略的有效性。

A. 相关文献

现有的排序与选择程序通常分为固定精度程序和固定预算程序。概述见 [10]–[12]。固定精度程序 [13], [14] 分配仿真预算，以保证达到预先设定的选择正确概率水平；而固定预算程序 [15], [16] 则旨在固定仿真预算的约束下最大化后验性能。序贯固定预算抽样策略的最优抽样与分配决策由马尔可夫决策过程的贝尔曼方程 [16], [17] 决定。在 [18], 中，研究者利用大偏差技术推导了选择最佳设计问题的渐近最优性条件。该研究表明，实现静态渐近最优抽样比例，能够使最佳设计的频率学派错误选择概率以最优指数速率衰减。因此，能够实现渐近最优抽样比例的抽样策略被视为渐近最优。近期研究将 [18] 中的结果拓展至选择前 m 个设计 [8] , 以及选择最佳上下文相关设计 [9]。已有研究证明，若干序贯固定预算抽样策略具有渐近最优性，其适用情形包括正态抽样分布 [8], [9],[16], [19]–[21], 和一般采样分布 [22]。证明某一抽样策略具有渐近最优性，通常需要证明其渐近抽样比例满足相应的渐近最优性条件。

尽管无上下文排序与选择问题已得到广泛研究并被充分理解，但关于上下文选择问题的文献仍然有限。文献中，上下文也称为协变量、个性化信息、辅助信息或辅助量。针对依赖上下文的选择问题，现有采样策略可根据上下文取值属于有限集还是无限集进行分类。在上下文空间有限的情况下，每个设计-情境对都可以进行采样。 [9], [23] 和 [24] 将 [18] 中的结果扩展至上下文选择问题。他们提出了采样策略，依次平衡渐近最优抽样比例所满足的条件。从频率学派的角度出发， [25]和 [26]基于无差异区间表述推导出一种两阶段算法。他们利用线性元模型刻画各设计与上下文之间的线性关系。沿着 [25], [27]的开创性工作， [21], 考虑了高维上下文，并将设定推广到线性函数之外更广泛的一类基函数。在 [9]中，研究者使用高斯混合模型将设计和上下文分组成聚类，以确保每个聚类中的设计-情境对具有相似的性能。该论文提出了一种动态采样策略，可满足 [28] 定义的渐近最优性条件。 [9]没有采用高斯混合元模型，而是针对高斯过程元模型提出了一种采样策略，以最大化后验PCS。该策略依次平衡与 [29], [30]中类似的渐近最优性条件。在上下文无限的情况下，对每个可能的上下文取值进行采样并不可行。 [31] 和从贝叶斯角度研究这一问题。他们采用期望改进和知识梯度等采集函数，这些函数是最优采样策略的短视近似，并推导出序贯采样策略，以在固定仿真预算下最大化良好选择的后验概率

约束。值得注意的是，现有研究均未考虑 top-m依赖上下文的选择问题。

我们的问题与机器学习中关于上下文多臂老虎机（MAB）的文献相关。上下文 MAB 问题通常假设存在某些结构，例如利普希茨结构 [32], [33], ，即预期奖励相对于上下文呈现利普希茨连续性；线性结构 [34], [35], ，即预期奖励与上下文呈线性依赖关系；以及使用神经网络 [36] 和随机森林 [37] 将上下文映射到设计的更一般映射。排序与选择（R&S）中不依赖上下文的前-m 选择，与 MAB 问题的一种纯探索版本相似，后者称为最佳-m臂识别。现有的最佳-m臂识别算法可分为固定置信度 [38]–[40] 和固定预算 [41] 两类方法。然而，关于依赖上下文的最佳-m臂识别的研究相对较少。据我们所知， [42] 是首项研究在固定置信度设置下考虑奖励与上下文呈线性依赖关系的最佳-m臂识别问题的工作。现有关于 MAB 的研究均未关注在固定预算设置下选择前-m 个依赖上下文的设计。

我们问题中使用的最坏情况性能指标考虑了针对攻击的极端事件风险。优化这一指标与鲁棒优化（RO）文献相关。鲁棒仿真是在模糊集 [43], [44] 内输入分布未知的情况下，对仿真模型进行最坏情况分析。在 RO 中，未知输入分布的参数被视为决策变量。相比之下，我们的工作侧重于根据最坏情况性能指标优化设计变量。上下文空间可以与一个模糊集相关联，该模糊集包含有限个表示未知输入分布的参数，其中包括真实的输入分布。由此，我们的问题设定与排序与选择文献中的分布鲁棒优化（DRO）相关。DRO 旨在从模糊集中找出最坏情况性能最佳的鲁棒设计 [45], [46], 或最坏情况性能排名前 m 的鲁棒设计 [47] ；这些具有最坏情况性能的鲁棒设计具有普适性。相比之下，我们的问题涉及依赖上下文的前 m 个设计，且输入分布固定。

本文其余部分安排如下。第二节对所研究的问题进行表述。第三节提出一种计算高效的动态抽样程序。第四节推导渐近最优抽样比例，并证明所提出的抽样程序具有渐近最优性。第五节展示数值示例和计算结果，第六节总结全文并展望未来研究方向。

II. 问题表述

假设存在 k 个相互竞争的设计，且每个设计的性能 $y_i(\boldsymbol{x}) \in \mathbb{R}$, $i = 1, \cdots, k$ 取决于一个上下文向量 $\boldsymbol{x} = (x_1, \cdots, x_d)^T$，其中 $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^d$，而 $y_i(\boldsymbol{x})$ 未知，只能通过蒙特卡洛仿真进行估计。我们假设上下文空间 $\mathcal{X}$ 包含有限个 q 可能的上下文 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_q$，并且在某个上下文 $\boldsymbol{x}_\ell$ $\ell = 1 \cdots q$ $y_{\langle 1 \rangle_\ell}(\boldsymbol{x}_\ell) \geq \cdots \geq y_{\langle m \rangle_\ell}(\boldsymbol{x}_\ell) >$,, $y_{\langle m+1 \rangle_\ell}(\boldsymbol{x}_\ell) \geq \cdots \geq y_{\langle k \rangle_\ell}(\boldsymbol{x}_\ell)$ $\langle i \rangle_\ell$ $i = 1 \cdots k$ 中存在唯一的一组前-m 个设计，即 ( ) ( ) ( ) ( )，其中，,, $\ell = 1, \cdots, q$是依据上下文相关设计的性能排序的索引。我们的目标是针对上下文 $\boldsymbol{x}$ 的所有可能取值，正确选出前-m个设计，即识别$\mathcal{F}_\ell^m \triangleq \{\langle 1 \rangle_\ell, \langle 2 \rangle_\ell, \cdots, \langle m \rangle_\ell\}$, $\ell = 1, \cdots, q$。图 1 展示了一个示例。在医疗资源分配中，例如，我们旨在针对每种急诊患者流（上下文），确定能使危重患者等待时间最短的前-m 个分配决策（设计）。该框架涵盖不同的 m 取值，以及每个上下文中不同数量的设计。不过，为简化记号，我们假设各上下文中的 m 和设计数量均相同。

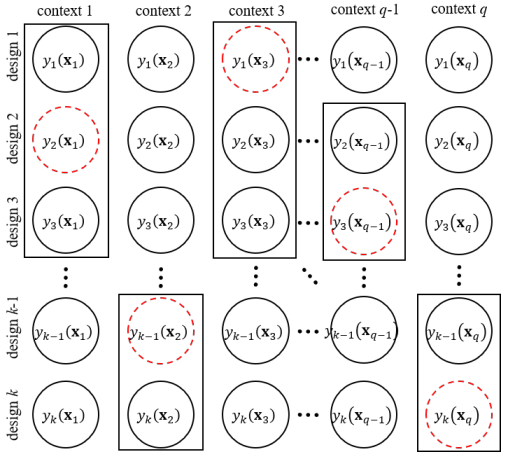


图 1. 在所有上下文中选取前 m 个设计的示意图。红色虚线圆圈表示最佳设计，黑色实线圈出每个上下文中的前 m 个设计，其中 m 的数量可能因上下文而异。

令 $Y_{i\ell}^h$, $i = 1, \cdots, k$, $\ell = 1, \cdots, q$, $h \leq T$ 表示设计 i 在情境 $\boldsymbol{x}_\ell$下第 h 次独立同分布（i.i.d.）仿真重复，其中 $T < \infty$ 为仿真预算总量。为简化起见，我们在 $Y_{i\ell}^h$ 的记号中省略 $\boldsymbol{x}_\ell$ 。我们假设，对于每个设计-情境对，仿真重复均服从 i.i.d. 正态抽样分布，即 $Y_{i\ell}^h \sim N\left(y_i(\boldsymbol{x}_\ell)\right.$ $\left.\sigma_i^2(\boldsymbol{x}_\ell)\right)$，且不同设计和情境之间的重复试验相互独立。在这项工作中，我们假设抽样分布中的 $\sigma_i^2(\boldsymbol{x}_\ell)$ 已知，并在实践中使用样本估计值作为真实值的代入估计量。正态性假设是排序与选择文献中的常见假设；对于非正态抽样分布，可通过批处理来证明正态性假设的合理性。我们引入贝叶斯框架来学习未知参数 $y_i(\boldsymbol{x}_\ell)$，并假设先验分布为共轭先验 $N(y_i^{(0)}(\boldsymbol{x}_\ell), (\sigma_i^{(0)}(\boldsymbol{x}_\ell))^2)$，其中超参数 $(\sigma_i^{(0)}(\boldsymbol{x}_\ell))^2$ 用于量化先验不确定性。在分配了 t, $t = 1, \cdots, T$ 次仿真后， $y_i(\boldsymbol{x}_\ell)$, $i = 1, \cdots, k$,$\ell = 1, \cdots, q$ 的后验分布 手机看片

重复试验为 $N(y_i^{(t)}(\boldsymbol{x}_\ell)$，$(\sigma_i^{(t)}(\boldsymbol{x}_\ell))^2)$，其中

$$(\sigma_i^{(t)}(\boldsymbol{x}_\ell))^2=\left(\frac{1}{(\sigma_i^{(0)}(\boldsymbol{x}_\ell))^2}+\frac{t_{i\ell}}{\sigma_i^2(\boldsymbol{x}_\ell)}\right)^{-1}, \quad (1)$$

$$y_i^{(t)}(\boldsymbol{x}_\ell)=(\sigma_i^{(t)}(\boldsymbol{x}_\ell))^2\left(\frac{y_i^{(0)}(\boldsymbol{x}_\ell)}{(\sigma_i^{(0)}(\boldsymbol{x}_\ell))^2}+\frac{t_{i\ell}\bar{Y}_{i\ell}}{\sigma_i^2(\boldsymbol{x}_\ell)}\right), \quad (2)$$

其中，$\bar{Y}_{i\ell}=\sum_{h=1}^{t_{i\ell}}Y_{i\ell}^h/t_{i\ell}$是样本均值，$t_{i\ell}$是在上下文$\ell$下分配给设计 i 的样本数量。如果 $\sigma_i^{(0)}(\boldsymbol{x}_\ell)\to\infty$，那么 $y_i^{(t)}(\boldsymbol{x}_\ell)=\bar{Y}_{i\ell}$。在这种情况下，先验不包含信息，后验分布仅包含样本信息，不涉及任何主观超参数。注意，后验均值 $y_i^{(t)}(\boldsymbol{x}_\ell)$是 $y_i^{(0)}(\boldsymbol{x}_\ell)$与 $\bar{Y}_{i\ell}$的加权平均值，并且后验方差（$\sigma_i^{(t)}(\boldsymbol{x}_\ell)$）^{2} 会随着 t 增加而单调不增，从而逐渐降低 $y_i(\boldsymbol{x}_\ell)$的不确定性。

抽样策略会智能地在设计-情境对之间分配固定的仿真预算，并根据收集到的样本选出估计的前m个依赖上下文的设计。在实际应用中，需要根据样本依次估计未知参数 $y_i(\boldsymbol{x}_\ell)$，因此，抽样决策实质上成为一种适应于分配样本所生成的信息滤过的策略。动态采样策略是一系列映射 $\mathcal{A}_T(\cdot)\triangleq(A_1(\cdot)\ \cdots,A_T(\cdot))$，其中 $A_t(\mathcal{E}_{t-1})\in\{(i,\ell):i\in\{1,\cdots,k\},\ \ell\in\{1,\cdots,q\}\}$ 根据信息滤过 $\mathcal{E}_{t-1}$，分配第 t次仿真重复，以估计上下文 $\boldsymbol{x}_\ell$ 下设计 i 的性能；其中，$\mathcal{E}_0$ 包含所有设计-情境对的先验分布的全部超参数。记 $t_{i\ell}\triangleq\sum_{h=1}^t A_{i\ell}^h(\mathcal{E}_{h-1})$为经过 t 步后相应的分配样本数量，其中 $A_{i\ell}^h(\mathcal{E}_{h-1})\triangleq\mathbb{1}(A_h(\mathcal{E}_{h-1})=(i,\ell))$，而1 (·)是一个指示函数，当括号中的事件为真时取值为1，否则取值为0。为简化表示，我们省略 $t_{i\ell}$ 对 $\mathcal{E}_{t-1}$ 的依赖。于是，信息滤过变为$\mathcal{E}_t\triangleq\{\mathcal{E}_0,Y(t)\ 11,\cdots,Y(t)_{1q},\cdots,Y(t)_{k1},\cdots,Y(t)_{kq}\}$, $t=1,\cdots,T$，其中 $Y(t_{i\ell}\triangleq(Y_{i\ell}^1,Y_{i\ell}^2,\cdots,Y_{i\ell}^{t_{i\ell}})$包含设计-情境对$(i,\ell)$的样本观测。在共轭先验下，$\mathcal{E}_t$完全由后验超参数 $y_i^{(t)}(\boldsymbol{x}_\ell)$和$(\sigma_i^{(t)}(\boldsymbol{x}_\ell))^2$确定。给定 $A_{t+1}\ \mathcal{E}_t=\widehat{i}\ \widehat{\ell}\quad\mathcal{E}_{t+1}=\mathcal{E}_t\cup Y_{\widehat{i}\widehat{\ell}}^{t_{\widehat{i}\widehat{\ell}}+1}\ \widehat{i}\in\{1,\cdots,k\}\ \widehat{\ell}\in\{1,\cdots,q\}\quad i\ \ell\ \neq\ \widehat{i}\ \widehat{\ell}\quad\sigma_i^{(t+1)}\ \boldsymbol{x}_\ell\ ^2\quad y_i^{(t+1)}\ \boldsymbol{x}_\ell$

( ) (,)和, , ,若(,) (,)，则(( ))和( )保持与$\sigma_i^{(t)}(\boldsymbol{x}_\ell))^2$和 $y_i^{(t)}(\boldsymbol{x}_\ell)$相同；否则，

$$\left(\sigma_i^{(t+1)}(\boldsymbol{x}_\ell)\right)^2=\left(\frac{1}{(\sigma_i^{(0)}(\boldsymbol{x}_\ell))^2}+\frac{t_{i\ell}+1}{\sigma_i^2(\boldsymbol{x}_\ell)}\right)^{-1}$$

$$y_i^{(t+1)}(\boldsymbol{x}_\ell)=(\sigma_i^{(t+1)}(\boldsymbol{x}_\ell))^2\left(\frac{y_i^{(t)}(\boldsymbol{x}_\ell)}{(\sigma_i^{(t)}(\boldsymbol{x}_\ell))^2}+\frac{Y_{i\ell}^{t_{i\ell}+1}}{\sigma_i^2(\boldsymbol{x}_\ell)}\right).$$

我们固定一种选择策略：在每个上下文中，耗尽所有 T 仿真重复后，选择后验均值最高的前 m 个设计。将上下文 $\boldsymbol{x}_\ell$ 中选出的前 m 个设计的集合记为

$\widehat{\mathcal{F}}_{\ell T}^m\triangleq\{\langle1\rangle_{\ell T},\langle2\rangle_{\ell T},\cdots,\langle m\rangle_{\ell T}\}$, $\ell=1,\ \cdots,\ q$，其中 $\langle i\rangle_{\ell T}$,$i=1,\ \cdots,\ k$是按后验均值排序的索引。对于某个上下文 $\boldsymbol{x}_\ell$，当 $\mathcal{F}_{\ell t}^m=\mathcal{F}_\ell^m$,$\ell\in\{1,\cdots,q\}$时，即实现了正确选择，而选择质量可以通过以下方式衡量：

$$\begin{aligned}&\mathrm{PCS}(\boldsymbol{x}_\ell)=\Pr\left\{\widehat{\mathcal{F}}_{\ell T}^m=\mathcal{F}_\ell^m\middle|\mathcal{E}_T\right\}\\&=\Pr\left\{\bigcap_{i=1}^m\bigcap_{j=m+1}^k\left(y_{\langle i\rangle_{\ell T}}^{(T)}(\boldsymbol{x}_\ell)>y_{\langle j\rangle_{\ell T}}^{(T)}(\boldsymbol{x}_\ell)\right)\middle|\mathcal{E}_T\right\}.\end{aligned}$$

我们问题中的PCS也依赖上下文。考虑到决策的稳健性，我们采用PCS_W来衡量在整个上下文空间 $\mathcal{X}$内的选择质量，即$\mathrm{PCS}_W=\min_{\ell=1,\ \cdots,\ q}\mathrm{PCS}(\boldsymbol{x}_\ell)$。无论是PCS $(\boldsymbol{x}_\ell)$、$\ell=1,\ \cdots,\ q$还是$\mathrm{PCS_W}$，都没有解析形式。我们的目标是提出一种动态采样策略 $\mathcal{A}_T(\cdot)$，以最大化PCS_W，从而为所有上下文选择前m 个设计。

III. 高效动态采样策略 Y

序贯采样决策可以表述为一个随机控制问题 [16], [17]。随机控制问题中动态采样策略 $\mathcal{A}_T(\cdot)$的期望收益可以递归定义为 $V_T(\mathcal{E}_T;\ \mathcal{A}_T(\cdot))\ \triangleq\min_{\ell=1,\ \cdots,\ q}\Pr\left\{\bigcap_{i=1}^m\bigcap_{j=m+1}^k\left(y_{\langle i\rangle_{\ell T}}^{(T)}(\boldsymbol{x}_\ell)>y_{\langle j\rangle_{\ell T}}^{(T)}(\boldsymbol{x}_\ell)\right)\middle|\mathcal{E}_T\right\}$，它是不同参数取值下$\mathrm{PCS_W}$的加权平均，即积分后验$\mathrm{PCS_W}$ (IPCS$_\mathrm{W}$)，且对于 $0\ \leq\ t\ <\ T$,

$h\ \in\ \{1,\cdots,k\}\ \ell\ \in\ \{1,\cdots,q\}\ V_t\ \mathcal{E}_t\ \mathcal{A}_T\ \cdot\quad\triangleq,,\ \mathbb{E}\left[V_{t+1}\left(\mathcal{E}_t\cup\{Y_{h\ell}^{t_{h\ell}+1}\}\ \mathcal{A}_T\ \cdot\right)\middle|\mathcal{E}_t\right]\big|_{(h,\ell)=A_{t+1}}\ \mathcal{E}_t$

(; ( )) ; 0()。

最优动态采样策略 $\mathcal{A}_T^*(\cdot)$ 可定义为随机控制问题的解，即 $\mathcal{A}_T^*(\cdot)\triangleq\arg\max_{\mathcal{A}_T(\cdot)}V_0(\mathcal{E}_0;\ \mathcal{A}_T(\cdot))$. 使用反向归纳法求解随机控制问题会遭遇维数灾难。为推导出具有解析形式的动态采样策略，我们采用近似动态规划（ADP）方案，该方案通过一步向前展望来寻找价值函数的单一特征，以分配下一次仿真重复，并使用确定性等价来近似价值函数 $V_T(\mathcal{E}_T;\ \mathcal{A}_T(\cdot))$。

假设任意步骤 t 都可能是最后一步。按照 [8], 中的VFA方案，针对某个上下文 $\boldsymbol{x}_\ell$, $\ell\in\{1,\cdots,q\}$，$\Pr\left\{\bigcap_{i=1}^m\bigcap_{j=m+1}^k\left(y_{\langle i\rangle_{\ell t}}^{(T)}(\boldsymbol{x}_\ell)>y_{\langle j\rangle_{\ell t}}^{(T)}(\boldsymbol{x}_\ell)\right)\middle|\mathcal{E}_t\right\}$是一个 $[m\times(k-m)]$ 维标准正态分布的密度在若干超平面所覆盖区域上的积分。由于正态密度呈指数衰减，覆盖区域内最大内切球上的积分可以很好地表示该积分。因此，$V_t(\mathcal{E}_t;\mathcal{A}_t(\cdot))$的一个近似值为 $\widetilde{V}_t(\mathcal{E}_t)\triangleq\min_{\ell=1,\ \cdots,\ q}$ 近似 $\mathrm{PCS}_\ell(\mathcal{E}_t)$，其中最大内切球的半径

$$\mathrm{APCS}_\ell(\mathcal{E}_t)\triangleq\min_{\substack{i=1,\cdots,m\\j=m+1,\cdots,k}}\frac{\left(y_{\langle i\rangle_{\ell t}}^{(t)}(\boldsymbol{x}_\ell)-y_{\langle j\rangle_{\ell t}}^{(t)}(\boldsymbol{x}_\ell)\right)^2}{\left(\sigma_{\langle i\rangle_{\ell t}}^{(t)}(\boldsymbol{x}_\ell)\right)^2+\left(\sigma_{\langle j\rangle_{\ell t}}^{(t)}(\boldsymbol{x}_\ell)\right)^2},$$

被视为PCS($\boldsymbol{x}_\ell$)的一项特征。在上下文 $\boldsymbol{x}_r$ 中为设计 h 分配一次仿真重复 $h\ \in\{1,\cdots,k\}\ r\ \in\ \{1,\cdots,q\}$后，采用确定性等价 [48]

随后应用于第 t 步的VFA一步前瞻:
$\mathbb{E}[\widetilde{V}_{t+1}\left(\mathcal{E}_t \cup \{Y_{hr}^{t_{hr}+1}\}\right)|\mathcal{E}_t] \approx \widetilde{V}_{t+1}\left(\mathcal{E}_t \cup \mathbb{E}\left[Y_{hr}^{t_{hr}+1}\middle|\mathcal{E}_t\right]\right) = \min_{\ell=1,\cdots,q}$近似$\mathrm{PCS}_\ell\left(\mathcal{E}_{t+1}^E\right)$，其中 $\mathcal{E}_{t+1}^E \triangleq \mathcal{E}_t \cup \mathbb{E}\left[Y_{hr}^{t_{hr}+1}\middle|\mathcal{E}_t\right]$，并且

$$\mathrm{APCS}_\ell\left(\mathcal{E}_{t+1}^E\right) \triangleq \min_{\substack{i=1,\cdots,m \\ j=m+1,\cdots,k}} \frac{\left(y_{\langle i\rangle_{\ell t}}^{(t)}(\boldsymbol{x}_\ell) - y_{\langle j\rangle_{\ell t}}^{(t)}(\boldsymbol{x}_\ell)\right)^2}{\left(\sigma_{\langle i\rangle_{\ell(t+1)}}^{(t+1)}(\boldsymbol{x}_\ell)\right)^2 + \left(\sigma_{\langle j\rangle_{\ell(t+1)}}^{(t+1)}(\boldsymbol{x}_\ell)\right)^2}, \tag{3}$$

其中，式 (3) 成立是因为额外的重复抽样 $Y_{hr}^{t_{hr}+1}$ 取其后验均值 $\mathbb{E}\left[Y_{hr}^{t_{hr}+1}\middle|\mathcal{E}_t\right]$的值，这 $\widetilde{}$意味着 $y_{\widetilde{i}}(\boldsymbol{x}_\ell)$、 $i = 1, \cdots, k$ 的所有后验均值均保持不变，只有 $y_h(\boldsymbol{x}_r)$ 的后验方差会更新。注意，在上下文 $\boldsymbol{x}_r$ 中为设计 h 分配一次仿真重复抽样，可以降低 $y_h(\boldsymbol{x}_r)$ 的后验方差，而 $\widetilde{V}_{t+1}\left(\mathcal{E}_t \cup \mathbb{E}\left[Y_{hr}^{t_{hr}+1}\middle|\mathcal{E}_t\right]\right)$则反映了所有上下文中的最差性能。令 $\widehat{V}_t(\mathcal{E}_t; (h, r)) \triangleq \widetilde{V}_{t+1}\left(\mathcal{E}_t \cup \mathbb{E}\left[Y_{hr}^{t_{hr}+1}\middle|\mathcal{E}_t\right]\right)$，则对于 $h = 1, \cdots, k$、 $r = 1, \cdots, q$ 以及 $t = 0, \cdots, T-1$，用于选择排名前 m 个上下文相关设计的渐近最优抽样策略 (AOAmc)——其最大化 VFA一步前瞻——定义如下:

$$A_{t+1}\left(\mathcal{E}_t\right) \in \arg\max_{(h,r)} \widehat{V}_t\left(\mathcal{E}_t; (h, r)\right) . \tag{4}$$

推论1。 采样规则 *(4)* 可改写为 $A_{t+1}\left(\mathcal{E}_t\right) = (\widehat{i}^{(t+1)}, \widehat{\ell}^{(t+1)})$，其中 $\widehat{i}^{(t+1)}$ 和 $\widehat{\ell}^{(t+1)}$ 分别是接收第 $(t+1)$ 次仿真重复的设计和上下文，

$$\widehat{\ell}^{(t+1)} \in \arg\min_{\ell=1,\cdots,q} APCS_\ell\left(\mathcal{E}_t\right) , \tag{5}$$

并令 $H^{(t+1)}$ 为包含满足$APCS_{\widehat{\ell}^{(t)}}\left(\mathcal{E}_t\right)$的索引的集合，即对于 $i = 1, \cdots, m,\ j = m+1, \cdots, k$,

$$H^{(t+1)} = \arg\min_{\{i,j\}} \frac{\left(y_{\langle i\rangle_{\widehat{\ell}^{(t+1)}_t}}^{(t)}\left(\boldsymbol{x}_{\widehat{\ell}^{(t+1)}}\right) - y_{\langle j\rangle_{\widehat{\ell}^{(t+1)}_t}}^{(t)}\left(\boldsymbol{x}_{\widehat{\ell}^{(t+1)}}\right)\right)^2}{\left(\sigma_{\langle i\rangle_{\widehat{\ell}^{(t+1)}_t}}^{(t)}\left(\boldsymbol{x}_{\widehat{\ell}^{(t+1)}}\right)\right)^2 + \left(\sigma_{\langle j\rangle_{\widehat{\ell}^{(t+1)}_t}}^{(t)}\left(\boldsymbol{x}_{\widehat{\ell}^{(t+1)}}\right)\right)^2} .$$

那么， $\forall\, h^{(t+1)} \in H^{(t+1)}$,

$$\widehat{i}^{(t+1)} \in \arg\max_{\langle h^{(t+1)}\rangle_{\widehat{\ell}^{(t+1)}_t}} \widehat{V}_t(\mathcal{E}_t; (\langle h^{(t+1)}\rangle_{\widehat{\ell}^{(t+1)}_t}, \widehat{\ell}^{(t+1)})) . \tag{6}$$

注释1。 推论*1*的证明可见附录。根据推论*1*，按照*AOAmc*进行分配后，分配的上下文 $\widehat{\ell}^{(t)}$在所有上下文中实现了最小的 $APCS_\ell\left(\mathcal{E}_t\right)$。在上下文 $\widehat{\ell}^{(t)}$下，按照*AOAmc*进行分配的设计 $\widehat{i}^{(t)}$ 在所有能够实现$APCS_{\widehat{\ell}^{(t)}}\left(\mathcal{E}_t\right)$的设计中，实现了最大的*VFA*一步前瞻值。*AOAmc*重点关注在具有所有上下文中最小近似选择正确概率$\widehat{}$的上下文下，最难以进行比较的设计。推论*1*表明，*AOAmc*无需计算所有设计-情境对的 $V_t\left(\mathcal{E}_t; (h, r)\right)$ 来确定一个分配的设计-情境对，而是在每一步所需的时间复杂度为 $O(kq+k)$，该复杂度与总量有关。

设计-情境对的数量，以及在分配的环境下设计的数量。

AOAmc 具有反映均值-方差权衡的解析形式。所提出的 AOAmc 程序如算法 1 所示。最初的 n_0 次重复试验分配给每个设计-情境对，以估计其未知参数。随后依次向设计-情境对分配额外的重复试验，直至所有重复试验用尽。当 m 的值和设计数量因上下文而异时，AOAmc 也适用；例如，请参见在线附录 A.3 节的补充实验 [49] 。

算法1AOAmc**输入：** k, q, m, n_0, T。

初始化： $t = 0$ 并对每个设计-情境对执行 n_0 次仿真复制。

当 $t < (T - k \times q \times n_0)$ 时，**循环执行：**

更新：根据（1）和（2）更新信息滤过。

分配：根据（5）和（6）确定分配的设计-情境对。

仿真：对分配的设计-情境对执行额外仿真。

$t = t + 1$。**循环**

结束。

下一个结果（如定理 1 所述）表明，随着仿真预算趋于无穷，AOAmc 能够正确识别所有上下文下的前-m 个设计。

定理 1. 所提出的用于选择前-m个上下文相关设计 *(4)* 的抽样策略是一致的，即

$$\lim_{t\to\infty} \widehat{\mathcal{F}}_{\ell t}^m = \mathcal{F}_\ell^m, \ \ a.s., \ \ell = 1, \cdots, q .$$

备注 2. 定理 *1* 的证明见在线附录 *[49]* 的 *A.1* 节。前-m 个上下文相关选择问题的一致性证明（定理 *1*）与前-m 个无上下文选择问题的证明（ *[8]* 中的定理 *1*）有以下不同之处：首先，我们证明，对于固定的上下文，几乎必然会对所有设计进行无限次抽样。其次，我们通过反证法将结论推广到任意上下文，并利用 PCS_W 的定义导出矛盾。

IV. 渐近最优性

除了证明一致性之外，我们还证明了所提出抽样策略的渐近最优性。在本节中，我们首先通过最大化 PFS_W 的大偏差率，推导一般采样分布下的静态渐近最优抽样比例。随后证明，在正态抽样分布下，AOAmc 的渐近抽样比例是渐近最优的。

定义抽样比例 $r_{h\ell} \triangleq t_{h\ell}/T$, $\sum_{h=1}^k \sum_{\ell=1}^q r_{h\ell} = 1$, $r_{h\ell} \geq 0$, 而 $\boldsymbol{r} = (r_{11}, \cdots, r_{1q}, \cdots, r_{k1}, \cdots, r_{kq})$ 是抽样比例向量。在静态优化框架下，假设每个上下文 $\boldsymbol{x}_\ell$ 中的设计 $\langle h\rangle$, $h = 1, \cdots, k$ 已知

具体而言，错误选择概率$_W$计算为错误选择概率$_W$ $=\max_{\ell=1,\cdots,q}\Pr\left\{\bigcup_{i=1}^{m}\bigcup_{j=m+1}^{k}\left(\bar{Y}_{\langle i\rangle\ell}\left(r_{\langle i\rangle\ell}T\right)\le\bar{Y}_{\langle j\rangle\ell}\left(r_{\langle j\rangle\ell}T\right)\right)\right\}$。

将 $\Lambda_{h\ell}(\lambda)\triangleq\log\mathbb{E}(\exp(\lambda\bar{Y}_{h\ell}(r_{h\ell}T)))$ 记为样本均值 $Y_{h\ell}(r_{h\ell}T)$,$h\in\{1,\cdots,k\}$, $\ell\in\{1,\cdots,q\}$ 的累积量生成函数。 $\Lambda_{h\ell}(\cdot)$ 的有效定义域为 $\mathcal{D}_{\Lambda_{h\ell}}\triangleq\{\lambda\in\mathbb{R}:\Lambda_{h\ell}(\lambda)<\infty\}$。定义$\mathcal{F}_{h\ell}\triangleq\left\{d\Lambda_{h\ell}(\lambda)/d\lambda:\lambda\in\mathcal{D}_{\Lambda_i}^{o}\right\}$，其中 $\mathbb{A}^o$ 表示集合 $\mathbb{A}$的内部。我们在分析中作出以下技术假设。

假设1。 $\Lambda_{h\ell}(\lambda)=\lim_{T\to\infty}\frac{1}{T}\Lambda_{h\ell}(T\lambda)$ 对于所有 λ，以及 $0\in\mathcal{D}_{\Lambda_{h\ell}^{o}}$, $h=1,\cdots,k$, $\ell=1,\cdots,q$，均作为扩展实数有定义。

假设2。 $\Lambda_{h\ell}(\lambda)$ 在 $\mathcal{D}_{\Lambda_h}^{o}$ 上严格凸、连续且陡峭，即 $\lim_{T\to\infty}|\Lambda'_{h\ell}(\lambda_T)|=\infty$, $h=1,\cdots,k$, $\ell=1,\cdots,q$，其中 $\{\lambda_T\}\in\mathcal{D}_{\Lambda_h}$ 是一个收敛到 $\mathcal{D}_{\Lambda_h}^{o}$ 边界点的序列。

假设3 $[y_{\langle k\rangle_\ell}\ \boldsymbol{x}_\ell\quad y_{\langle 1\rangle_\ell}\ \boldsymbol{x}_\ell\ ]$。区 $\subset\bigcap_{i=1}^{k}\mathcal{F}_{i\ell}^{o}$ $\ell=1\cdots q$ 间 ()(),,,。

根据 G¨artner-Ellis 定理， [50], 假设1和假设2表明，样本均值 $Y_{h\ell}$ 满足速率函数为 $\Lambda_{h\ell}^{*}(x)=\sup_{\lambda\in\mathbb{R}}\{\lambda x-\Lambda_{h\ell}(\lambda)\}$ 的大偏差原理。假设2表明，对于 $x\in\mathcal{F}_{h\ell}^{o}$， $\Lambda_{h\ell}^{*}(x)$ 严格凸且连续。假设3确保每个设计-情境对的样本均值可以取 $[y_{\langle k\rangle_\ell}\ \boldsymbol{x}_\ell\quad y_{\langle 1\rangle_\ell}\ \boldsymbol{x}_\ell\ ]$, $\{\bar{Y}_{\langle j\rangle\ell}\left(r_{\langle j\rangle\ell}T\right)\ge\bar{Y}_{\langle i\rangle\ell}\left(r_{\langle i\rangle\ell}T\right)\}>0$ $i=1\cdots m$ $j=m+1\cdots k$ ( ) ( ) 中的任意值，并且 Pr ,,,以及 ,,。然后，我们可以使用大偏差范式分析错误选择概率$_W$的收敛速率。

错误选择概率$_W$ 的下界为

$$lb\triangleq\max_{\ell=1,\cdots,q}\max_{\substack{i=1,\cdots,m\\ j=m+1,\cdots,k}}\Pr\left\{\bar{Y}_{\langle i\rangle\ell}\left(r_{\langle i\rangle\ell}T\right)\le\bar{Y}_{\langle j\rangle\ell}\left(r_{\langle j\rangle\ell}T\right)\right\},$$

并且上界为 $m\times(k-m)\times lb$。尽管该上界可能大于 1，但这表明 lb自身的收敛速率刻画了 PFS$_W$的收敛速率。根据关于 $\bar{Y}_{\langle j\rangle\ell}\left(r_{\langle j\rangle\ell}T\right)<\bar{Y}_{\langle i\rangle\ell}\left(r_{\langle i\rangle\ell}T\right)$的 G¨artner-Ellis 定理[50],，存在一个速率函数 $\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell}\ r_{\langle j\rangle\ell}\right)$ $\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell}\ r_{\langle j\rangle\ell}\right)=$，使得，$\overline{\lim}_{T\to\infty}\frac{1}{T}\log\Pr\left\{\bar{Y}_{\langle j\rangle\ell}\left(r_{\langle j\rangle\ell}T\right)\ge\bar{Y}_{\langle i\rangle\ell}\left(r_{\langle i\rangle\ell}T\right)\right\}$。 [8] 进一步表明，$\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell}\ r_{\langle j\rangle\ell}\right)=r_{\langle i\rangle\ell}\Lambda_{\langle i\rangle\ell}^{*}\left(x\left(r_{\langle i\rangle\ell}\ r_{\langle j\rangle\ell}\right)\right)+,,,,$ 其中，$r_{\langle j\rangle\ell}\Lambda_{\langle j\rangle\ell}^{*}\left(x\left(r_{\langle i\rangle\ell}\ r_{\langle j\rangle\ell}\right)\right)$ $x\left(r_{\langle i\rangle\ell}\ r_{\langle j\rangle\ell}\right)$ 是满足 $r_{\langle i\rangle\ell}\frac{d\Lambda_{\langle i\rangle\ell}^{*}(\ x}{dx}\Big|_{x=x(r_{\langle i\rangle\ell},r_{\langle j\rangle\ell}}\)\)$ $+_{()}$ $r_{\langle j\rangle\ell}\frac{d\Lambda_{\langle j\rangle\ell}^{*}\ x}{dx}\Big|_{x=x(r_{\langle i\rangle\ell},r_{\langle j\rangle\ell})}=0$ 的唯一解。因此，PFS$_W$的收敛速率可表示为

$$\lim_{T\to\infty}\frac{1}{T}\log\left(\max_{\ell}\max_{i,j}\Pr\left\{\bar{Y}_{\langle j\rangle\ell}\left(r_{\langle j\rangle\ell}T\right)\ge\bar{Y}_{\langle i\rangle\ell}\left(r_{\langle i\rangle\ell}T\right)\right\}\right)$$
$$=\max_{\ell}\max_{i,j}\lim_{T\to\infty}\frac{1}{T}\log\left(\Pr\left\{\bar{Y}_{\langle j\rangle\ell}\left(r_{\langle j\rangle\ell}T\right)\ge\bar{Y}_{\langle i\rangle\ell}\left(r_{\langle i\rangle\ell}T\right)\right\}\right)$$
$$=-\min_{\ell=1,\cdots,q}\min_{\substack{i=1,\cdots,m\\ j=m+1,\cdots,k}}\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell},r_{\langle j\rangle\ell}\right)\triangleq-\mathcal{R}\left(\boldsymbol{r}\right).$$

因此，选择前 m 个上下文相关设计的 PFS$_W$ 呈指数收敛，其速率函数为 $\boldsymbol{r}$，即 $-\lim_{T\to\infty}\frac{1}{T}\log\text{PFS}_W=\mathcal{R}(\boldsymbol{r})$。注意， $\bar{\mathcal{R}}(\boldsymbol{r})$ 被定义为所有上下文中最困难的成对比较所对应的最慢速率函数。为推导静态优化问题下的渐近最优抽样比例，我们考虑以下优化问题：

$$\begin{aligned}\max\quad&\mathcal{R}(\boldsymbol{r})\\ \text{s.t.}\quad&\sum_{h=1}^{k}\sum_{\ell=1}^{q}r_{\langle h\rangle\ell}=1,\\ &r_{\langle h\rangle\ell}\ge0,\quad h=1,\cdots,k,\ \ell=1,\cdots q\ ,\end{aligned}\tag{7}$$

其中，该约束等价于 $\sum_{h=1}^{k}\sum_{\ell=1}^{q}t_{\langle h\rangle\ell}=T$。(7) 的等价表述可以写为

$$\begin{aligned}\max\quad&z\\ \text{s.t.}\quad&\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell},r_{\langle j\rangle\ell}\right)-z\ge0,\quad i=1,\cdots,m,\\ &\qquad j=m+1,\cdots,k,\ \ell=1,\cdots,q,\\ &\sum_{h=1}^{k}\sum_{\ell=1}^{q}r_{\langle h\rangle\ell}=1,\\ &r_{\langle h\rangle\ell}\ge0,\quad h=1,\cdots,k\ .\end{aligned}\tag{8}$$

$\mathcal{R}_{ij}^{\ell}(\cdot,\cdot)$ 是关于 $\left(r_{\langle i\rangle\ell},r_{\langle j\rangle\ell}\right)$的严格凹连续函数。令 $\widetilde{\mathcal{R}}_{ij}^{\ell}(\boldsymbol{r})=\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell},r_{\langle j\rangle\ell}\right)$，则 $\widetilde{\mathcal{R}}_{ij}^{\ell}(\boldsymbol{r})$ 是关于 $\boldsymbol{r}$ 的凹函数。(7) 中的目标函数$\mathcal{R}(\boldsymbol{r})$ 对于 $\boldsymbol{r}\ge0$ 是凹函数，因为凹函数 $\widetilde{\mathcal{R}}_{ij}^{\ell}(\boldsymbol{r})$ 的最小值同样是凹函数。因此，(8) 是一个凸优化问题。存在解$\boldsymbol{r}=(1/k,\cdots,1/k)$ 和 $z=0$，使得 (8) 中的不等式约束严格成立，因此 Slater 条件对于 (8) 成立。随后，我们可以研究 (8) 的 Karush-Kuhn-Tucker (KKT) 条件，该条件是最优性的充要条件。

引理 1。 在假设 *1-3* 下，*(8)* 的最优解满足，对于 i、$i'=1$、$\cdots$、m、 j、$j'=m+1$、$\cdots$、k、 ℓ、$\ell'=1$，$\cdots$， q

$$\min_{j=m+1,\cdots,k}\mathcal{R}_{ij}^{\ell}(r_{\langle i\rangle\ell}^{*},r_{\langle j\rangle\ell}^{*})=\min_{i'=1,\cdots,m}\mathcal{R}_{i'j'}^{\ell}(r_{\langle i'\rangle\ell}^{*},r_{\langle j'\rangle\ell}^{*})=$$
$$\min_{j=m+1,\cdots,k}\mathcal{R}_{ij}^{\ell'}(r_{\langle i\rangle\ell'}^{*},r_{\langle j\rangle\ell'}^{*})=\min_{i'=1,\cdots,m}\mathcal{R}_{i'j'}^{\ell'}(r_{\langle i'\rangle\ell'}^{*},r_{\langle j'\rangle\ell'}^{*}),\tag{9}$$

where $r_{h\ell}^{*}>0$, $h=1,\cdots,k$, $\ell=1,\cdots,q$ are asymptotically optimal sampling ratios.

证明：首先，我们证明 $r_{h\ell}^{*}>0$。对于 $r_{h\ell}^{*}=1/k$,$h=1,\cdots,k$, $\ell=1,\cdots,q$，我们有 $z>0$。如果存在$r_{\langle h\rangle\ell}^{*}=0$, $h\in\{1,\cdots,k\}$, $\ell\in\{1,\cdots,q\}$，则有 $\mathcal{R}_{hj}^{\ell}\ r_{\langle h\rangle\ell}^{*}\ r_{\langle j\rangle\ell}^{*}=0$ $\mathcal{R}_{ih}^{\ell}\ r_{\langle i\rangle\ell}^{*}\ r_{\langle h\rangle\ell}^{*}=0$ $z=0$ $i=(,)$ 或 $(,)$ $1\cdots m$ $\cdots k$，以及 ,,,, $j=m+1$,,, 这与旨在最大化 z 的目标函数相矛盾。因此，优化问题 (8) 的最优解必须满足 $r_{h\ell}^{*}>0$。根据 KKT 条件，存在乘子 γ 和 $\lambda_{ij}^{\ell}\ge0$，使得对于 $i=1,\cdots,m$, $j=m+1,\cdots,k$, $\ell=1,\cdots,q$,

$$\sum\nolimits_{h=1}^{k}\sum\nolimits_{\ell=1}^{q}r_{\langle h\rangle\ell}^{*}=1\ ,$$

$$\sum_{i=1}^{m}\sum_{j=m+1}^{k}\sum_{\ell=1}^{q}\lambda_{ij}^{\ell}=1\ , \tag{10}$$

$$\sum_{j=m+1}^{k}\lambda_{ij}^{\ell}\left.\frac{\partial\mathcal{R}_{ij}^{\ell}\left(y,r_{\langle j\rangle\ell}^{*}\right)}{\partial y}\right|_{y=r_{\langle i\rangle\ell}^{*}}=\gamma\ , \tag{11}$$

$$\sum_{i=1}^{m}\lambda_{ij}^{\ell}\left.\frac{\partial\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell}^{*},y\right)}{\partial y}\right|_{y=r_{\langle j\rangle\ell}^{*}}=\gamma\ , \tag{12}$$

$$\lambda_{ij}^{\ell}\left[z-\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell}^{*},r_{\langle j\rangle\ell}^{*}\right)\right]=0\ . \tag{13}$$

根据 (10)，我们有 $\max_{i,j,\ell}\lambda_{ij}^{\ell}>0$，$i=1,\cdots,m$，$j=m+1,\cdots,k$，$\ell=1,\cdots,q$。由于 $\frac{\partial\mathcal{R}_{ij}^{\ell}(r_{\langle i\rangle\ell}^{*},y)}{\partial y}\Big|_{y=r_{\langle j\rangle\ell}^{*}}>0$，且 $\frac{\partial\mathcal{R}_{ij}^{\ell}(y,r_{\langle j\rangle\ell}^{*})}{\partial y}\Big|_{y=r_{\langle i\rangle\ell}^{*}}>0$，因此可得 $\gamma>0$。因此，由 (11) 和 (12) 可得，$\max_{j=m+1,\cdots,k}\lambda_{ij}^{\ell}>0$，$i=1,\cdots,m,\ell=1,\cdots,q$ 且 $\max_{i=1,\cdots,m}\lambda_{ij}^{\ell}>0$，$j=m+1,\cdots,k$，$\ell=1,\cdots,q$。接着，由 (13) 得到 $\mathcal{R}_{ij}^{\ell}(r_{\langle i\rangle\ell}^{*},r_{\langle j\rangle\ell}^{*})=z,i=1,\cdots,m,\ell=1,\cdots,q$，且 $\min_{i=1,\cdots,m}\mathcal{R}_{ij}^{\ell}(r_{\langle i\rangle\ell}^{*},r_{\langle j\rangle\ell}^{*})=z,j=m+1,\cdots,k$，$\ell=1,\cdots,q$。综上，定理得证。■

方程（9）针对某个上下文以及不同上下文，调整分配给 top-m 和其他（$k-m$）设计的仿真重复次数。注意，由于（11）和（12）中的求和，KKT 条件中的某些拉格朗日乘子 λ_{ij}^{ℓ} 可能等于 0，因此（9）包含不同情形，具体取决于 $\lambda_{ij}^{\ell}=0$ 的可能数量，即非活跃不等式约束的数量 $\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell},\ r_{\langle j\rangle\ell}\right)-z\geq 0$。情形数量会随着 k和 q 的增大而增加。以 $k=4$、$q=2$和 $m=2$ 为例，（9）变为

$$\begin{aligned}&\min\left\{\mathcal{R}_{13}^{1}(\cdot,\cdot)\ \ \mathcal{R}_{14}^{1}(\cdot,\cdot)\right\}=\min\left\{\mathcal{R}_{23}^{1}(\cdot,\cdot)\ \ \mathcal{R}_{24}^{1}(\cdot,\cdot)\right\}\\&=\min\left\{\mathcal{R}_{13}^{1}(\cdot,\cdot)\ \ \mathcal{R}_{23}^{1}(\cdot,\cdot)\right\}=\min\left\{\mathcal{R}_{14}^{1}(\cdot,\cdot)\ \ \mathcal{R}_{24}^{1}(\cdot,\cdot)\right\}=\\&\min\left\{\mathcal{R}_{13}^{2}(\cdot,\cdot)\ \ \mathcal{R}_{14}^{2}(\cdot,\cdot)\right\}=\min\left\{\mathcal{R}_{23}^{2}(\cdot,\cdot)\ \ \mathcal{R}_{24}^{2}(\cdot,\cdot)\right\}=\\&\min\left\{\mathcal{R}_{13}^{2}(\cdot,\cdot)\ \ \mathcal{R}_{23}^{2}(\cdot,\cdot)\right\}=\min\left\{\mathcal{R}_{14}^{2}(\cdot,\cdot)\ \ \mathcal{R}_{24}^{2}(\cdot,\cdot),\right\}\end{aligned} \tag{14}$$

它考虑 $\min\{\cdot\}$ 函数中的所有可能情形，共包含 49 种情况。与 $q=1$ 相比，这包含的情况数量要多得多，因此相较于无上下文问题，分析也更加复杂。请注意，(9) 是根据互补松弛条件 (13) 推导得出的，而该条件并未完全刻画所有 KKT 条件。具体而言，它无法确保所有拉格朗日乘子$\lambda_{ij}^{\ell}\geq 0$。不过，当 $m=1$ 且$m=k-1$ [9]时，可以精确求出所有拉格朗日乘子，并证明它们均为正， [23]。相反，当 $m\neq 1$且 $m\neq k-1$时，根据 (9) 的某些情形得出的静态抽样比例可能导致 $\lambda_{ij}^{\ell}<0$，从而违反 KKT 条件。这种违反表明，这类抽样比例并非渐近最优。因此，并非 (9) 的所有可能情形都会导出某个 top-m 依赖上下文的选择问题的渐近最优抽样比例。具体而言，对于 top-m 依赖上下文的选择问题，(9) 仅是渐近最优抽样比例的必要条件（尤其是 $m\neq 1$ 和 $m\neq k-1$）。备注 5 中给出了一个采用正态抽样分布的速率函数示例，以支持这些论述。

然后，我们在定理 2 中研究正态抽样分布下渐近最优抽样比例所满足的条件。

定理 2。 在假设 *1-3* 下，对于具有正态底层分布的、用于选择前 m 个依赖上下文的设计的 错误选择概率$_W$，使大偏差率达到最优的渐近最优抽样比例满足：对于 i、$i'=1$、$\cdots$、m、j、$j'=m+1$、$\cdots$、k、ℓ、$\ell'=1$、$\cdots$、q，

$$\begin{aligned}&\min_{j=m+1,\cdots,k}\frac{\left(y_{\langle i\rangle_\ell}(\boldsymbol{x}_\ell)-y_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{\sigma_{\langle i\rangle_\ell}^2(\boldsymbol{x}_\ell)\Big/r_{\langle i\rangle\ell}^{*}+\sigma_{\langle j\rangle_\ell}^2(\boldsymbol{x}_\ell)\Big/r_{\langle j\rangle\ell}^{*}}=\\&\min_{i'=1,\cdots,m}\frac{\left(y_{\langle i'\rangle_\ell}(\boldsymbol{x}_\ell)-y_{\langle j'\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{\sigma_{\langle i'\rangle_\ell}^2(\boldsymbol{x}_\ell)\Big/r_{\langle i'\rangle\ell}^{*}+\sigma_{\langle j'\rangle_\ell}^2(\boldsymbol{x}_\ell)\Big/r_{\langle j'\rangle\ell}^{*}}=\\&\min_{j=m+1,\cdots,k}\frac{\left(y_{\langle i\rangle_{\ell'}}(\boldsymbol{x}_{\ell'})-y_{\langle j\rangle_{\ell'}}(\boldsymbol{x}_{\ell'})\right)^2}{\sigma_{\langle i\rangle_{\ell'}}^2(\boldsymbol{x}_{\ell'})\Big/r_{\langle i\rangle\ell'}^{*}+\sigma_{\langle j\rangle_{\ell'}}^2(\boldsymbol{x}_{\ell'})\Big/r_{\langle j\rangle\ell'}^{*}}=\\&\min_{i'=1,\cdots,m}\frac{\left(y_{\langle i'\rangle_{\ell'}}(\boldsymbol{x}_{\ell'})-y_{\langle j'\rangle_{\ell'}}(\boldsymbol{x}_{\ell'})\right)^2}{\sigma_{\langle i'\rangle_{\ell'}}^2(\boldsymbol{x}_{\ell'})\Big/r_{\langle i'\rangle\ell'}^{*}+\sigma_{\langle j'\rangle_{\ell'}}^2(\boldsymbol{x}_{\ell'})\Big/r_{\langle j'\rangle\ell'}^{*}}\ ,\end{aligned} \tag{15}$$

$$\sum_{i=1}^{m}\frac{\left(r_{\langle i\rangle\ell}^{*}\right)^2}{\sigma_{\langle i\rangle_\ell}^2(\boldsymbol{x}_\ell)}=\sum_{j=m+1}^{k}\frac{\left(r_{\langle j\rangle\ell}^{*}\right)^2}{\sigma_{\langle j\rangle_\ell}^2(\boldsymbol{x}_\ell)},\quad \ell=1,\cdots,q\ . \tag{16}$$

证明：根据正态底层分布下的 [50]，有 $\Lambda_{\langle i\rangle\ell}^{*}(x)=(x-y_{\langle i\rangle_\ell}(\boldsymbol{x}_\ell))^2/(2\sigma_{\langle i\rangle_\ell}^2(\boldsymbol{x}_\ell))$，进而可推得

$$\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell},r_{\langle j\rangle\ell}\right)=\frac{\left(y_{\langle i\rangle_\ell}(\boldsymbol{x}_\ell)-y_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{2\left(\sigma_{\langle i\rangle_\ell}^2(\boldsymbol{x}_\ell)\Big/r_{\langle i\rangle\ell}+\sigma_{\langle j\rangle_\ell}^2(\boldsymbol{x}_\ell)\Big/r_{\langle j\rangle\ell}\right)}\ .$$

将 $\mathcal{R}_{ij}^{\ell}\left(r_{\langle i\rangle\ell},\ r_{\langle j\rangle\ell}\right)$、$i=1$、$\cdots$、$m$1、$\cdots$、$k$、$\ell=1$、$\cdots$、$q$ 代入（9），即可得到（15）成立。然后，对于每个上下文 $\ell\in\{1,\cdots,q\}$，按照定理 3 中的类似证明， [8]（15）中的所有情形均满足（16）。■

若 $q=1$，则（15）和（16）对应于 [8] 在单一上下文下选择前-m 个设计时的渐近最优性条件。（16）确立了针对每个上下文 $\boldsymbol{x}_\ell$，分配给前-m 个设计与其他（$k-m$）设计的仿真预算之间的某种平衡关系 $\ell=1$，$\cdots$，q。尽管（15）和（16）是在正态性假设下推导出的，但（9）对于伯努利分布和指数分布等一般采样分布仍然成立。然而，（9）并不能完全刻画渐近最优性条件，使用经验量估计速率函数可能会导致较大的估计误差。这些计算问题阻碍了仅基于（9）制定抽样程序。

令 $r_\ell^{*}=\sum_{h=1}^{k}r_{h\ell}^{*}>0$，$\sum_{\ell=1}^{q}r_\ell^{*}=1$ 为每个上下文 $\boldsymbol{x}_\ell$，$\ell\in$的渐近最优抽样比例

$\{1,\cdots,q\}$，并令 z^* 为优化问题 (8) 中由渐近最优抽样比例 $\boldsymbol{r}^* \triangleq (r_{11}^*,\ \cdots r_{1q}^*,\ \cdots,\ r_{k1}^*,\ \cdots,\ r_{kq}^*) > 0$ 所达到的目标函数最优值。定义 $\alpha_{h\ell}^* > 0,\ h = 1,\ \cdots,\ k,$ $\ell \in \{1,\cdots,q\}$，其中 $\sum_{h=1}^{k} \alpha_{h\ell}^* = 1$ 是相应的无上下文 top-m 选择问题的渐近最优抽样比例。记 $\boldsymbol{\alpha}_\ell^* \triangleq (\alpha_{1\ell}^*,\ \cdots,\ \alpha_{k\ell}^*) > 0$，并令 $z_\ell(\boldsymbol{\alpha}_\ell^*)$ 为由 $\boldsymbol{\alpha}_\ell^*$ 实现的、用于选择 top-m 无上下文设计的最优大偏差率。根据在正态抽样分布下于 [8]，建立的结果，$\boldsymbol{\alpha}_\ell^*$ 满足

$$\begin{aligned}&\min_{j=m+1,\cdots,k} \frac{\left(y_{\langle i\rangle_\ell}(\boldsymbol{x}_\ell) - y_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{\sigma^2_{\langle i\rangle_\ell}(\boldsymbol{x}_\ell)\Big/\alpha^*_{\langle i\rangle\ell} + \sigma^2_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\Big/\alpha^*_{\langle j\rangle\ell}} \\ &= \min_{i'=1,\cdots,m} \frac{\left(y_{\langle i'\rangle_\ell}(\boldsymbol{x}_\ell) - y_{\langle j'\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{\sigma^2_{\langle i'\rangle_\ell}(\boldsymbol{x}_\ell)\Big/\alpha^*_{\langle i'\rangle\ell} + \sigma^2_{\langle j'\rangle_\ell}(\boldsymbol{x}_\ell)\Big/\alpha^*_{\langle j'\rangle\ell}},\end{aligned} \tag{17}$$

且 $z_\ell(\boldsymbol{\alpha}_\ell^*) = \min_{j=m+1,\ \cdots,\ k}\left(\frac{y_{\langle i\rangle_\ell}\boldsymbol{x}_\ell - y_{\langle j\rangle_\ell}\boldsymbol{x}_\ell}{\sigma^2_{\langle i\rangle_\ell}\boldsymbol{x}_\ell\Big/\alpha^*_{\langle i\rangle\ell} + \sigma^2_{\langle j\rangle_\ell}\boldsymbol{x}_\ell}\right)^2$。此外，对于任何满足 (17) 的 $\widetilde{\boldsymbol{\alpha}}_\ell^* > 0$，$z_\ell(\boldsymbol{\alpha}_\ell^*) > z_\ell\left(\widetilde{\boldsymbol{\alpha}}_\ell^*\right)$。以下推论阐明了在正态抽样分布下，具有 q 个上下文的前 m 项上下文相关选择问题的渐近最优抽样比例，与 q 个相互独立的前 m 项无上下文选择问题之间的联系。

推论 2。在假设 *1-3* 和正态性假设下，我们有 $\alpha_{h\ell}^* = r_{h\ell}^*/r_\ell^*,\ h = 1,\cdots,k,\ \ell = 1,\cdots,q$，其中

$$r_\ell^* = \frac{1/z_\ell\left(\boldsymbol{\alpha}_\ell^*\right)}{\sum_{\widetilde{\ell}=1}^{q} 1\Big/z_{\widetilde{\ell}}\left(\alpha_{\widetilde{\ell}}^*\right)},\quad z^* = \frac{1}{\sum_{\widetilde{\ell}=1}^{q} 1\Big/z_{\widetilde{\ell}}\left(\alpha_{\widetilde{\ell}}^*\right)}\ .$$

注 3. 推论 *2* 的证明见附录。根据推论 *2*，在底层分布为正态分布的前-m 个上下文相关选择问题中，每个上下文的渐近最优抽样比例与选择前-m 个无上下文设计时相应的最优错误选择概率大偏差率成反比。此外，PFS_W的最优大偏差率与 q 个相互独立的前-m 个无上下文选择问题的 PFS 最优大偏差率相关。推论 *2* 提供了一种计算具有 q 个上下文的前-m 个上下文相关选择问题的渐近最优抽样比例的方法：分别求解 q 个相互独立的前-m个无上下文选择问题的渐近最优抽样比例。

以下定理表明，在正态性假设下，随着仿真预算趋于无穷大，AOAmc 的抽样比例可以达到渐近最优抽样比例。

定理 3。假设 *(15)* 和 *(16)* 确定了唯一解。在假设 *1-3* 和正态性假设下，当 $t \to \infty$ 时，采用所提出的 *AOAmc* 程序，每个设计-情境对的抽样比例都将依次达到渐近最优抽样比例，该比例可优化 PFS_W 的大偏差率，即，

$$\lim_{t\to\infty} r_{\langle h\rangle\ell}^{(t)} = r_{\langle h\rangle\ell}^*,\ a.s.,\quad h = 1,\cdots,k,\ \ell = 1,\cdots,q\ ,$$

其中 $r_{h\ell}^{(t)} = t_{h\ell}/t$、$\sum_{h=1}^{k}\sum_{\ell=1}^{q} r_{h\ell}^{(t)} = 1$、$r_{h\ell}^* > 0$，且 $\boldsymbol{r}^*$满足（*15*）和（*16*）。

注释 4。定理 *3* 的证明相当技术性且篇幅较长，详见在线附录 *[49]* 的 *A.1* 节。下面概述依赖上下文的前-m 选择问题（定理 *3*）与无上下文的前-m 选择问题（ *[8]*中的定理 *4*）在证明渐近最优性方面的关键技术差异。首先，在证明收敛比例为正并满足（*15*）时，我们采用反证法，并依赖 PCS_W 的定义来导出矛盾。其次，在证明收敛比例满足（*16*）时，由于需要考虑所有设计-情境对的收敛比例，隐函数集合和方程组的构造更加复杂。对于依赖上下文的前-m 选择问题而言，假设（*15*）和（*16*）的解具有唯一性至关重要，因为并非所有将（*15*）与（*16*）结合得到的情形都会导出渐近最优抽样比例。当 $m = 1$ 时，这一困难并不存在，如 *[9], [21],* 和 *[23]*所示。

备注 5。接下来，我们给出一个示例，以说明 *(15)* 和 *(16)* 的解可能并不唯一。以 $k = 4$、$q = 2$ 和 $m = 2$ 为例。所有设计-上下文对的平均表现如下：

$$\begin{aligned}\boldsymbol{y} &= \left\{y_{\langle 1\rangle_1},\cdots,y_{\langle 4\rangle_1},y_{\langle 1\rangle_2},\cdots,y_{\langle 4\rangle_2}\right\} \\ &= \{6.7,5.9,3.5,1.9,9.6,8.2,5.4,4.3\}\ ,\end{aligned}$$

具有方差：

$$\begin{aligned}\boldsymbol{\sigma}^2 &= \{\sigma^2_{\langle 1\rangle_1},\cdots,\sigma^2_{\langle 4\rangle_1},\sigma^2_{\langle 1\rangle_2},\cdots,\sigma^2_{\langle 4\rangle_2}\} \\ &= \{3.6,1.1,4.4,4.7,3.7,8.0,6.5,4.6\}\ .\end{aligned}$$

(15) 和 *(16)* 的四个解及其对应的大偏差率 PFS_W 如表 *I* 所示。表 *I* 中的解 *1-4* 分别对应 $\mathcal{R}_{13}^1 = \mathcal{R}_{14}^1 = \mathcal{R}_{24}^1 = \mathcal{R}_{13}^2 = \mathcal{R}_{14}^2 = \mathcal{R}_{24}^2 = 0.04748$、$\mathcal{R}_{13}^1 = \mathcal{R}_{14}^1 = \mathcal{R}_{24}^1 = \mathcal{R}_{13}^2 = \mathcal{R}_{23}^2 = \mathcal{R}_{24}^2 = 0.06081$、$\mathcal{R}_{13}^1 = \mathcal{R}_{23}^1 = \mathcal{R}_{24}^1 = \mathcal{R}_{13}^2 = \mathcal{R}_{14}^2 = \mathcal{R}_{24}^2 = 0.05382$ 和 $\mathcal{R}_{13}^1 = \mathcal{R}_{23}^1 = \mathcal{R}_{24}^1 = \mathcal{R}_{13}^2 = \mathcal{R}_{23}^2 = \mathcal{R}_{24}^2 = 0.07163$。因此，根据 $\boldsymbol{r}^*$ 的定义，解 *4* 表示此示例中的渐近最优抽样比例，而其他三个解虽然满足 *(15)* 和 *(16)*，却是渐近非最优解。在此示例中，*(14)* 中 *49* 种可能情况里，只有 *4* 种情况能得到解析解，而仅有一种情况得出渐近最优抽样比例。解 *1-3* 中对应的拉格朗日乘子为负值，这进一步说明了这些解为何渐近非最优。例如，在解 *1* 中，$\lambda_{14}^1 = -0.4058$ 和 $\lambda_{14}^2 = -0.4564$；在解 *2* 中，$\lambda_{14}^1 = -0.4057$；在解 *3* 中，$\lambda_{14}^2 = -1.3297$。为了在每个上下文中将排名前2 的设计与其他 *2* 个设计区分开来，设计-情境对 $\langle 2\rangle_1$、$\langle 3\rangle_1$、$\langle 2\rangle_2$ 和 $\langle 3\rangle_2$ 往往会获得更多仿真

重复试验。然而，表 *I* 显示，在三个渐近非最优解中，分配给四个设计-情境对中的某些对的仿真重复次数过多。例如，在解 *1* 中，仿真重复次数过多地分配给了上述所有四个设计-情境对；在解 *2* 中，仿真重复次数过多地分配给了 $\langle 2\rangle_1$ 和 $\langle 3\rangle_1$；在解 *3* 中，仿真重复次数过多地分配给了 $\langle 2\rangle_2$ 和 $\langle 3\rangle_2$。与排名前 m 的无上下文选择问题相比，排名前 m 的依赖上下文的选择问题中潜在的渐近非最优解更多（尤其是 $m \neq 1$ 和 $m \neq k-1$）。*(15)* 和 *(16)* 的解可能不唯一，这给刻画针对排名前 m 的依赖上下文选择问题（尤其是 $m \neq 1$ 和 $m \neq k-1$）所导出的任何抽样策略的渐近最优性带来了挑战。然而，这一挑战不会影响所提出的 *AOAmc* 的表现，因为它并非基于 *(15)* 和 *(16)* 推导而来。在在线附录的 *A.2* 节 *[49]*, 中，我们通过仿真表明，基于此示例，随着仿真预算增加，所提出的 *AOAmc* 往往会达到渐近最优抽样比例，而非渐近非最优的抽样比例。如果 *(15)* 和 *(16)* 的解不唯一，则可以通过验证 *AOAmc* 的渐近解是否满足引理 *1* 证明中的 *KKT* 条件，来数值检验其渐近最优性。

表 I 备注 5 中示例里 (15) 和 (16) 的四个解

抽样比例	解 1	解 2	解 3	解 4
$r_{\langle 1\rangle 1}$	0.04210	0.05393	0.05930	**0.07892**
$r_{\langle 2\rangle 1}$	0.09680	0.12399	0.05762	**0.07668**
$r_{\langle 3\rangle 1}$	0.19700	0.25233	0.12785	**0.17015**
$r_{\langle 4\rangle 1}$	0.02991	0.03831	0.03628	**0.04829**
$r_{\langle 1\rangle 2}$	0.02285	0.03241	0.02590	**0.03817**
$r_{\langle 2\rangle 2}$	0.30417	0.23814	0.34483	**0.28050**
$r_{\langle 3\rangle 2}$	0.27281	0.21059	0.30927	**0.24805**
$r_{\langle 4\rangle 2}$	0.03436	0.05030	0.03895	**0.05924**
z	0.04748	0.06081	0.05382	**0.07163**

V. 数值实验

在本节中，我们进行数值实验，以测试不同采样策略在选择前-m 个依赖上下文的设计时的表现。我们将提出的 AOAmc 与均匀分配（EA）、在上下文之间均匀分配以选择前-m 个设计的最优计算预算分配（E-OCBAm）、在上下文之间均匀分配以选择前-m 个设计的渐近最优分配策略（E-AOAm）、线性多臂老虎机的前-m 个识别算法（m-LinGapE），以及用于选择前-m个依赖上下文的设计的平衡最优大偏差率函数（BOLDmc）进行比较。

具体而言，均匀分配会将仿真复制均匀分配给每个设计-情境对，以估计其性能，即大约为 $t_{h\ell} = T/(k\times q)$, $h = 1, \cdots, k$, $\ell = 1, \cdots, q$。OCBAm 程序 [5] 旨在识别单个情境中的前 m 个设计，这需要一个分离参数来区分前 m 个设计与其余 $(k-m)$ 个设计。我们对其稍作改进，以用于前 m 个依赖上下文的选择问题：在不同上下文之间均匀分配仿真复制，并在每个上下文中运行 OCBAm 程序。在 E-OCBAm 的实现中，我们将其与启发式“最匮乏优先”序贯规则 [51] 结合使用，并采用 [8], 提出的分离参数；已有研究表明，该参数能够提升 OCBAm 程序的性能。E-AOAm 是 AOAm 程序 [8],的扩展，它在不同上下文之间均匀分配仿真复制，并针对每个上下文运行 AOAm 程序。 [42], 提出的 m-LinGapE 首先找到一个设计-情境对 (i,ℓ)，使得 $\min_\ell \max_{i,j} x_\ell^T(\widehat{\theta}_i^\lambda,_t - \widehat{\theta}_j^\lambda,_t)^{\sim}$，然后找到与 $($ i ℓ $\ell\,\widetilde{\ell}$ i t $+,)$ $C_{t\delta}\sqrt{(\boldsymbol{x}_\ell - \boldsymbol{x}_{\widetilde{\ell}})^T \widetilde{\sigma}^2 (V_t^\lambda)^{-1}(\boldsymbol{x}_\ell - \boldsymbol{x}_{\widetilde{\ell}})}$ $\widehat{\theta}_h^\lambda$ 竞争最激烈的设计-情境对 (j,ℓ)，使得 $\min$, $\max$,$_j$ $\boldsymbol{x}_\ell^T\widehat{\theta}_i^\lambda$, $_t$ $-$ $\boldsymbol{x}_{\widetilde{\ell}}^T\widehat{\theta}_j^\lambda$, , where , t $=$ $(V_t^\lambda)^{-1}\sum_{\ell'=1}^q \bar{Y}_{h\,t}$ $\boldsymbol{x}_{\ell'}$ $\boldsymbol{x}_{\ell'}$ $V_t^\lambda = \lambda I_d + \sum_{\ell'=1}^q t_{\ell'}\boldsymbol{x}_{\ell'}\boldsymbol{x}_{\ell'}^T$, ( ) , ,$t_{\ell'} = \sum_{\ell'=1}^q t_{h\ell'}$, $C_{t\delta} = \sqrt{2\ln((\ln t+1)/\delta)}$，且 $\widetilde{\sigma}^2$ 为上下文值的方差。随后按序确定一个分配的设计-情境对，使得 $\max_{\ell''} C_{t\delta}\sqrt{\boldsymbol{x}_{\ell''}^T\widetilde{\sigma}^2(V_t^\lambda)^{-1}\boldsymbol{x}_{\ell''}}$, $\ell'' = \ell, \widetilde{\ell}$。最后，我们考虑将 [9] 中的 CR&S 程序扩展至前 m 个依赖上下文的选择问题。CR&S 程序在概念上与 [22], 中的 BOLD 程序相似，因此我们将扩展后的程序记为 BOLDmc。BOLDmc 分配仿真复制，以便按序平衡渐近最优性条件 (15) 和 (16) 中的各个方程。在每一步中，它首先找到满足条件的设计-情境对 $(\langle i\rangle_{\ell^* t}, \langle j\rangle_{\ell^* t})$

$$\arg\min_{\ell=1,\cdots,q}\ \min_{\substack{i=1,\cdots,m\\ j=m+1,\cdots,k}} \frac{\left(\bar{Y}_{\langle i\rangle_{\ell t}\ell} - \bar{Y}_{\langle j\rangle_{\ell t}\ell}\right)^2}{\widehat{\sigma}^2_{\langle i\rangle_{\ell t}}\Big/ t_{\langle i\rangle_{\ell t}\ell} + \widehat{\sigma}^2_{\langle j\rangle_{\ell t}}\Big/ t_{\langle j\rangle_{\ell t}\ell}}, \quad (18)$$

其中，$\widehat{\sigma}_h^2$ 是样本方差，随后在满足 $\sum_{i=1}^m \frac{t^2_{\langle i\rangle_{\ell^* t}\ell^* t}}{\widehat{\sigma}^2_{\langle i\rangle_{\ell^* t}}} < \sum_{j=m+1}^k \frac{t^2_{\langle j\rangle_{\ell^* t}\ell^* t}}{\widehat{\sigma}^2_{\langle j\rangle_{\ell^* t}}}$ 时将一次仿真重复分配给设计-情境对（$\langle i\rangle_{\ell^* t}$, ℓ^*）；否则，分配给（$\langle j\rangle_{\ell^* t}$, ℓ^*）。

均匀分配不利用任何样本信息。E-OCBAm 和 E-AOAm 均通过求解 q 个相互独立的 Top-m 无上下文选择问题推导得出，这些问题仅利用设计之间的样本信息，而不考虑上下文之间的信息。m-LinGapE 假设每个设计的性能与上下文之间存在线性依赖关系，并将上下文参数作为辅助信息。BOLDmc 利用设计-情境对的样本均值和方差信息。根据推论1，当先验无信息时，BOLDmc 中的公式（18）与公式（5）右侧相同，这表明在 BOLDmc 和 AOAmc 下，每一步分配的上下文都是相同的。BOLDmc 与 AOAmc 的主要区别在于它们确定每一步分配的设计的方法。前者侧重于渐近最优性，而后者则同时追求理想的单步最优性和渐近最优性。在在线附录的 A.2 节 [49], 中，我们基于备注 5 中的示例，通过仿真预算的增加，实证检验 BOLDmc 的抽样比例变化趋势。数值结果表明，与备注 5 中 AOAmc 的结果类似，BOLDmc 的抽样比例也趋于收敛于

随着仿真预算的增加，渐近最优抽样比例，而非渐近非最优抽样比例。然而，考虑到 (15) 和 (16) 仅作为针对前-m上下文相关选择问题而设计的固定预算抽样策略实现渐近最优性的必要条件（特别是，$m \neq 1$和$m \neq k-1$），仅基于该必要条件推导 BOLDmc 的合理性可能存在问题。此外，尽管在备注 5 的测试示例中，BOLDmc 趋于实现渐近最优抽样比例——这是一个简单示例，因为不同的设计-情境对易于区分——但 (15) 和 (16) 的解存在非唯一性，可能会导致 BOLDmc 在更复杂的问题设置中产生误导性结果。确定是否存在某些问题设置，使得 BOLDmc 的抽样比例收敛于渐近非最优抽样比例，仍是一个悬而未决的问题。相比之下，AOAmc 的推导不依赖于 (15) 和 (16)，但仍可证明其具有理想的渐近最优性。

在所有数值实验中，我们将每个设计-情境对的初始仿真复制次数设置为 $n_0 = 10$ 。对于 m-LinGapE，上下文参数 $\boldsymbol{x}_\ell$ 被设置为由 $\boldsymbol{x}_\ell \sim N\left(5, 1^2\right)$生成的单维变量。m-LinGapE 涉及的其他参数被指定为 $\delta = 0.05$和 $\lambda = \tilde{\sigma}/20$。在在线附录的实验 A.3.3 中[49]，，我们通过实证证明，当仿真预算量较小时，每步分配一次仿真复制通常会使 E-OCBAm 和 BOLDmc 获得更好的性能。此外，两种抽样策略在增量仿真复制次数方面均表现出稳健性。实验 1 至 3（以及在线附录中的实验 A.3.1 和 A.3.2 [49]）包含合成示例。为了检验 AOAmc 在不同参数设置下性能的稳健性，我们在每次独立的宏观实验中，从正态先验中随机生成每个设计-情境对的性能。每种采样策略的统计效率通过 IPCS$_W$进行衡量。示例 4 和 5 中每个设计-情境对的性能是确定性的，不同采样策略的统计效率则使用经典选择正确概率进行衡量$_W$。全部

性能指标通过100，000次独立宏观实验进行估计。用于实施数值实验的代码可在 https://github.com/gongbozhang-pku/Context-Dependent-Selection 找到。

实验*1*： 10×5 设计-情境对。我们在一个包含10个竞争设计和5种情境的合成示例中测试采样策略。我们考虑在每种情境下选出排名前3的设计。在每次宏观实验中，每个设计-情境对的表现由$y_h(\boldsymbol{x}_\ell) \sim N\left(0, 6^2\right)$, $h = 1, \cdots, 10$, $\ell = 1, \cdots, 5$生成。仿真重复独立地从正态分布 $N\left(y_h(\boldsymbol{x}_\ell)\ \sigma_h^2(\boldsymbol{x}_\ell)\right)$中抽取，其中 $\sigma_h(\boldsymbol{x}_\ell) = 6$。仿真预算为 $T = 2500$。真实均值生成分布的标准差衡量其离散程度。在实验1中，不同竞争设计-情境对之间的真实均值差异，与仿真重复的标准差处于同一数量级。从统计学角度来看，这表明真实均值之间的差异会相对较大，因此实验1可归类为高置信度情境。

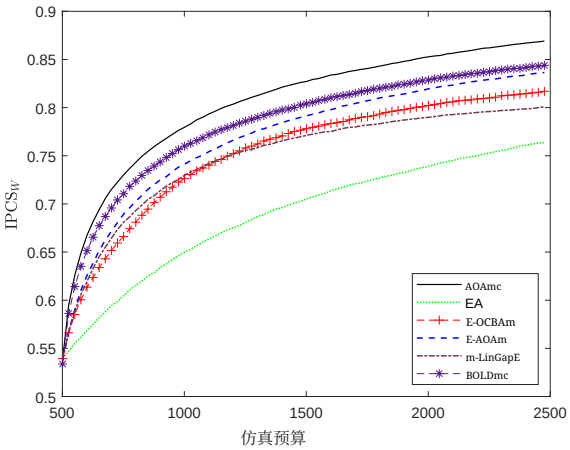


图2. 实验1中六种采样策略的IPCS$_W$。

从图2可以看出，均匀分配在所有采样策略中表现最差。m-LinGapE和E-OCBAm在初期表现相当，但随着仿真预算增加，后者逐渐超过前者。E-AOAm优于E-OCBAm和m-LinGapE；随着仿真预算增加，E-AOAm与BOLDmc之间的差距逐渐缩小。BOLDmc优于除AOAmc之外的所有采样策略，而AOAmc在所有采样策略中表现最佳。与E-AOAm相比，AOAmc的性能提升表明，对于依赖上下文的选择问题，求解 q 无上下文选择问题的效率较低，尤其是在仿真成本高昂时。AOAmc仅需1950次仿真重复即可达到IPCS$_W = 0.85$，而其他采样策略需要超过2500次仿真重复才能达到相同的IPCS$_W$ 水平，即AOAmc将仿真重复的消耗降低了22.0%以上。AOAmc的性能提升可归因于随机控制框架，该框架在有限样本预算约束下制定序贯采样决策。此外，AOAmc兼具良好的单步最优性和渐近最优性，能够充分利用设计-情境对之间的样本信息。

为展示参数 m 对不同采样策略的影响，表 II 报告了六种采样策略在五个不同 m 值下的 IPCS$_W$ 值：$m = 1$, 3, 5, 7, 9, 当$T = 2500$。从表 II 可以观察到，在所有 m 值下，AOAmc 的表现始终优于其他采样策略，展现出其稳健性。

实验*2*： 10×5 设计-情境对。 在一个包含10个设计和5个上下文的合成示例中测试采样策略。数值设置与实验1相同，但每次宏观实验中，每个设计-情境对的表现由 $y_h(\boldsymbol{x}_\ell) \sim N(0, (\sqrt{3.6})^2)$,$h = 1, \cdots, 10$, $\ell = 1, \cdots, 5$ 生成。在实验2中，从统计学角度来看，不同竞争设计-情境对之间的真实均值差异

表 IIIPCS$_W$ 六种采样策略在不同 m 值下的

m	1	3	5	7	9
AOAmc	0.7838	0.8699	0.8335	0.8712	0.9469
EA	0.5320	0.7645	0.7269	0.7632	0.8800
E-OCBAm	0.6872	0.8176	0.7834	0.8193	0.9259
E-AOAm	0.6968	0.8373	0.8039	0.8360	0.9295
m-LinGapE	0.6555	0.8011	0.7607	0.8055	0.9183
BOLDmc	0.7814	0.8449	0.8114	0.8446	0.9454

相较于仿真重复的标准差会比较小。因此，实验2的数值设置可归类为低置信度情景。

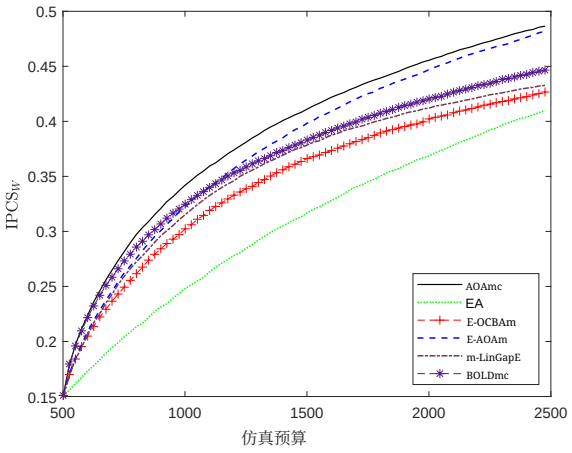


图3. 实验2中六种抽样策略的选择正确概率$_W$。

从图3中可以看出，EA 的 IPCS$_W$ 增长较为缓慢，随着仿真预算的增加，EA 与 E-OCBAm 之间的差距逐渐缩小。E-AOAm、E-OCBAm 和 m-LinGapE 在初始阶段的表现相近；然而，随着仿真预算的增加，m-LinGapE 的表现逐渐优于 E-OCBAm，而 E-AOAm 在三种采样策略中表现最佳。BOLDmc 在初始阶段优于 E-AOAm，但随着仿真预算的增加，后者的表现逐渐超过前者。AOAmc 在初始阶段略优于 BOLDmc，且随着仿真预算的增加，两者之间的差距逐渐扩大。总体而言，AOAmc 优于所有其他参与比较的采样策略。AOAmc 仅需 1930 次仿真重复即可达到 IPCS$_W = 0.45$，而其他采样策略则需要超过 2040 次仿真重复才能达到相同的 IPCS$_W$ 水平，即 AOAmc 可将仿真重复的消耗降低 5.4% 以上。当不同设计-情境对难以区分时，AOAmc 所表现出的稳健性进一步凸显了其优越性及其在实际应用中的潜力。

比较图 2 和图 3 可以观察到，BOLDmc 和 E-OCBAm 在低置信度情形下的性能均有所下降。这凸显了序贯采样策略的渐近性质与有限样本行为之间的差异。具体而言，这一观察从经验上表明，在渐近最优性条件中，用样本估计值替代未知均值和方差所得到的序贯采样策略，其有限样本估计并不能保证良好的有限样本性能。

实验 *3*： 50×10 设计-情境对。 在一个规模更大的合成示例中测试了这些采样策略，该示例包含 50 个设计和 10 个上下文。我们考虑选择排名前列的——

每个上下文有 5 个设计。在每次宏观实验中，每个设计-情境对的性能由 $y_h(\boldsymbol{x}_\ell) \sim N\left(30, 10^2\right)$, $h = 1, \cdots, 50$, $\ell = 1, \cdots, 10$ 生成。对于设计 h和上下文 ℓ，仿真复制独立地取自正态分布 $N\left(y_h\left(\boldsymbol{x}_\ell\right)\ \sigma_h^2\left(\boldsymbol{x}_\ell\right)\right)$，其中$\sigma_h\left(\boldsymbol{x}_\ell\right) \sim U(4,6)$。仿真预算为 $T = 10000$。在实验 3 中，每次独立宏观实验都会随机生成仿真复制的方差，而在实验 1中，该方差在所有独立宏观实验中保持不变。实验3的数值设置可归类为高置信度情景。

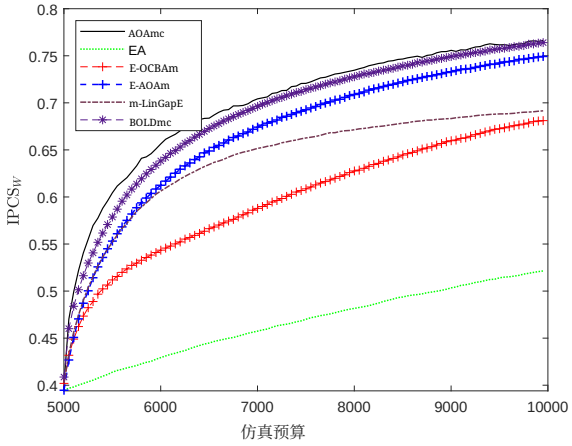


图4. 实验3中六种采样策略的IPCS$_W$。

在图 4 中可以看到，当设计和上下文数量较多时，EA 的 IPCS$_W$ 增长缓慢。E-OCBAm 的表现优于 EA，并随着仿真预算增加而逐渐赶上 m-LinGapE。起初，E-AOAm 和 m-LinGapE 的表现相当；随着仿真预算增加，前者的 IPCS$_W$ 增长速度快于后者。BOLDmc 的表现优于除 AOAmc 以外的所有对比采样策略，并且随着仿真预算增加，BOLDmc 与 AOAmc 之间的差距逐渐缩小。AOAmc 优于其他采样策略。AOAmc 仅需 6915 次仿真重复即可达到 IPCS$_W = 0.7$，而其他采样策略则需要超过 7095 次仿真重复才能达到相同的 IPCS$_W$ 水平；也就是说，AOAmc 可将仿真重复的消耗减少 2.5% 以上。AOAmc 的稳健性凸显了其优越性，以及其在设计和上下文数量较多的场景中用于实际应用的潜力。

实验*4*：蜜罐欺骗博弈示例。具有无线连接的设备容易受到恶意黑客攻击。在网络中部署蜜罐是一种防御网络攻击的欺骗手段。蜜罐会诱使攻击者偏离目标，并帮助防御者了解攻击者使用的黑客技术。在此示例中，我们使用博弈论来刻画防御者与对手之间的策略互动——

在启用蜜罐的网络中，双方采取的行动。我们考虑一个包含5个节点和7条边的网络，如图5所示，其中每个节点代表与某台设备相关联的一个漏洞，而每条边连接两个网络漏洞。攻击者的确切位置对防御者而言未知，并且攻击者在发起攻击前不会收集任何网络信息。因此，行动组合的奖励事先未知。仿真模型的输入包括设计参数（防御行动）和上下文参数（攻击行动）。我们的目标是为设计者确定一些防御行动，以保护网络免受已知攻击类别的攻击。[52] 和 [53]中也研究了这一示例。奖励矩阵中的参数基于 [54]。

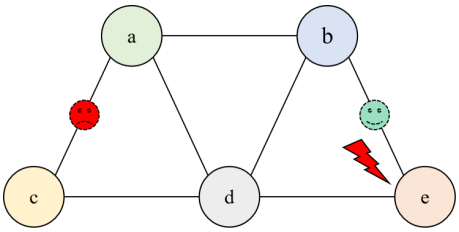


图5. 实验4中的启用蜜罐的网络。如果被攻击的节点是 e，且蜜罐部署在边 be上，则防御者获得捕获收益；如果蜜罐部署在边 ac上，则攻击者获得成功攻击收益。

该策略博弈被建模为一个双人零和静态博弈（$\mathcal{N}$，$\mathcal{A}$，$\mathcal{R}$），其中 $\mathcal{N}$涉及一名防御者和一名攻击者，$\mathcal{A}$ 包含可用博弈行动，而 $\mathcal{R}$ 是由两名参与者采取的行动对（a_i，a_j）所确定的奖励矩阵。防御行动包括决定将蜜罐部署在哪条边上，攻击行动则包括决定攻击哪个节点。此外，考虑到对手行动存在不确定性，防御者可以选择不分配行动，攻击者也可以选择不攻击行动，以避免可能的成本损失。

蜜罐分配成本记为 P_c，攻击成本记为 A_c。如果防御者在连接被攻击节点的边上部署蜜罐，则会获得捕获收益，记为 Cap。否则，攻击者会获得成功攻击收益，记为 Att。收益根据被攻击节点的重要性进行加权，以考虑网络中不同节点的重要性差异。具体而言，防御者的未知奖励表示为 $\mathcal{R}_{ij} = y_{ij} + \epsilon_{ij}, i = 1, \cdots, k, j = 1, \cdots, q$，其中 ϵ_{ij} 服从正态分布，并且

$$y_{ij}(a_i, a_j) = \begin{cases} -P_c + A_c + Cap * w_u & , \ a_i = e_{hu}, a_j = u, \ h \neq u \\ -P_c + A_c - Att * w_u & , \ a_i = e_{h\ell}, a_j = u, \ h, \ell \neq u \\ -P_c & , \ a_i = e_{h\ell}, a_j = 0 \\ A_c - Att * w_u & , \ a_i = 0, a_j = u \\ 0 & , \ a_i = 0, a_j = 0 \end{cases},$$

其中 h 和 ℓ 是连接节点 u 的边，$a_i = 0$ 表示防御者的不分配行动，$a_j = 0$ 表示攻击者的不攻击行动。在此示例中，有 $k = 8$ 个防御行动和 $q = 6$ 个攻击行动。对于每个攻击行动，我们考虑选择奖励最高的前 3 个防御行动。参数设置如下：$P_c = (5, 1, 6, 7, 3, 4)$，$A_c = 4$，$Cap = 10$，$Att = 10$，$w = (0.2401, 0.2669, 0.0374, 0.2692, 0.1864)$，and $\epsilon_{ij} \overset{i.i.d.}{\sim} N(0, 3^2)$。真实奖励矩阵如表 III 所示。仿真预算为 $T = 3000$。

表 III 蜜罐欺骗博弈示例中由元素 y_{ij} (a_i, a_j) 构成的奖励矩阵。

y_{ij}	a	b	c	d	e	无攻击
e_{ab}	1.401	1.669	-1.374	-3.692	-2.864	-5
e_{ac}	5.401	0.331	3.374	0.308	1.136	-1
e_{ad}	-5.265	-4.669	-2.374	0.692	-3.864	-6
e_{bd}	-5.401	-0.331	-3.374	-0.308	-4.864	-7
e_{be}	-1.401	3.669	0.626	-1.692	2.864	-3
e_{cd}	-0.401	-0.669	2.374	4.692	0.136	-2
e_{de}	-2.401	-2.669	-0.374	2.692	1.864	-4
无需分配	1.599	1.331	3.626	1.308	2.136	0

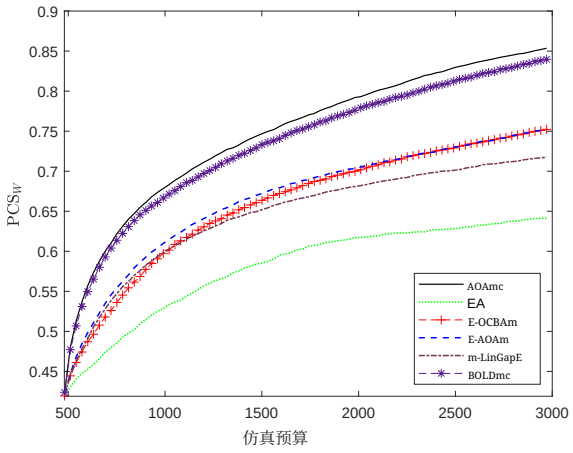
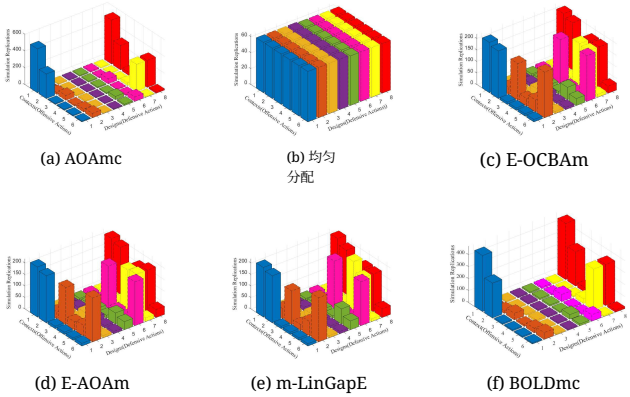


图 6。蜜罐欺骗博弈示例中六种采样策略的选择正确概率$_W$。



(a) AOAmc (b) 均匀分配 (c) E-OCBAm

(d) E-AOAm (e) m-LinGapE (f) BOLDmc

图7. 蜜罐欺骗博弈示例中，各采样策略分配给每个设计-上下文对的仿真重复次数：(a) AOAmc；(b) 均匀分配；(c) E-OCBAm；(d) E-AOAm；(e) m-LinGapE；(f) BOLDmc。

在图6中可以观察到，均匀分配在所有抽样策略中表现最差。E-AOAm 起初略优于 E-OCBAm，随着仿真预算增加，两者之间的差距逐渐缩小。

m-LinGapE 起初略优于 E-OCBAm，但随着仿真预算增加，后者逐渐超过前者。BOLDmc 和 AOAmc 起初表现相近，但随着仿真预算增加，前者的提升速度略慢于后者。AOAmc 的表现优于其他所有对比抽样策略。图7展示了各抽样策略分配给每个设计-情境对的仿真重复次数，该数据取自 100,000 次独立宏观实验的平均值。从图7可以看出，E-OCBAm、E-AOAm 和 m-LinGapE 将更多仿真重复分配给的设计-情境对相似。尽管 AOAmc 和 BOLDmc 分配更多重复试验的设计-情境对相似，但分配给每个设计-情境对的仿真重复次数实际相差显著，导致选择正确概率$_W$存在差异。这一观察与实验1之前的讨论一致：当先验信息不足时，AOAmc 和 BOLDmc 在每一步都会将一次仿真重复分配给同一上下文，但分配的设计可能不同。AOAmc 会将更多仿真重复分配给那些表现难以学习的设计-情境对。针对 top-m 依赖上下文的选择问题，高效的抽样程序应同时考虑设计与上下文之间的抽样信息。

实验*5*：医疗资源分配示例。新冠疫情给医疗系统带来了巨大压力，尤其导致急诊科医疗资源短缺。除了维持急诊科的正常运转外，还需分配医疗资源，以确保收治的患者未感染新冠病毒。急诊科医疗资源分配的主要目标是缩短危重患者的等待时间。急诊科中的分配设计与不同的急诊患者流相关，而在新冠疫情期间，这些患者流具有高度变异性。在本示例中，我们使用离散事件仿真，根据每种急诊患者流场景中的不同分配决策，评估危重患者的等待时间。仿真模型的输入包括设计参数（医疗资源分配决策）和上下文参数（到达率）。

图8展示了急诊科急诊患者的治疗流程。乘坐救护车抵达急诊科的患者称为救护车送来的患者，其他抵达的患者则归类为自行就诊患者。自行就诊患者经入口进入，护士会评估他们是否有 COVID-19 症状。无症状者前往接待台，而救护车送来的患者则立即赶往接待台。接待员收集患者的个人信息，并开展流行病学调查，以识别疑似 COVID-19 病例。所有在入口和接待台识别出的 COVID-19 疑似病例都会被送往发热门诊，随后离开急诊科。离开接待台后，救护车送来的患者中病情严重者会被送往急诊室，而检查室的医生则评估其他患者的病情。医生决定患者是否需要接受进一步检查，这些检查由实验室的实验室技术人员执行。接受检查的患者随后返回检查室接受进一步评估。之后，病情不严重的救护车送来的患者会被送往治疗室；自行就诊患者则根据病情，在接受药物治疗后离开急诊科，或被转送至治疗室或急诊室。治疗室的患者完成治疗后离开急诊科，而急诊室的患者则会被转送至医院接受进一步治疗，之后离开急诊科。患者的路由概率如图8所示。对于系统仿真，我们采用2天的预热期，随后再进行7天的仿真（每天24小时）。本示例改编自 [2] 和 [55]。

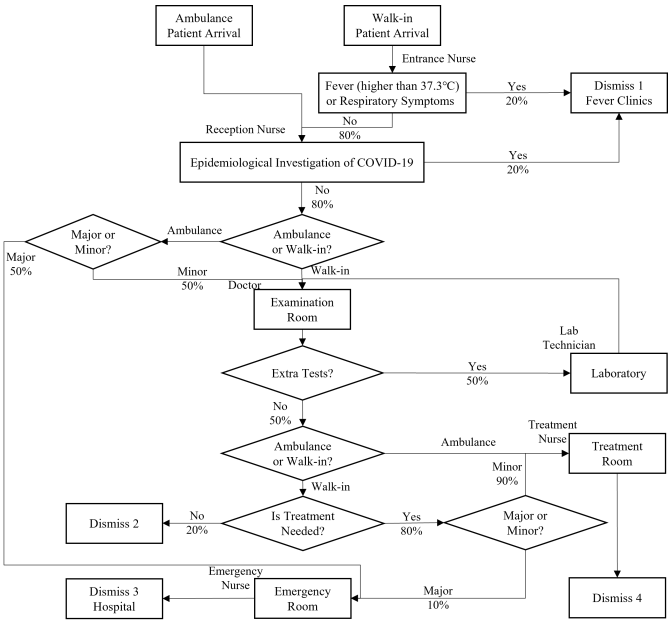


图 8. 实验5中急诊科患者的治疗流程。

在线附录 A.4 节的表 2 [49]展示了流程各阶段服务时间的分布。患者按到达顺序接受服务。如果当前阶段没有可用服务，患者将留在队列中，直到有服务可用。急诊科的到达过程遵循齐次泊松过程，其速率参数如在线附录 A.4 节的表 3 所示 [49]。急诊科配备了各种医疗资源，包括入口护士、接待员、医生、实验室技术人员、治疗床位和急诊床位。每种医疗资源都会在流程的相应阶段为患者提供服务。在线附录 A.4 节的表 4 [49]列出了 20 种医疗资源分配设计。危重患者指被分流至急诊室的患者。目标是确定能够使危重患者的期望

等待时间最短的前两种设计，适用于每种急诊患者流场景。在线附录 A.4 节的表 5[49] 展示了每种急诊患者流下前四种设计的表现，每种设计的表现均通过 100,000 次独立实验估算。我们可以看到，前两种设计因急诊患者流而异，且它们的表现差异很小。仿真总预算为 $T = 5000$。

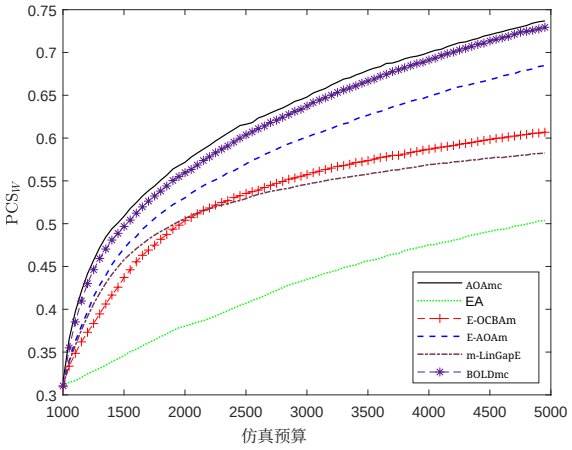


图 9. 医疗资源分配示例中六种采样策略的选择正确概率$_W$。

从图 9 可以看出，均匀分配在所有采样策略中表现最差。起初，m-LinGapE 的表现优于 E-OCBAm；但随着仿真预算增加，后者逐渐超过前者。起初，E-AOAm 略优于 E-OCBAm 和 m-LinGapE，且随着仿真预算增加，其提升速度快于后两种采样策略。BOLDmc 的选择正确概率$_W$ 增速略低于 AOAmc 的选择正确概率$_W$ 。在所有比较的采样策略中，AOAmc 表现最佳。AOAmc 消耗 4000 次仿真重复即可达到选择正确概率$_W$ $= 0.7$，而其他采样策略需要超过 4200 次仿真重复才能达到相同的选择正确概率$_W$ 水平；也就是说，AOAmc 将仿真重复的消耗降低了 4.8% 以上。

VI. 结论

本文研究了一个仿真预算分配问题，旨在选出排名前m的上下文相关设计。在贝叶斯框架下，我们将序贯采样决策表述为一个随机控制问题。本文提出了一种名为 AOAmc 的高效序贯采样程序，该程序具有计算可行性，并被证明是一致的。为分析所提采样策略的渐近最优性，我们严格推导了能够优化最坏情形下 PFS 大偏差率的渐近最优抽样比例。结果表明，在正态抽样分布下，所提分配策略的渐近抽样比例是渐近最优的。数值实验表明，AOAmc 能显著提高在所有上下文中学习排名前m 个设计的效率。利用并行计算，使每个设计-情境对能够同时接收多个仿真重复，并利用设计与上下文之间的函数关系来提高采样效率，值得在未来进一步研究。

致谢

本研究部分得到了国家自然科学基金委员会（NSFC）项目（编号：72250065、72022001、71901003）的资助，以及中国国家留学基金管理委员会（CSC）奖学金项目（编号：CSC202206010152）的资助。合著者Loo Hay Lee在本文撰写期间不幸去世。我们感谢他对本研究的贡献。

附录 I: 推论1的证明

Proof: Notice that if $A_{t+1}(\mathcal{E}_t) = (\langle\widetilde{i}\rangle_{rt}, r)$, $\widetilde{i} \in \{1, \cdots, m\}$, $r = 1, \cdots, q$, we have $\left(\sigma^{(t+1)}_{\langle\widetilde{i}\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2 < \left(\sigma^{(t)}_{\langle\widetilde{i}\rangle_{rt}}(\boldsymbol{x}_r)\right)^2$, $\left(\sigma^{(t+1)}_{\langle\ell\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2 = \left(\sigma^{(t)}_{\langle\ell\rangle_{rt}}(\boldsymbol{x}_r)\right)^2$, $\widetilde{\ell} = 1, \cdots, k$, $\widetilde{\ell} \neq \widetilde{i}$,

$$\min_{j=m+1,\cdots,k} \frac{\left(y^{(t)}_{\langle\widetilde{i}\rangle_{rt}}(\boldsymbol{x}_r) - y^{(t)}_{\langle j\rangle_{rt}}(\boldsymbol{x}_r)\right)^2}{\left(\sigma^{(t+1)}_{\langle\widetilde{i}\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2 + \left(\sigma^{(t+1)}_{\langle j\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2} > \min_{j=m+1,\cdots,k} \frac{\left(y^{(t)}_{\langle\widetilde{i}\rangle_{rt}}(\boldsymbol{x}_r) - y^{(t)}_{\langle j\rangle_{rt}}(\boldsymbol{x}_r)\right)^2}{\left(\sigma^{(t)}_{\langle\widetilde{i}\rangle_{rt}}(\boldsymbol{x}_r)\right)^2 + \left(\sigma^{(t)}_{\langle j\rangle_{rt}}(\boldsymbol{x}_r)\right)^2},$$

and if $A_{t+1}(\mathcal{E}_t) = (\langle\widetilde{j}\rangle_{rt}, r)$, $\widetilde{j} \in \{m+1, \cdots, k\}$, $r = 1, \cdots, q$, we have $\left(\sigma^{(t+1)}_{\langle\widetilde{j}\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2 < \left(\sigma^{(t)}_{\langle\widetilde{j}\rangle_{rt}}(\boldsymbol{x}_r)\right)^2$, $\left(\sigma^{(t+1)}_{\langle\widetilde{\ell}\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2 = \left(\sigma^{(t)}_{\langle\widetilde{\ell}\rangle_{rt}}(\boldsymbol{x}_r)\right)^2$, $\widetilde{\ell} = 1, \cdots, k$, $\widetilde{\ell} \neq \widetilde{j}$,

$$\min_{i=1,\cdots,m} \frac{\left(y^{(t)}_{\langle i\rangle_{rt}}(\boldsymbol{x}_r) - y^{(t)}_{\langle\widetilde{j}\rangle_{rt}}(\boldsymbol{x}_r)\right)^2}{\left(\sigma^{(t+1)}_{\langle i\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2 + \left(\sigma^{(t+1)}_{\langle\widetilde{j}\rangle_{r(t+1)}}(\boldsymbol{x}_r)\right)^2} > \min_{i=1,\cdots,m} \frac{\left(y^{(t)}_{\langle i\rangle_{rt}}(\boldsymbol{x}_r) - y^{(t)}_{\langle\widetilde{j}\rangle_{rt}}(\boldsymbol{x}_r)\right)^2}{\left(\sigma^{(t)}_{\langle i\rangle_{rt}}(\boldsymbol{x}_r)\right)^2 + \left(\sigma^{(t)}_{\langle\widetilde{j}\rangle_{rt}}(\boldsymbol{x}_r)\right)^2}.$$

令 $\widetilde{\ell}^{(t+1)} \notin \arg\min_{\ell=1,\cdots,q} APCS_\ell(\mathcal{E}_t)$、$\widehat{\ell}^{(t+1)} \in \arg\min_{\ell=1,\cdots,q} APCS_\ell(\mathcal{E}_t)$，即 $APCS_{\widetilde{\ell}^{(t+1)}}(\mathcal{E}_t) > APCS_{\widehat{\ell}^{(t+1)}}(\mathcal{E}_t)$。假设遵循采样规则 (4)，$A_{t+1}(\mathcal{E}_t) = (\widetilde{h}, \widetilde{\ell}^{(t+1)})$ $\widetilde{h} \in \{1, \cdots, k\}$，则有

$$APCS_{\widetilde{\ell}^{(t+1)}}\left(\mathcal{E}^E_{t+1}\right) = APCS_{\widetilde{\ell}^{(t+1)}}\left(\mathcal{E}_t \cup \mathbb{E}\left[Y^{t_{\widetilde{h}\widetilde{\ell}^{(t+1)}}+1}_{\widetilde{h}\widetilde{\ell}^{(t+1)}} \middle| \mathcal{E}_t\right]\right)$$

$$\geq APCS_{\widetilde{\ell}^{(t+1)}}(\mathcal{E}_t) > APCS_{\widehat{\ell}^{(t+1)}}(\mathcal{E}_t),$$

导致，对于 $\widetilde{\ell}=1,\ \cdots,\ k, \forall \widehat{\ell}^{(t+1)},\ \widehat{\ell}^{(t+1)}$,

$$\begin{aligned}\widehat{V}_t\left(\mathcal{E}_t;\left(\widetilde{\ell},\widehat{\ell}^{(t+1)}\right)\right)&=\min_{\ell=1,\cdots,q} APCS_\ell\left(\mathcal{E}_t\cup\mathbb{E}[Y^{t_{\widetilde{\ell}\widehat{\ell}^{(t+1)}}+1}_{\widetilde{\ell}\widehat{\ell}^{(t+1)}}|\mathcal{E}_t]\right)\\&=APCS_{\widehat{\ell}^{(t+1)}}\left(\mathcal{E}_t\right)=\min_{\ell=1,\cdots,q} APCS_\ell\left(\mathcal{E}_t\right),\end{aligned}$$

这与采样规则 (4) 相矛盾，因为根据该规则，分配的 $\widehat{\ }\quad\widetilde{\ }\quad\widetilde{\ }$设计-情境对具有最大的 $V_t(\mathcal{E}_t;(\ell,\ r))$，$\ell=1,\ \cdots,\ k, r=1,\ \cdots,\ q$。因此，遵循采样规则 (4) 的设计-情境对所分配的上下文为 $\widehat{\ell}^{(t+1)}\in \arg\min_{\ell=1,\cdots,q} APCS_\ell(\mathcal{E}_t)$。

令集合 $\widetilde{H}^{(t+1)}=\{1,\cdots,k\}\setminus H^{(t+1)}$，其中 $B\setminus A$ 表示集合 B 中集合 A 的相对补集。假设遵循采样规则 ($4\ \ A_{t+1}\ \mathcal{E}_t=\widetilde{h}^{(t+1)}\ \widetilde{\ell}^{(t+1)}\ \ \forall\widetilde{h}^{(t+1)}\in$)，( ) (，)，则有 $\widetilde{H}^{(t+1)}$

$$APCS_{\widehat{\ell}^{(t+1)}}\left(\mathcal{E}_t\cup\mathbb{E}[Y^{t_{\widetilde{h}(t+1)\widetilde{\ell}(t+1)}+1}_{\widetilde{h}(t+1)\widetilde{\ell}(t+1)}|\mathcal{E}_t]\right)=APCS_{\widehat{\ell}^{(t+1)}}\left(\mathcal{E}_t\right),$$

从而，$\forall\ \widetilde{h}^{(t+1)}$,

$$\begin{aligned}&\widehat{V}_t\left(\mathcal{E}_t;\left(\widetilde{h}^{(t+1)},\widetilde{\ell}^{(t+1)}\right)\right)\\&=\min_{\ell=1,\cdots,q} APCS_\ell\left(\mathcal{E}_t\cup\mathbb{E}\left[Y^{t_{\widetilde{h}(t+1)\widetilde{\ell}(t+1)}+1}_{\widetilde{h}(t+1)\widetilde{\ell}(t+1)}\middle|\mathcal{E}_t\right]\right)\\&=APCS_{\widehat{\ell}^{(t+1)}}\left(\mathcal{E}_t\right)=\min_{\ell=1,\cdots,q} APCS_\ell\left(\mathcal{E}_t\right)\ ,\end{aligned}$$

这与采样规则 (4) 相矛盾。根据该规则，分配的 $\widehat{\ }\quad\widetilde{\ }\quad\widetilde{\ }$设计-情境对具有最大的 $V_t(\mathcal{E}_t;(\ell,r))$，$\ell=1,\cdots,k,\ r=1,\cdots,q$。因此，按照采样规则 (4)，情境 $\widetilde{\ell}^{(t+1)}$ 下分配的设计 $\widehat{i}^{(t+1)}$ 属于集合 $H^{(t+1)}$，并可计算如下：$\forall\ h^{(t+1)}\in H^{(t+1)}$,

$$\widehat{i}^{(t+1)}\in\arg\max_{\langle h^{(t+1)}\rangle_{\widehat{\ell}(t+1)_t}}\widehat{V}_t\left(\mathcal{E}_t;\left(\left\langle h^{(t+1)}\right\rangle_{\widehat{\ell}(t+1)_t},\widehat{\ell}^{(t+1)}\right)\right)\ .$$

综上所述，推论得证。 ■

附录 II：推论 2 的证明

证明：根据（15），对于 $\ell\in\{1,\cdots,q\}$，我们有

$$\begin{aligned}&\min_{j=m+1,\cdots,k}\frac{r^*_\ell\left(y_{\langle i\rangle_\ell}(\boldsymbol{x}_\ell)-y_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{\sigma^2_{\langle i\rangle_\ell}(\mathbf{x}_\ell)\Big/(r^*_{\langle i\rangle_\ell}/r^*_\ell)+\sigma^2_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\Big/(r^*_{\langle j\rangle_\ell}/r^*_\ell)}\\&=\min_{i'=1,\cdots,m}\frac{r^*_\ell\left(y_{\langle i'\rangle_\ell}(\boldsymbol{x}_\ell)-y_{\langle j'\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{\sigma^2_{\langle i'\rangle_\ell}(\boldsymbol{x}_\ell)\Big/(r^*_{\langle i'\rangle_\ell}/r^*_\ell)+\sigma^2_{\langle j'\rangle_\ell}(\mathbf{x}_\ell)\Big/(r^*_{\langle j'\rangle_\ell}/r^*_\ell)},\end{aligned}$$

然后，令 $\alpha^*_{\langle h\rangle\ell}=r^*_{\langle h\rangle\ell}\Big/r^*_\ell, h=1,\ \cdots,\ k$，即可得到 (17)。由 $r^*_{\langle h\rangle\ell}>0$ 和 $\sum_{h=1}^k r^*_{\langle h\rangle\ell}=r^*_\ell>0$，可得 $\boldsymbol{\alpha}^*_\ell>0$ 和 $\sum_{h=1}^k\alpha^*_{h\ell}=1$。此外，由 (15) 还可推出，对于 $\ell,\ \ell'=1,\ \cdots,\ q$,

$$\begin{aligned}&\min_{j=m+1,\cdots,k}\frac{r^*_\ell\left(y_{\langle i\rangle_\ell}(\boldsymbol{x}_\ell)-y_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\right)^2}{\sigma^2_{\langle i\rangle_\ell}(\mathbf{x}_\ell)\Big/(r^*_{\langle i\rangle_\ell}/r^*_\ell)+\sigma^2_{\langle j\rangle_\ell}(\boldsymbol{x}_\ell)\Big/(r^*_{\langle j\rangle_\ell}/r^*_\ell)}=\\&\min_{j=m+1,\cdots,k}\frac{r^*_{\ell'}\left(y_{\langle i\rangle_{\ell'}}(\boldsymbol{x}_{\ell'})-y_{\langle j\rangle_{\ell'}}(\boldsymbol{x}_{\ell'})\right)^2}{\sigma^2_{\langle i\rangle_{\ell'}}(\mathbf{x}_{\ell'})\Big/(r^*_{\langle i\rangle_{\ell'}}/r^*_{\ell'})+\sigma^2_{\langle j\rangle_{\ell'}}(\boldsymbol{x}_{\ell'})\Big/(r^*_{\langle j\rangle_{\ell'}}/r^*_{\ell'})},\end{aligned}$$

根据 $z_\ell(\cdot)$ 的定义，可得 $r^*_\ell z_\ell(\alpha^*_\ell)=r^*_{\ell'}z_{\ell'}(\alpha^*_{\ell'})$.结合 $r^*_\ell>0$ 和 $\sum_{\ell=1}^q r^*_\ell=1$，可得

$$r^*_\ell=\frac{1/z_\ell(\boldsymbol{\alpha}^*_\ell)}{\sum_{\widetilde{\ell}=1}^q 1\Big/z_{\widetilde{\ell}}\left(\alpha^*_{\widetilde{\ell}}\right)},\quad \ell=1,\cdots,q\ .$$

根据 z^*的定义，并在正态性假设下，z^* 可表示为：$z^*=$ 最小值$_{j=m+1,\cdots,k}\frac{\left(y_{\langle i\rangle_\ell}\boldsymbol{x}_\ell-y_{\langle j\rangle_\ell}\boldsymbol{x}_\ell{}^2()()\right)()()\Big/r^*_{\langle j\rangle\ell}}{\sigma^2_{\langle i\rangle_\ell}\boldsymbol{x}_\ell\Big/r^*_{\langle i\rangle\ell}+\sigma^2_{\langle j\rangle_\ell}\boldsymbol{x}_\ell}$，从而得到

$$z^*=r^*_\ell z_\ell\left(\alpha^*_\ell\right)=\frac{1}{\sum_{\widetilde{\ell}=1}^q 1\Big/z_{\widetilde{\ell}}\left(\alpha^*_{\widetilde{\ell}}\right)}\ .$$

■

参考文献

[1] G. Melman、A. Parlikad 和 E. Cameron，“在新冠疫情期间利用离散事件仿真平衡医院稀缺资源”，《医疗保健管理科学》，第 24 卷，第 2 期，第 356–374 页，2021年。[2] T. R. Bovim、A. N. Gullhav、H. Andersson、J. Dale 和 K. Karlsen，“新冠疫情期间急诊患者流的仿真”，《仿真杂志》，第 1–15 页，2022年。[3] Z. Chen、C. Choe 和 N. Matsushima，“竞争性个性化定价”，《管理科学》，第 66 卷，第 9 期，第 4003–4023 页，2020年。[4] X. Chen、Z. Owen、C. Pixton 和 D. Simchi-Levi，“ 收益管理中个性化的统计学习方法 ”，《管理科学》，第 68 卷，第 3 期，第 1923–1937 页，2022年。[5] C.-H. Chen、D. He、M. Fu 和 L. H. Lee，“用于选择最优子集的高效仿真预算分配”，《*INFORMS*计算杂志》，第 20 卷，第 4 期，第 579–595 页，2008年。[6] S. Zhang、L. H. Lee、E. P. Chew、J. Xu 和 C.-H. Chen，“一种提高最优子集选择效率的仿真预算分配程序”，《*IEEE* 自动控制汇刊》，第 61 卷，第 1 期，第 62–75 页，2015年。[7] S. Gao 和 W. Chen，“一种用于选择前列仿真设计方案的新型预算分配框架”，《*IIE*汇刊》，第 48 卷，第 9 期，第 855–863 页，2016年。[8] G. Zhang、Y. Peng、J. Zhang 和 E. Zhou，“用于选择前m个备选方案的渐近最优抽样策略”，*arXiv* 预印本 *arXiv: 2111.15172*，2021年。[9] J. Du、S. Gao 和 C.-H. Chen，“速率最优的情境排序与选择”，*arXiv* 预印本 *arXiv:2206.12640*，2022年。[10] S.-H. Kim 和 B. L. Nelson，“选择最佳系统”，《运筹学与管理科学手册》，第 13 卷，第 501–534 页，2006年。[11] S. R. Hunter 和 B. L. Nelson，“并行排序与选择”，载于《建模与仿真进展》。施普林格，2017年，第 249–275 页。[12] L. J. Hong、W. Fan 和 J. Luo，“排序与选择综述：一种新视角”，《工程管理前沿》，第 8 卷，第 3 期，第 321–343 页，2021年。[13] S.-H. Kim 和 B. L. Nelson，“用于仿真中无差异区选择的完全序贯程序”，*ACM* 建模与计算机仿真汇刊（*TOMACS*），第 11 卷，第 3 期，第 251–273 页，2001年。[14] P. I. Frazier，“一种用于无差异区排序与选择的完全序贯淘汰程序，并对正确选择概率给出紧确界”，《运筹学》，第 62 卷，第 4 期，第 926–942 页，2014年。[15] C.-H. Chen、J. Lin、E. Y¨ucesan 和 S. E. Chick，“进一步提高序优化效率的仿真预算分配”，《离散事件动态系统杂志》，第 10 卷，第 3 期，第 251–270 页，2000年。[16] Y. Peng、E. K. Chong、C.-H. Chen 和 M. C. Fu，“作为随机控制的排序与选择”，《*IEEE* 自动控制汇刊》，第 63 卷，第 8 期，第 2359–2373 页，2018年。[17] Y. Peng、C.-H. Chen、M. C. Fu 和 J.-Q. Hu，“动态抽样分配与设计选择”，《*INFORMS*计算杂志》，第 28 卷，第 2 期，第 195–208 页，2016年。[18] P. Glynn 和 S. Juneja，“从大偏差视角看序优化”，载于《*2004*年冬季仿真会议论文集（*WSC*）》，第 1 卷。IEEE，2004年。

[19] Y. Chen 和 I. O. Ryzhov，“完全期望改进收敛至最优预算分配”，《应用概率进展》，第 51 卷，第 1 期，第 209–235 页，2019年。[20] H. Avci、L. Nelson、Barry 和 A. W¨achter，“在排序与选择中实现‘速率最优’”，载于《2021年冬季仿真会议（WSC）论文集》，第 1 卷。IEEE，2021年，第 1–12 页。[21] H. Li、H. Lam 和 Y. Peng，“依赖上下文的设计的高效学习、聚类与优化”，《运筹学》，第 1–22 页。https://doi.org/10.1287/opre.2022.2368，2022年。[22] Y. Chen 和 I. O. Ryzhov，“序贯选择中最优大偏差的平衡”，《管理科学》，提前在线发表，2022年。[23] S. Gao、J. Du 和 C.-H. Chen，“在协变量条件下选择最优系统设计”，载于《第15届自动化科学与工程国际会议（CASE）论文集》。IEEE，2019年，第 547–552 页。[24] X. Jin、H. Li 和 L. H. Lee，“仿真分析中的最优预算分配*”，载于《第15届自动化科学与工程国际会议（CASE）论文集》，2019年，第 178–182 页。[25] H. Shen、L. J. Hong 和 X. Zhang，“具有协变量的排序与选择”，载于《2017年冬季仿真会议（WSC）论文集》。IEEE，2017年，第 2137–2148 页。[26] ——，“面向个性化决策的具有协变量的排序与选择”，《INFORMS计算杂志》，第 33 卷，第 4 期，第 1500–1519 页，2021年。[27] X. Li、X. Zhang 和 Z. Zheng，“数据驱动的排序与选择：高维协变量与一般依赖关系”，载于《2018年冬季仿真会议（WSC）论文集》。IEEE，2018年，第 1933–1944 页。[28] S. Cakmak、S. Gao 和 E. Zhou，“基于高斯过程的情境排序与选择”，arXiv 电子预印本，第 arXiv: 2201.07782 页，2022年。[29] R. Hu 和 M. Ludkovski，“响应曲面排序的序贯设计”，《SIAM/ASA不确定性量化杂志》，第 5 卷，第 1 期，第 212–239 页，2017年。[30] M. Pearce 和 J. Branke，“具有相关性的连续多任务贝叶斯优化”，《欧洲运筹学杂志》，第 270 卷，第 3 期，第 1074–1085 页，2018年。[31] L. Ding、L. J. Hong、H. Shen 和 X. Zhang，“具有协变量的选择问题的知识梯度：一致性与计算”，《海军研究后勤学》，第 69 卷，第 3 期，第 496–507 页，2022年。[32] E. Hazan 和 N. Megiddo，“利用先验知识进行在线学习”，载于《第20届年度学习理论会议（COLT）论文集》。施普林格，2007年，第 499–513 页。[33] A. Krishnamurthy、J. Langford、A. Slivkins 和 C. Zhang，“具有连续动作的上下文老虎机：平滑、缩放与自适应”，《机器学习研究杂志》，第 21 卷，第 1 期，第 5402–5446 页，2020年。[34] L. Li、W. Chu、J. Langford 和 R. E. Schapire，“一种用于个性化新闻文章推荐的上下文老虎机方法”，载于《第19届万维网国际会议（WWW）论文集》，2010年，第 661–670 页。[35] Y. Abbasi-Yadkori、D. P´al 和 C. Szepesv´ari，“改进的线性随机老虎机算法”，《神经信息处理系统进展》，第 24 卷，第 2312–2320 页，2011年。[36] R. Allesiardo、R. F´eraud 和 D. Bouneffouf，“用于上下文老虎机问题的神经网络委员会”，载于《第21届神经信息处理国际会议（ICONIP）论文集》。施普林格，2014年，第 374–381 页。[37] R. F´eraud、R. Allesiardo、T. Urvoy 和 F. Cl´erot，“用于上下文老虎机问题的随机森林”，载于《第19届人工智能与统计国际会议（AISTATS）论文集》。PMLR，2016年，第 93–101 页。[38] S. Kalyanakrishnan、A. Tewari、P. Auer 和 P. Stone，“随机多臂老虎机中的 PAC 子集选择”，载于《第30届机器学习国际会议（ICML）论文集》，第 12 卷，2012年，第 655–662 页。[39] E. Kaufmann 和 S. Kalyanakrishnan，“老虎机子集选择中的信息复杂度”，载于《第26届年度学习理论会议（COLT）论文集》。PMLR，2013年，第 228–251 页。[40] L. Chen、J. Li 和 M. Qiao，“Top-k 臂选择的近似实例最优样本复杂度界”，载于《第20届人工智能与统计国际会议（AISTATS）论文集》。PMLR，2017年，第 101–110 页。[41] S. Bubeck、T. Wang 和 N. Viswanathan，“多臂老虎机中的多重识别”，载于《第31届机器学习国际会议（ICML）论文集》。PMLR，2013年，第 258–265 页。

[42] C. R´eda、E. Kaufmann 和 A. Delahaye-Duriez，“线性多臂老虎机的前 m 名识别”，载于第 *24* 届人工智能与统计国际会议论文集（*AISTATS*）。PMLR，2021年，第 1108–1116 页。[43] H. Lam，“对输入过程序列依赖性的敏感性：一种稳健方法”，《管理科学》，第 64 卷，第 3 期，第 1311–1327 页，2018年。[44] S. Ghosh 和 H. Lam，“随机模拟中的稳健分析：计算与性能保证”，《运筹学》，第 67 卷，第 1 期，第 232–249 页，2019年。[45] S. Gao、H. Xiao、E. Zhou 和 W. Chen，“采用最优计算预算分配的稳健排序与选择”，《自动化》，第 81 卷，第 30–36 页，2017 年。[46] W. Fan 、 L. J. Hong 和 X. Zhang，“ 最佳方案的分布鲁棒选择”，《管理科学》，第 66 卷，第 1 期，第 190–208 页，2020年。[47] H. Xiao 和 S. Gao，“在输入不确定性下选择前 m 个设计的仿真预算分配”，《*IEEE* 自动控制汇刊》，第 63 卷，第 9 期，第 3127–3134 页，2018年。[48] D. P. Bertsekas，《动态规划与最优控制》。Athena scientific，贝尔蒙特，马萨诸塞州，1995年，第 1 卷，第 2 期，命题 1.3.1。[49] G. Zhang、S. Chen、K. Huang 和 Y. Peng，“《高效学习以选择前m 个上下文相关设计》的在线附录”，2022年，可通过 SSRN 获取：https://ssrn.com/abstract=4316394。[50] O. Zeitouni 和 A. Dembo，《大偏差技术及其应用》，第 2 版。施普林格出版社，纽约，2010年，第 38 卷。[51] C.-H. Chen 和 L. H. Lee，《随机模拟优化：最优计算预算分配》。世界科学出版社，2011年，第 1 卷。[52] C. Kiekintveld、V. Lis`y 和 R. Pıbil，“网络安全中蜜罐策略性使用的博弈论基础”，载于《网络战》。施普林格，2015年，第 81–101 页。[53] Q. D. La、T. Q. Quek、J. Lee、S. Jin 和 H. Zhu，“物联网蜜罐网络中的欺骗攻防博弈”，《*IEEE* 物联网期刊》，第 3 卷，第 6 期，第 1025–1035 页，2016年。[54] A. H. Anwar、C. Kamhoua 和 N. Leslie，“网络欺骗博弈中基于攻击图的蜜罐分配”，载于*2020* 年计算、网络与通信国际会议论文集（*ICNC*）。IEEE，2020年，第 502–506 页。[55] M. A. Ahmed 和 T. M. Alkhamis，“科威特急诊科医疗单元的仿真优化”，《欧洲运筹学杂志》，第 198 卷，第 3 期，第 936–942 页，2009年。