

本科毕业论文（设计）

UNDERGRADUATE THESIS (PROJECT)

题 目：

学 院：

专 业：

学 号：

学生姓名：

指导教师：

起讫日期：

|  |  |
| --- | --- |
| 姓 名： | 学号： |
| 论文题目： | |

**原 创 性 声 明**

本人声明：所呈交的论文是本人在指导教师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签 名： 日 期：

**本论文使用授权说明**

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

（**保密的论文在解密后应遵守此规定**）

签 名： 指导教师签名： 日期：

摘 要

摘要的内容需作者简要介绍本论文的主要内容，主要为本人所完成的工作和创新点。

……

(注：标题黑体小二号，正文宋体小四，行距20磅)

**关键词：**学位论文；论文格式；规范化；模板

（注：不宜多，最好3~5个）

ABSTRACT

The content of the abstract requires the author to briefly introduce the main content of this paper, mainly for my work and innovation.

…….

(Times New Roman，小四号，行距20磅)

**Keywords:** Dissertation; Dissertation format; Standardization; Template

目 录

[摘 要 I](#_Toc2372)

[ABSTRACT II](#_Toc26076)

[1 绪论 1](#_Toc88)

[1.1 研究意义 1](#_Toc10954)

[1.1.1 在线问题的现实意义 1](#_Toc23983)

[1.1.2 在线最优选择问题 2](#_Toc3714)

[1.2 本文主要研究内容 4](#_Toc16118)

[1.3 本文组织结构 6](#_Toc18050)

[2 研究背景 7](#_Toc18)

[2.1 节一级标题（黑体四号） 7](#_Toc20988)

[2.1.1 节二级标题（黑体小四号） 7](#_Toc21215)

[2.1.2 节二级标题（黑体小四号） 7](#_Toc21759)

[2.1.3 节二级标题（黑体小四号） 8](#_Toc17972)

[2.2 节一级标题（黑体四号） 8](#_Toc14853)

[2.2.1 节二级标题（黑体小四号） 8](#_Toc15798)

[2.2.2 节二级标题（黑体小四号） 9](#_Toc13544)

[2.2.3 节二级标题（黑体小四号） 9](#_Toc10614)

[3 算法设计 12](#_Toc23178)

[3.1 不可回溯情况下选取最大值 13](#_Toc897)

[3.1.1 证明最优停止策略的存在性 13](#_Toc23222)

[3.1.2 最优停止策略的描述 14](#_Toc2743)

[3.2 不可回溯且观察消耗成本情况下选取最大收益 16](#_Toc5633)

[3.2.1 证明最优停止策略的存在性 16](#_Toc11696)

[3.2.2 最优停止策略的描述 17](#_Toc22449)

[3.2.3 均匀分布下的选择规则 18](#_Toc28022)

[3.2.3 指数分布下的选择规则 19](#_Toc31740)

[3.3 有失败回弹风险 21](#_Toc32603)

[3.3.1 证明最优停止策略的存在性 21](#_Toc19739)

[3.3.2 最优停止策略的描述 22](#_Toc28214)

[3.3.3 指数分布下的选择规则 24](#_Toc30206)

[4 算法验证 29](#_Toc16075)

[4.1 不可回溯情况下选取最大值问题的算法与实验 29](#_Toc8107)

[4.1.1 算法设计 29](#_Toc30759)

[4.1.2 实验方案 30](#_Toc13965)

[2.1.3 实验指标 30](#_Toc14396)

[4.1.4 实验结果 30](#_Toc32573)

[4.2 不可回溯且观察消耗成本的选择算法与实验 33](#_Toc32258)

[4.2.1 算法设计 33](#_Toc7816)

[4.2.2 实验方案 35](#_Toc24033)

[4.2.3 实验指标 35](#_Toc5831)

[4.2.4 实验结果 36](#_Toc14901)

[4.3 有失败回弹风险的算法验证 41](#_Toc3259)

[4.3.1 算法设计 41](#_Toc10537)

[4.3.2 实验方案 42](#_Toc30863)

[4.3.3 实验指标 42](#_Toc14912)

[4.3.4 实验结果 43](#_Toc19377)

[结论（黑体小二号） 47](#_Toc8513)

[参考文献（黑体小二号） 48](#_Toc6811)

[附 录（黑体小二号） 51](#_Toc14687)

[致 谢 52](#_Toc4553)

注意：

正文各章节应拟标题，每章结束后应另起一页。标题要简明扼要，不应使用标点符号。各章、节、条的层次，可以按照“1……、1.1……、1.1.1……”标识，条以下具体款项的层次依次按照“1.1.1.1”或“（1）”、“①”等标识。各学院根据实际情况，可自行规定层次格式，但学院之内建议格式统一，以清晰无误为准。

正文是毕业论文的主体和核心部分，不同学科专业和不同的选题可以有不同的写作方式。正文一般包括以下几个方面。

1.引言或背景

引言是论文正文的开端，引言应包括毕业论文选题的背景、目的和意义；对国内外研究现状和相关领域中已有的研究成果的简要评述；介绍本项研究工作研究设想、研究方法或实验设计、理论依据或实验基础；涉及范围和预期结果等。要求言简意赅，注意不要与摘要雷同或成为摘要的注解。

2.主体

论文主体是毕业论文的主要部分，必须言之成理，论据可靠，严格遵循本学科国际通行的学术规范。在写作上要注意结构合理、层次分明、重点突出，章节标题、公式图表符号必须规范统一。论文主体的内容根据不同学科有不同的特点，一般应包括以下几个方面：

（1）毕业设计（论文）总体方案或选题的论证；

（2）毕业设计（论文）各部分的设计实现，包括实验数据的获取、数据可行性及有效性的处理与分析、各部分的设计计算等；

（3）对研究内容及成果的客观阐述，包括理论依据、创新见解、创造性成果及其改进与实际应用价值等；

（4）论文主体的所有数据必须真实可靠，自然科学论文应推理正确、结论清晰；人文和社会学科的论文应把握论点正确、论证充分、论据可靠，恰当运用系统分析和比较研究的方法进行模型或方案设计，注重实证研究和案例分析，根据分析结果提出建议和改进措施等。

3.结论

结论是毕业论文的总结，是整篇论文的归宿。应精炼、准确、完整。着重阐述自己的创造性成果及其在本研究领域中的意义、作用，还可进一步提出需要讨论的问题和建议。

# **绪论**

## 研究意义

### 1.1.1 在线问题的现实意义

随着信息技术的不断进步，人们获得信息的渠道越来越多样化，这给我们提供了更多选择的机会，但也带来了选择的困扰。例如，社交媒体的普及使我们可以与全球各地的人们联系，获取各种信息，但也使我们面临更多的信息过载和选择焦虑。因此，研究如何在信息爆炸的环境中做出理性、高效的选择变得尤为重要。

在社交媒体和电子商务领域，在线选择问题的研究可以帮助我们理解用户行为和偏好，为平台提供个性化的服务和内容推荐。例如，通过分析用户的浏览历史和点击行为，电子商务平台可以向用户推荐更符合其兴趣和需求的商品，从而提高用户满意度和购买转化率。在社交媒体领域，了解用户的好友关系和兴趣爱好，可以帮助平台更好地推送相关内容，增加用户的参与度和粘性。

此外，从决策科学的角度来看，在线选择问题涉及到人们的决策过程和选择行为，是决策科学研究的一个重要领域。通过对在线选择问题的研究，我们可以深入了解人们是如何做出决策的，以及他们在做出选择时所面临的困难和挑战。这有助于我们制定出更加科学和有效的决策策略，提高个人和组织的决策能力。例如，在金融投资领域，了解投资者的选择行为和决策偏好，可以帮助投资者制定更明智的投资策略，降低投资风险，提高收益率。

个性化推荐系统的发展为在线选择问题的研究提供了一个重要的背景。个性化推荐系统是通过分析用户的历史行为和偏好，向他们推荐最相关的内容或产品的系统。在线选择问题的研究可以为个性化推荐系统的设计和优化提供理论支持，从而提高推荐的准确性和用户满意度。例如，在视频流媒体平台上，个性化推荐系统可以根据用户的观看历史和喜好，推荐最适合他们的影视作品，提高用户的观看体验和平台的留存率。

最后，从经济学和市场营销的角度来看，在线选择问题涉及到市场竞争、消费者行为和产品定价等重要问题。通过研究在线选择问题，我们可以更好地理解市场运作的机制，为企业制定更加有效的产品设计和营销策略提供参考。例如，在线购物平台可以根据消费者的选择偏好和购买行为，调整产品的定价和促销策略，提高销售额和市场份额。

综上所述，在线选择问题是信息时代背景下的一个重要议题，涉及到社交媒体、电子商务、决策科学、个性化推荐、经济学和市场营销等多个领域。通过研究在线选择问题，我们可以更好地理解人们的选择行为和决策过程，指导平台设计和优化，提高个人和组织的决策能力，促进市场竞争和经济发展。

### 1.1.2 在线最优选择问题

在线最优选择问题是计算机科学中一个重要且引人关注的领域，它探索了在面临未知未来信息的情况下，如何做出最佳决策的方法和策略。这个问题在众多实际场景中都有广泛的应用，包括但不限于股票交易、资源分配、在线匹配等。通过研究在线最优选择问题，我们能够深入理解在线系统的工作原理、性能评价以及优化方法，从而提高在线系统的效率和性能。

首先，让我们来深入了解在线最优选择问题的基本概念。在这个问题中，我们面临的挑战是在不知道未来信息的情况下，通过一系列的选择来达到最优的结果。换句话说，我们必须根据已有的信息和经验，做出每一步的最佳选择，以期望在整个过程中达到最优的结果。这个问题的复杂性在于，我们无法预测未来的发展趋势和变化，因此必须依靠对已有信息的合理利用和决策。

为了解决在线最优选择问题，我们需要设计和分析相应的算法。这些算法需要具备一定的智能和灵活性，能够根据当前情况做出合理的决策。在算法设计中，通常会结合各种启发式方法、贪婪算法或者动态规划等技术，以实现在不确定环境下的最优决策。同时，对算法性能的评估和分析也是至关重要的，我们需要考虑算法的时间复杂度、空间复杂度以及实际应用中的表现。

竞争比分析是评估在线最优选择算法性能的重要方法之一。竞争比指的是在线算法在最坏情况下的性能与最优离线算法性能之间的比值。通过竞争比分析，我们可以评估在线算法的效率和实用性，了解它们在不同情况下的表现优劣，并进一步优化和改进算法设计。

在线最优选择问题的研究不仅有助于我们理解在线系统的工作原理和机制，还能够帮助我们发现在线系统的实际价值和潜力。例如，在股票交易领域，一个有效的在线最优选择算法可以帮助投资者根据当前市场情况做出及时的交易决策，最大程度地获取利润。在资源分配和在线匹配等场景中，优秀的在线最优选择算法也可以提高资源利用效率，实现资源的合理分配和任务的高效执行。

此外，对在线最优选择问题的应用场景进行深入研究还能够为我们提供改进和创新的方向。通过分析不同领域的需求和挑战，我们可以针对性地设计和优化在线算法，提高其适用性和效率。同时，结合新技术和方法，也可以拓展在线最优选择问题的解决范围，推动在线系统的发展和应用。

在线最优选择算法是信息检索、推荐系统、在线广告等领域中至关重要的问题之一。随着互联网的快速发展，用户和数据量的增加，以及应用场景的不断变化，设计出具有更好竞争比和性能的在线最优选择算法成为了一项迫切的挑战。本文将探讨在线最优选择算法的发展历程、不同应用场景下的算法选择以及未来研究的方向和挑战。

在过去的几十年里，学术界和工业界已经提出了许多在线最优选择算法，其中包括随机排列算法、阈值算法、均匀采样算法、指数加权算法等。这些算法在不同的应用场景和模型下具有不同的竞争比和性能。例如，随机排列算法和阈值算法适用于选择单个最优元素的情况，而均匀采样算法和指数加权算法适用于选择多个最优元素的情况。这些算法的设计和分析旨在寻求更好的竞争比和性能，并且已经取得了一定的成果。

然而，随着在线系统的应用越来越广泛和复杂，传统的在线最优选择算法面临着新的挑战。例如，现实世界中的应用往往涉及到元素之间的相关性、不确定性、动态性、多目标性等因素。在这种情况下，传统的算法可能无法有效地适应复杂的环境，因此需要设计出更灵活、更智能的在线最优选择算法。

同时，现代在线系统还具有反馈、学习、自适应等特性，这为在线最优选择算法的设计带来了新的机遇和挑战。例如，在线广告系统可以根据用户的反馈来调整广告投放策略，推荐系统可以根据用户的行为来更新推荐结果。这些特性使得在线最优选择算法可以通过学习和自适应来不断优化性能，但同时也需要解决如何有效地利用反馈信息和学习算法的设计问题。

因此，未来的研究方向之一是设计出更智能、更高效的在线最优选择算法，以应对复杂的应用场景和挑战。这可能涉及到机器学习、强化学习等技术的应用，以及对算法的理论分析和实验验证。另外，还需要进一步研究在线最优选择算法的竞争比的下界和上界，以更好地理解算法的性能和局限性。

除了算法设计和分析之外，还需要关注在线最优选择算法在实际应用中的效果评估和优化。这可能涉及到大规模实验、仿真模拟等方法，以及与行业合作进行实际系统的部署和测试。通过不断地实践和反思，可以进一步改进和优化在线最优选择算法，使其能够更好地应对现实世界中的挑战和需求。

综上所述，随着在线系统的发展和应用场景的变化，设计出更好竞争比和性能的在线最优选择算法成为了一项重要的研究课题。未来的研究需要继续关注算法设计、理论分析、实验评估等方面，以不断推动在线最优选择算法的发展和应用。

## 本文主要研究内容

停止选择问题是在线选择问题的一个重要分支，其关注点在于在持续接收数据流的过程中，如何在适当的时机选择停止节点，以最大化总体收益。本文旨在深入探讨停止选择问题的各种情形，针对各种情形设计选择策略，并设计实验验证选择算法的性能。

在线选择问题是指在动态环境下，系统需要根据不断到来的信息做出即时决策的情境。而停止选择问题则进一步强调了在连续流数据中选择适当的时机停止，并在停止时获取最大利益的策略。在实际应用中，停止选择问题具有多种情形和应用场景。举例而言，可以将其应用于公司的员工招聘问题。在招聘过程中，公司需要在面试了一定数量的候选人之后做出决策，是否继续招聘或者停止招聘并选择其中一位候选人。在这种情况下，选择适当的停止时机可以帮助公司节省时间和资源，并确保最终选择到最合适的员工。另一个例子是股票买卖问题。投资者需要在不断变化的市场条件下做出买入或卖出股票的决策。在这种情况下，选择合适的停止点可以帮助投资者最大化利润，避免因持有时间过长或过短而造成损失。类似地，房屋出售问题也可以看作是停止选择问题的一个实例。在房地产市场上，卖方需要在适当的时机决定是否接受某个报价并完成交易，或者继续等待更好的机会。通过有效的停止选择策略，卖方可以最大程度地优化自己的收益。

除了以上提及的情形外，停止选择问题还可以应用于企业资金链的运营、假设检验样本的选取等各种领域。在企业资金链运营中，管理者需要在资金供给和需求之间取得平衡，并在适当的时机进行投资或者回收资金。而在假设检验中，研究人员需要在实验进行到一定阶段时停止并得出结论，以节省资源并确保实验结果的准确性。

针对停止选择问题的复杂性和多样性，本文将设计和验证多种选择策略，以应对不同情形下的挑战。通过对现有算法的分析和改进，本文将提出一系列有效的停止选择策略，并通过实验验证其性能和可行性。这将为解决实际应用中的停止选择问题提供有力的理论支持和实践指导。

本文主要研究三个比较典型的抽象模型，它们在实际应用中具有重要意义：

1、不可反悔情形中的最大值选取：这个模型涉及到在做出决策后无法撤销的情况下，如何选择最优解。在现实生活中，这种情形很常见。例如，一家公司在制定战略方案时，可能需要在多个方案中选择一个，在选择后便无法更改。这就要求决策者综合考虑各种因素，包括风险、收益、市场变化等，以确保所做出的选择是最具有潜在回报的。

2、在不可反悔情形中增加成本值：在这个模型中，除了考虑到不可更改的决策外，还需要考虑到决策可能带来的成本。这就要求决策者在做出选择时，不仅要考虑到最终的结果，还要权衡可能产生的额外成本。例如，在电力市场能源交易中，电力公司需要权衡市场需求、发电成本、环境因素等多种因素，以在不可更改的情况下做出最佳决策。

3、在有失败回弹风险的情形中的最大累计值的选取：这个模型涉及到在存在失败风险的情况下，如何最大化累积价值。在现实生活中，风险是不可避免的，因此在做出决策时需要考虑到可能的失败情况，并采取相应的措施进行风险管理。例如，一个企业在推出新产品时，可能会面临市场竞争、技术失败等风险，因此需要在产品设计、市场推广等方面进行充分准备，以确保即使面临失败，也能最大程度地保留累积价值。

这些实际应用情形展示了这三个抽象模型在不同领域的广泛应用，它们可以帮助人们在面对各种决策时更好地理解并优化其选择。本文旨在深入研究该问题，提出有效的选择策略，并测试验证算法的表现性能。

## 本文组织结构

本文全篇由五个章节组成，每个章节的主要内容结构如下：

第一章：绪论。本章会介绍在线问题背景知识，并着重介绍在线选择问题与在线选择问题中的停止选择问题。

第二章：研究背景。本章会介绍国内外现在与过去在在线选择问题上的已有的研究，同时也会介绍本文在三四章所参考与应用到的知识与数学工具。

第三章：算法设计。本章对三个不同情形下的在线最优停止问题进行定义，并且介绍了不同情形下选择策略与算法的设计与证明。

第四章：算法验证。本章会设计实验复现第三章中所提出的选择算法，并且使用多种指标多方面评估这些算法的表现。

第五章：总结与展望。本章会对全文研究内容进行概况，归纳算法特性，并对在线选择算法的未来研究方向进行展望。

# **研究背景**

## 2.1 国内外研究

在线最优选择算法指的是在网络环境下，通过一系列算法实现对网络资源的最优选择，以提高网络性能和用户体验。国内外对于在线最优选择算法的研究现状主要包括以下几个方面：

### 2.1.1 基于博弈论的算法研究

博弈论可以描述多方参与决策的情景，因此也被应用在在线最优选择算法的研究中。通过博弈模型，可以分析网络中各方的利益和策略，从而设计出更加合理的资源选择策略。

最优停止理论是博弈论中的一个经典问题，关注的是如何基于顺序观察的随机变量选择一个行动的时间，以最大化预期收益或最小化预期成本。这个理论在经济学、金融领域使用非常广泛，例如美式期权在股票交易中看涨看跌，执行期权，基本都使用停止理论来求解。近年来，最优停止理论也在网络欺骗防御策略优化等领域得到了应用。最优停止理论为我们提供了一个在不确定性环境下做出最优选择的方法，它在许多领域都有广泛的应用。然而，这只是理论上的最优解，实际应用中可能会受到许多其他因素的影响，因此在实际应用中需要根据具体情况进行调整。

### 2.1.2 基于机器学习的算法研究

随着机器学习技术的发展，越来越多的研究开始将机器学习方法应用于在线最优选择算法中。这类算法通过对历史数据的学习，预测不同资源选择的性能，从而实现最优选择。

基于机器学习的在线最优停止问题是指在一个连续的决策过程中，利用机器学习技术来实现在何时停止决策以达到最优的目标。这种问题在实际应用中广泛存在，比如在线广告投放、网络流量管理、自动驾驶等领域。通常采用马尔可夫决策过程或者部分可观察的马尔可夫决策过程等形式构建决策模型，使用历史数据来学习决策规则，并根据目标函数进行优化。

### 2.1.3 基于深度学习的算法研究

深度学习技术的出现为在线最优选择算法提供了新的思路和方法。通过深度学习模型对网络状态进行建模和预测，可以实现更加精准的资源选择。

基于深度学习的最优停止问题方法，可以直接从蒙特卡洛样本中学习最优停止规则。在高维情况下，如伯努利最大期权定价、可召回多重障碍可逆转换证券定价和分数布朗运动的最优停止问题，该方法展现出了准确性和快速计算的优势。深度学习方法对于理解和解决高维最优停止问题提供了新的视角，尤其是在金融工程和风险管理领域，展示了深度学习在复杂决策问题中的潜力。

实际应用与性能评估：除了理论研究，国内外也进行了大量的实际应用与性能评估工作。通过在真实网络环境下的实验，评估不同算法的性能和效果，为算法的进一步优化提供参考。

在国内，互联网行业的快速发展和大规模应用，推动了在线最优选择算法的研究与实践。一些大型互联网企业和研究机构积极投入到该领域，推动了相关技术的不断创新和进步。与此同时，国外的学术界也在持续关注和研究在线最优选择算法，不断提出新的理论和方法，丰富了该领域的研究内容和成果。

## 2.2 预备知识

**（黑体四号，左对齐）**

×××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××**（格式：宋体，Times New Roman小四号字，两边对齐，首行缩进2个字符，行距23磅，字符间距为“标准”）**

.......

### 2.2.1 节二级标题（黑体小四号）

**（黑体小四号，左对齐）**

×××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××**（格式：宋体，Times New Roman小四号字，两边对齐，首行缩进2个字符，行距23磅，字符间距为“标准”）**

.......

### 2.2.2 节二级标题（黑体小四号）

**（黑体小四号，左对齐）**

×××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××**（格式：宋体，Times New Roman小四号字，两边对齐，首行缩进2个字符，行距23磅，字符间距为“标准”）**

.......

### 2.2.3 节二级标题（黑体小四号）

**（黑体小四号，左对齐）**

×××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××××**（格式：宋体，Times New Roman小四号字，两边对齐，首行缩进2个字符，行距23磅，字符间距为“标准”）**

.......

# **算法设计**

上文已经介绍了最优停止问题在现实中的普遍存在性以及最优选择策略在统计学和运筹学领域的重要性。这一章会介绍三种十分常见的最优停止问题情形，以及相应问题的较优解方案。在给出针对最优停止问题的最优选择策略以前，我们需要先给出此类问题的严格数学定义。

最优停止问题定义：

（1）已知随机变量序列 假设其联合分布已知；

（2）已知实值奖励函数序列：

决策者可逐个观察随机序列。并且对于任意第n个观测对象，，决策者可以执行两个动作：

1. 选择停止观察并获得奖励。
2. 选择继续观察下一观测对象。

如果决策者不做任何观察就选择停止观察，则会获得奖励；如果决策者一直选择继续观察而从不停止，则会获得奖励。

决策者需在恰当时刻选择停止观察以获得奖励，使得满足条件：

通过带入满足不同分布的随机样本以及设定不同的实值奖励函数，可以将此抽象问题具体为情况不一的实际问题。此类问题情形很多，包括但不限于：

（1）不可回溯情况下选取最大值；

（2）不可回溯且观察消耗成本情况下选取最大收益；

（3）有概率收益清零的情况下选取最大和值。

本章后续也将围绕上述这三种十分常见的典型情况设计较优的选择策略，并分析不同选择策略的性能与共同点。

## 3.1 不可回溯情况下选取最大值

在观察消耗成本情况下选取最大收益的情形下的最优停止问题可以描述为如下问题：

待观察的随机序列为：

其中随机样本中各对象之间相互独立且服从同一分布，分布存在期望与方差，且取值范围为。

相应的实值奖励函数序列：

其中是表示一次观察都不做的收益是无穷小，因而至少需要进行一次观察，表示选择停止时的奖励值为该时刻时所得到的观察值，若在观察到最后一个随机样本前从未选择停止，则在观察到最后一个观察对象后强制选择停止。

### 3.1.1 证明最优停止策略的存在性

由已知文献可知，对于待观察随机序列和实值奖励序列存在最优停止策略的充分条件为以下两个条件：

A1：

A2：

由于，故，又因为应当满足，因此成立。

同理，由于，故，又因为应当满足，因此。因而假设A1和A2均被满足，所以最优停止规则存在。

由此可得，该问题背景下满足条件A1和A2，因此最优停止策略存在，使得观察者以尽可能大的概率选择出最大收益。**（证明过于简陋）**

### 3.1.2 最优停止策略的描述

首先观察问题，在任意时刻选择继续观察表示放弃获取当前的奖励，接着观察到并面临与时刻时一样的处境。因此可以设表示时刻获得的奖励在序列中所处的排名。因此该问题可转换为选择出排名尽可能靠前的奖励值。又由于若选中的是全部候选值中的最大值，那么同时也也是局部前个观察值中的最大值。可以假设存在一个阈值表示在时刻以后出现的局部最大值是全局最大值的可能性最大，因此选择策略为：

1. 当时刻时的观察值为前个观察值中的最大值时，表明此时选中的奖励值是全局最大值的可能性最大，可以认为选择继续观察出现更大值的概率小于不会出现更大值的概率，因此选择停止。
2. 当时刻时的观察值为前个观察值的局部最大值时，表明此时选择继续观察能观察到更大值的概率比不会出现更大值的概率更大，因而选择继续观察。

因此由最优性原理得出的最佳停止策略为在处停止，其中应当满足如下条件：

因此现在的问题只是计算出阈值。在计算阈值之前先提出以下假设与已知条件：

1. 为放弃前个候选值后能够选中最大值得概率，
2. 在(为局部最大值)处选择停止时能选到全局最大值得概率为（表示所有观察值得数量），
3. 选择停止的时刻应满足条件，
4. 是全局最大值的概率为，
5. 被选中的概率等价于在没有出现比前个候选值更大值且比前个候选值中的最大值更大的概率。

由上述已知得，因此对于任意都应满足如下条件：

则可设表示阈值条件下以此选择策略能选中最大值得概率，且由定义可得的表达式如下：

又当时是一个无穷级数，且可计算得

因此

设，则可得

将对求导得

因此当时，取得极大值。因此解得

由上述推导结果得，当的取值为时，该选择策略能选中最大值的概率。最优停止策略为：拒绝前的样本后遇到最大值时选中，选中值为最大值的概率为，大约。

## 3.2 不可回溯且观察消耗成本情况下选取最大收益

在观察消耗成本情况下选取最大收益的情形下的最优停止问题可以描述为如下问题：

待观察的随机序列为：

其中随机样本中各对象之间相互独立且服从同一分布，且分布存在期望与方差。

相应的实值奖励函数序列：

其中常数为每次观察所需的代价，令是表示一次观察都不做的收益是无穷小，因而至少需要进行一次观察，表示若从不停止观察则所需的代价为无穷大，因而收益为负无穷大。

### 3.2.1 证明最优停止策略的存在性

由已知文献可知，对于待观察随机序列和实值奖励序列存在最优停止策略的充分条件为以下两个条件：

A1：

A2：

由存在知，可得

又由存在知

由此可得，该问题背景下满足条件A1和A2，因此最优停止策略存在，使得观察者以尽可能大的概率选择出最大收益。**（证明过于简陋，因为这块还不是很懂）**

### 3.2.2 最优停止策略的描述

首先观察问题，在任意时刻选择继续观察表示放弃获取当前的回报，并且支付观察代价，接着观察到并面临与时刻时一样的处境。因此可以说明，在i时刻选择继续观察与在时刻选择继续观察的收益期望是相同的，故可以假设为最优停止策略中任意时刻选择继续观察时的收益期望。由此可得选择策略为：

1. 当时刻t时的观察值时，表明此时选择继续所能获得的回报期望大于选择停止获得的回报，可以认为选择继续观察会有更大的概率获取更高的收益，因此选择继续观察。
2. 当时刻t时的观察值时，表明此时选择继续所能获得的回报期望不大于选择停止获得的回报，因此选择停止观察接受回报。

因此由最优性原理得出的最佳停止策略为在处停止，其中N应当满足如下条件：

因此现在的问题只是计算出收益期望，而由定义为最优停止策略中任意时刻选择继续观察时的收益期望可得出在观察时获得的收益期望为：

又知=，故有

也即

又有

其中是所服从分布的密度函数，故将上式代入可得

也即

又有

故上述方程可化为

化简得

也即

至此，观察上式可知收益期望的取值既与待观察随机样本所服从的分布也与观察代价有关。下文将通过两个具体的例子算出均匀分布与指数分布下的收益期望。

### 3.2.3 均匀分布下的选择规则

案例一：假如观察样本独立同分布且服从区间的均匀分布，也即

则有

故当时，

化简得

当时，

化简得

综上得

在该案例中，当时，选择方案为在第一次观察到时选择停止；当时，选择方案为在第一次观察到时选择停止。

### 3.2.3 指数分布下的选择规则

案例二：假如观察样本独立同分布且服从参数为的指数分布，也即

则有

故当时，

化简得

故当时，

化简得

综上得

在该案例中，当时，选择方案为在第一次观察到时选择停止；当时，选择方案为在第一次观察到时选择停止。

## 3.3 有失败回弹风险

在观察消耗成本且有失败回弹风险情况下选取最大收益的情形下的最优停止问题可以描述为如下问题：

待观察的随机序列为：

其中相互独立且服从同一分布，且分布存在期望与方差；也相互独立且服从伯努利分布，且。

在这个情况当中，表示观察者在第i次选择观察所能获得的回报，是一个指示器函数，表示在i时刻观察者是否成功。只要失败未发生（也即的值一直为1），则观察者每次观察所获得的回报将会累加。但是当失败发生时（也即）时，则回报值将会重新降低为0。回报值降为0后观察者可以选择继续观察从而继续累加收益，直到下次发生失败时。

设为观察者在时刻i所能获得的总回报值，则有

也即观察者在时刻能获得的总回报值等于第时刻的总回报值加上第时刻的回报值再乘以成功概率。

相应的实值奖励函数序列：

其中常数为每次观察所需的代价，令表示一次观察都不做时的初始奖励，表示若从不停止观察则所需的代价为无穷大，因而奖励为负无穷大。

### 3.3.1 证明最优停止策略的存在性

由已知文献可知，对于待观察随机序列和实值奖励序列存在最优停止策略的充分条件为以下两个条件：

A1：

A2：

由问题定义，可得

又由存在知

由此可得，该问题背景下满足条件A1和A2，因此最优停止策略存在，使得观察者以尽可能大的概率选择出最大收益。**（证明过于简陋）**

### 3.3.2 最优停止策略的描述

首先观察问题，在任意时刻选择继续观察表示以一定概率回报降低为0的风险获取更多的回报，并且支付观察代价，接着观察到并面临与时刻时一样的处境。因此可以说明，在i时刻选择继续观察与在时刻选择继续观察的收益期望是相关的，故可以定义为当初始奖励为时奖励值的期望，相应的为当n时刻时的奖励值为r时的期望，且可认为与满足如下关系：

由此可得到一个不错的选择策略：

在时刻t时观察到后计算t时刻的回报与奖励，

若满足条件也即时，则选择停止，此时选中的回报值大于以初始奖励为时奖励值的期望。由此可得，在选择停止时满足条件，因此即为按该选择策略所能选中的回报值的下确界。

现已知是关于变量回报值的函数，因此该选择策略问题可迭代为寻找实值，使得以下两条件得以成立：

（1）满足的同时使得尽可能大（取到极大值）；

（2）当时，。

假如当时满足也即，则对于任意时刻时满足条件的回报值，都满足如下条件：

当第一次失败发生的时间在之前时（），则

因此，

当第一次失败发生的时间在之后时（），则

因此，

(上述推导中是初始奖励为在时间时的奖励值，是初始奖励为在时间时的回报值)

由此可得

又由于

因此可得

又由假设已知

（与意义相同，均表示以初始奖励为的奖励期望）

故可得

至此，已证明若

则

由最优性原理得出的最佳停止策略可进一步迭代为在观察到处时停止，其中N应当满足如下条件：

其中，满足条件的同时，使得期望尽可能大。

因此现在的问题只是计算出奖励期望，而由被定义为在最优停止策略中，以初始奖励为为前置条件在观察者选择停止时所能获得的奖励期望可得出：

其中选择停止时间受初始奖励值的影响，为关于的函数。

### 3.3.3 指数分布下的选择规则

案例：假如观察样本独立同分布且服从参数为的指数分布，也即

由上一节可知在满足条件且初始时的情况下，奖励期望满足如下等式：

其中，由于在时做出了停止的决策，因此条件应被满足。故而有

又由且可知，在是已知常数。故而有

其中，服从指数分布，由指数分布的无记忆性可得：

计算出后若想计算的值，则还需计算出的表达式。要计算则需考虑两种情况，第一种情况为在回报值累加至之前失败从未发生，第二种情况为在累加至前发生过失败使得曾降低为0。

因此，假设在第一种情况下，时选择停止，其中停止时刻的表达式如下：

其中，为前个观察样本的累加和。

由于是服从参数为的指数分布，因此当时-1服从参数为的泊松分布，也即

又令表示第一次失败发生的时间，则可用下式表示：

其中表示选择停止时间在第一次失败发生时间之前的条件期望，则是停止时间在第一次失败发生时间之后的调价期望。

由于服从参数为的伯努利分布，因此第一次失败发生的时间服从几何分布：

接下来可以发现，若停止时间在第一次失败发生时间之后，则在时间为时失去全部已积累的回报值，与初始时相同。因此，当停止时间在第一次失败发生时间之后时的条件期望应满足：

若停止时间在第一次失败发生时间之前，则表面在回报累加到前从未发生过失败，因而在时间为时有：

至此，可将与代入的表达式中得到：

故

也即得到的表达式：

最后，计算与即可，其中

而可由以下推导得出：

此时先计算得

因此，将以上推导结果代入可得：

至此将与代入得

再将与代入得

其中和均为已知常数，为未知数。

将对求导得

故当取得极大值时，

其中，因此。而当时，

此时是关于的递减函数，故若要使得尽可能大，的取值应为。

综上得

在该案例中，当时，选择方案为在第一次观察到时选择停止；当时，选择方案为在第一次观察到

时立即选择停止。

# **4 算法验证**

本章将根据上一章描述的选择策略设计实现算法。并通过多次重复独立仿真实验探索与验证不同选择算法的性能表现。

## 4.1 不可回溯情况下选取最大值问题的算法与实验

该情形下待观察的随机序列服从存在期望与方差的同一分布，奖励函数为。下文将设计算法程序复现该情形下的选择策略，并通过实验分析验证该算法的性质。

### 4.1.1 算法设计

在前大约的数据中，只观察不停止，并且实时记录更新观察到的最大值。从大约的数据开始，每次获取新观察对象时都与上一过程记录的最大值进行比较，若实时最大值被更新，也即当前实时观察值大于前数据中的最大值时，最新观察值为观察到目前为止的最大值，此时选择停止并选中该值。根据上一章中推理结论，该选中值为全局最大值的概率大约为。伪代码如下：

|  |
| --- |
| 算法1：不可回溯的最大值在线选取算法 |
| 1. INITIALIZE： 当前最大值初始化为无穷小 2. 已知数据量大小 3. 将观察对象接入在线数据流 4. WHILE 获取新观察对象 DO 5. IF 当前观察对象序号总观察对象数0.37 6. ，当前观察对象 7. ELSE IF当前观察对象序号总观察对象数0.37 当前观察值 8. 当前观察对象 9. 选择完成 程序停止 10. ELSE 11. PASS 12. END IF 13. END WHILE |

### 4.1.2 实验方案

实验代码采用Python语言编写，用Random随机数生成服从均匀分布的随机数序列模拟在线数据流。以该数据流作为在线观察序列进行千次重复独立实验，每次实验观察对象的数据量在每组之间，以尽可能模拟现实情形中的随机状况。每次实验选择完成以后，将组内所有观察值进行从大到小排序，记录实验中选中观察值的大小以及在所有观察值中所在的位次。

### 2.1.3 实验指标

最大值命中率：选中的观察值在从大到小排序后排名第一的次数与重复独立实验次数的比值。

前命中率：选中的观察值在从大到小排序后排名前的次数与重复独立实验次数的比值。

前命中率：选中的观察值在从大到小排序后排名前的次数与重复独立实验次数的比值。

### 4.1.4 实验结果

实验结束后对实验数据进行统计分析得到如下实验结果：

表1**：不同数据量级下各项命中率指标的实验值**

|  |  |  |  |
| --- | --- | --- | --- |
| 观察样本数量 | 最大值命中率 | 前命中率 | 前命中率 |
| 100 | 39% | 39% | 68% |
| 1000 | 37.4% | 63.1% | 66.7% |
| 10000 | 36.9% | 63.4% | 67.3% |
| 随机 | 36.7% | 62.8% | 65.2% |

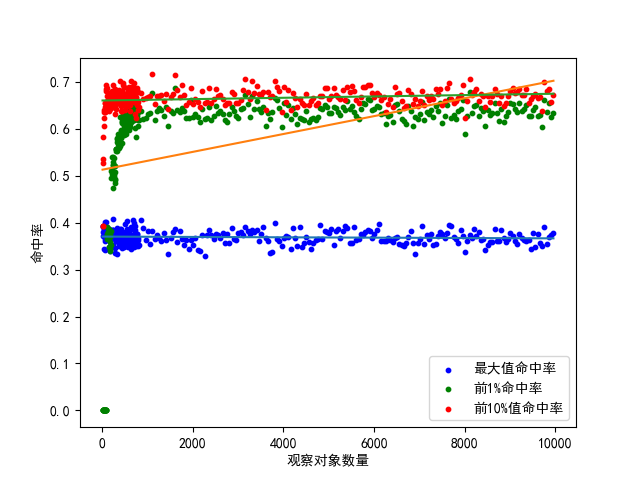


图4.1 各项命中率指标-观察样本数量曲线图

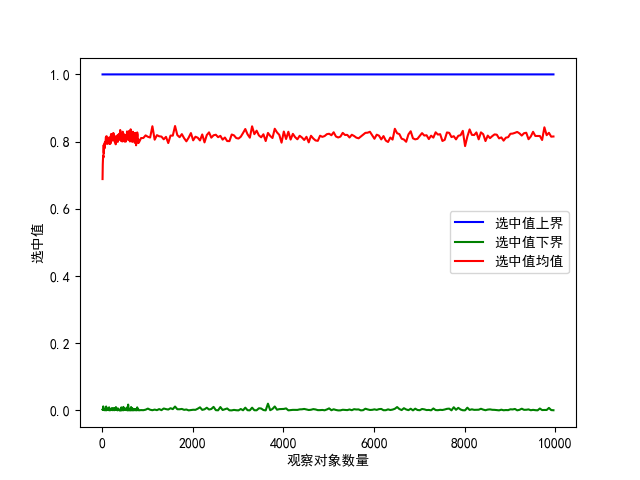


图4.2 选中值上界、下界、均值-观察样本数量曲线图

总结实验结果得：1、最大值命中率与理论值大致相符。2、条件不苛刻情况下，当观察样本量大于500时前与前命中率稳定较高，命中率稳定在以上，可见该策略有较大概率选到较为上乘的结果。3、随着数据量的增大，命中率逐渐收敛到一定值且趋于稳定。4、千次重复独立实验中，命中值的均值在观察样本达到一定规模后趋于稳定，稳定在附近（其中为期望，为均匀分布的下界）。

## 4.2 不可回溯且观察消耗成本的选择算法与实验

该情形下待观察的随机序列服从存在期望与方差的同一分布，奖励函数为，其中常数为每次观察所需的代价。下文将使用Python复现该情形下的选择策略，并通过实验分析验证该算法的性质。

### 4.2.1 算法设计

在均匀分布模型中，每一个观察对象都需要进行归一化处理使得观察样本服从均匀分布，而观察成本也需按归一化处理等比例进行放缩。标准化数据后再进行数值计算与判断决策。

当非负观察成本时，判断阈值为。对于每一个标准化的观察对象，如果大于或者等于阈值，则记录受益值，否则继续运行程序，循环上述操作。

当观察成本时，判断阈值为。对于每一个标准化的观察对象，如果大于或者等于阈值，则记录受益值，否则继续运行程序，循环上述操作。

在指数分布模型中，每一个观察对象都需要进行归一化处理至观察样本服从分布，观察成本同样除以原指数分布期望值进行等比例进行放缩。标准化数据后再进行后续的选择实验。

当非负观察成本时，判断阈值为。对于每一个标准化的观察对象，如果大于或者等于阈值，则记录受益值，否则继续运行程序，循环上述操作。

当观察成本时，判断阈值为。对于每一个标准化的观察对象，如果大于或者等于阈值，则记录受益值，否则继续运行程序，循环上述操作。

上述算法伪代码如下：

|  |
| --- |
| 算法2：不可回溯且观察消耗成本的在线选取算法（均匀分布模型） |
| 1. INITIALIZE： 当前最大收益初始化为无穷小 2. 观察样本服从的初始均匀分布的期望 4. 将观察对象接入在线数据流 5. IF 7. ELSE 9. END IF 10. WHILE 获取新观察对象 DO 11. 观察对象 13. IF 14. 停止观察并选取当前受益实值 15. ELSE 16. PASS 17. END IF 18. END WHILE |

|  |
| --- |
| 算法3：不可回溯且观察消耗成本的在线选取算法（指数分布模型） |
| 1. INITIALIZE： 当前最大收益初始化为无穷小 2. 观察样本服从的初始指数分布的系数 4. 将观察对象接入在线数据流 5. IF 7. ELSE 9. END IF 10. WHILE 获取新观察对象 DO 11. 观察对象 13. IF 14. 停止观察并选取当前受益实值 15. ELSE 16. PASS 17. END IF 18. END WHILE |

### 4.2.2 实验方案

用random随机数仿真标准化处理后的数据，使得观察值均服从均匀分布或者均服从指数分布，并且将随机数序列作为在线数据流。以该数据流作为在线观察序列进行千次重复独立实验，每次独立实验的观察对象数量服从均匀分布，使得仿真环境尽可能符合在线大数据量情形。每次实验完成后，记录选择后得到的收益值，然后计算组内所有观测值被选择的理论受益值，并且对所有收益值进行排序，再记录实验选中收益值在所有收益值中所在的位次。

### 4.2.3 实验指标

选择收益值：实验中选中的收益值。

表示选择收益值的上界，表示选择收益值的下界，表示选择收益值的均值。

最大值命中率：选中的收益值在从大到小排序后排名第一的次数与重复独立实验次数的比值。

前命中率：选中的收益值在从大到小排序后排名前的次数与重复独立实验次数的比值。

前命中率：选中的收益值在从大到小排序后排名前的次数与重复独立实验次数的比值。

### 4.2.4 实验结果

均匀分布情形实验结果：

在千次随机数据量的独立实验中，不同观察成本值的情况下，各项命中率指标也各不相同。

表2**：不同观察成本值各项命中率指标的实验值**

|  |  |  |  |
| --- | --- | --- | --- |
| 观察成本值 | 最大值命中率 | 前命中率 | 前命中率 |
| 0.0001 | 69.5% | 99% | 100% |
| 0.001 | 68.6% | 99.8% | 100% |
| 0.01 | 70.2% | 99.8% | 100% |
| 0.1 | 74.1% | 99.6% | 100% |
| 0.5 | 86.6% | 99.7% | 100% |
| 0.7 | 95.5% | 100% | 100% |

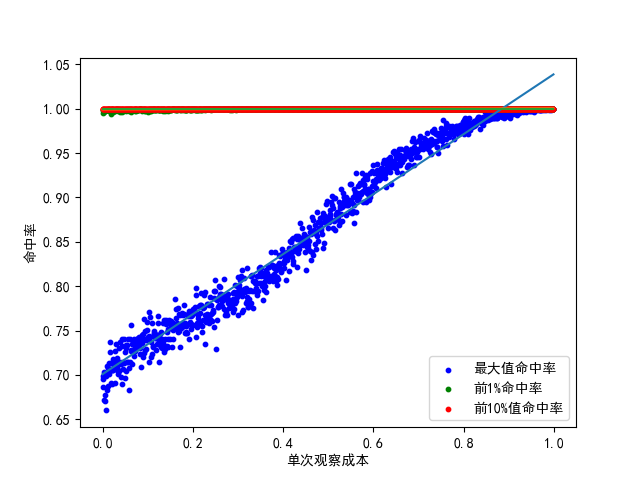


图4.3 各项命中率指标-观察成本值曲线

在千次随机数据量的独立实验中，选择收益值与观察成本值之间的关系如下图所示。

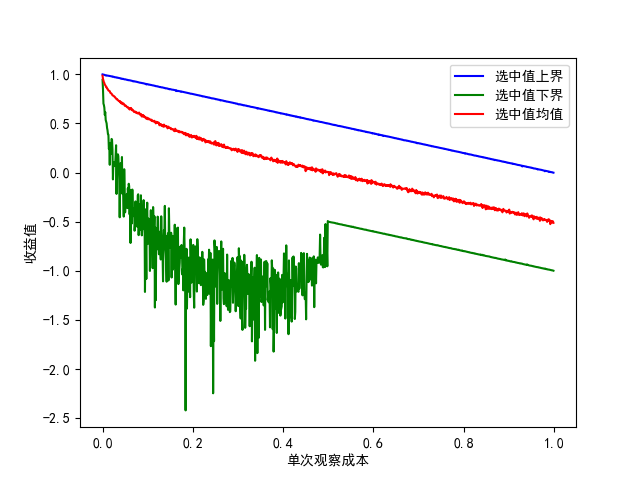


图4.4 收益值上界、下界、均值-观察成本值曲线

总结实验结果得：在观察对象服从均匀分布时：1、最大收益命中率随观察成本变大而变大且大致。2、前命中率与前命中率在时几乎稳定为。3、观察成本值时，收益下限随的增大总体呈减小后增大的趋势，但极其不稳定，时，收益下限随观察成本增大而减小且呈线性变化。4、选择收益上界与观察成本值的关系符合收益上界随单调减小的一次函数。5、选择收益的均值随观察成本的增加大致呈减小趋势，且减小速率也随增大而减小至趋于一定值。6、当时该策略下的收益均值大概率为正数，时收益均值为负值。

指数分布情形实验结果

在千次随机数据量的独立实验中，不同观察成本值的情况下，各项命中率指标也各不相同。

表3**：不同观察成本值各项命中率指标的实验值**

|  |  |  |  |
| --- | --- | --- | --- |
| 观察成本值 | 最大值命中率 | 前命中率 | 前命中率 |
| 0.0001 | 11% | 11.3% | 11.3% |
| 0.001 | 52.1% | 69.7% | 71.6% |
| 0.01 | 54.7% | 96.1% | 99.6% |
| 0.1 | 55.9% | 96.8% | 99.9% |
| 0.5 | 57.1% | 99.7% | 100% |
| 0.7 | 61.4% | 100% | 100% |

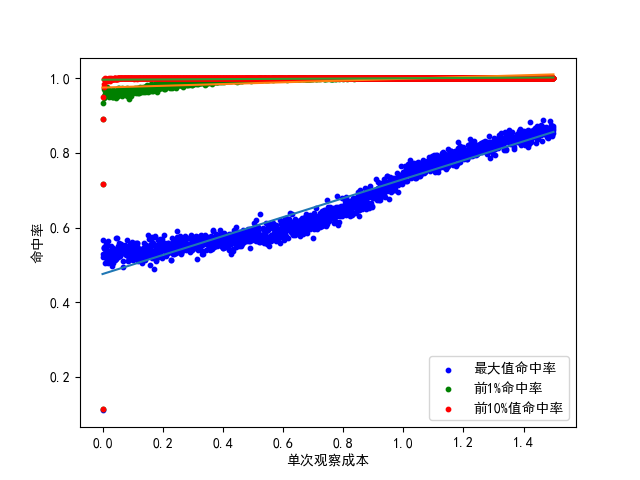


图4.5 各项命中率指标-观察成本值曲线

在千次随机数据量的独立实验中，选择收益值与观察成本值之间的关系如下图所示。

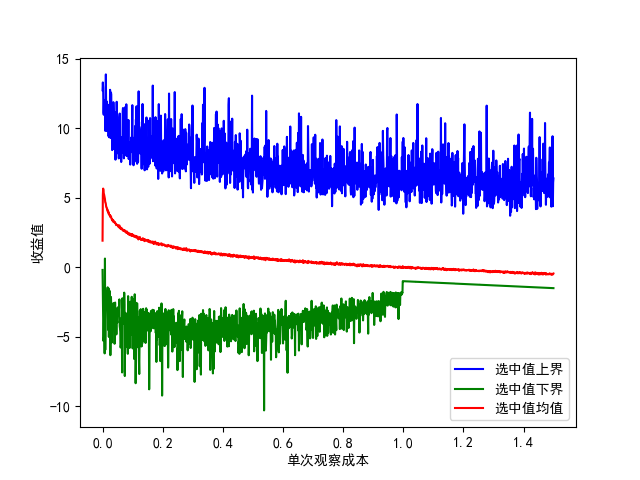


图4.6 收益值上界、下界、均值-观察成本值曲线

总结实验结果得：在观察对象服从指数分布时：1、时最大收益命中率随观察成本变大而变大，与时，最大收益命中率随观察成本增长较为缓慢，时最大收益命中率随观察成本增长相对较快，时，最大值命中率大概率大于。2、前命中率与前命中率在时稳定且可达以上。3、观察成本值时收益下限随观察成本先变小后变大，时收益下限与观察成本呈线性递减的关系，时收益值下限可取到极大值。5、收益值上界随增大总体上呈递减趋势，但方差较大。6、选择收益均值由于样本数量较大因而方差极小大致符合光滑曲线，在时，选择收益均值随的增大先增大后减小且在之间取得极大值。6、时，收益均值不小于。（为指数分布的期望，即）

## 4.3 有失败回弹风险的算法验证

该情形下待观察的随机序列服从存在期望与方差的同一分布，也相互独立且服从伯努利分布。回报函数为，其中常数为每次观察所需的代价。下文将设计算法程序复现该情形下的选择策略，并通过实验分析验证该算法的性质。

### 4.3.1 算法设计

在指数分布模型中，每一个观察对象都需要进行归一化处理使得观察样本服从均匀分布，而观察成本也需按归一化处理等比例进行放缩。标准化数据后再进行数值计算与判断决策。

在每次观察新的随机对象与时，先更新实时回报值，然后再根据伯努力分布的参数采用不同的判断方案。

当非负观察成本时，判断阈值为。对于每一个标准化的观察对象，如果大于或者等于阈值，则记录受益值，否则继续运行程序，循环上述操作。

当观察成本时，判断阈值为。对于每一个标准化的观察对象，如果大于或者等于阈值，则记录受益值，否则继续运行程序，循环上述操作。

上述算法伪代码如下：

|  |
| --- |
| 算法4：有失败回弹风险的在线选取算法 |
| 1. INITIALIZE： 当前回报值初始化为0 3. 为伯努利分布参数 4. 将观察对象接入在线数据流 5. IF 7. ELSE 9. END IF 10. WHILE 获取新观察对象 DO 12. IF 13. 停止观察并选取当前受益实值 14. ELSE 15. PASS 16. END IF 17. END WHILE |

### 4.3.2 实验方案

用random随机数模拟两组随机数，第一组仿真标准化处理后的收益数据，使得观察值均服从指数分布，另一组生成0-1数据流使得的概率为，并且将两组随机数序列作为在线数据流。以该数据流作为在线观察序列进行千次重复独立实验，每次独立实验的观察对象数量服从均匀分布，使得仿真环境尽可能符合在线大数据量情形。每次实验完成后，记录选择后得到的选择奖励值，然后依次计算组内所有选择方案的理论回报值与相应的奖励值，并且对所有选择方案的奖励值进行排序，再查看并记录实验选择所获奖励值在所有奖励值中所在的位次。

### 4.3.3 实验指标

选择奖励值：实验中算法选择所获的奖励值。

表示选择奖励值的上界，表示选择奖励值的下界，表示选择奖励值的均值。

最大值命中率：选择奖励值在从大到小排序后排名第一的次数与重复独立实验次数的比值。

前命中率：选择奖励值在从大到小排序后排名前的次数与重复独立实验次数的比值。

前命中率：选择奖励值在从大到小排序后排名前的次数与重复独立实验次数的比值。

### 4.3.4 实验结果

在千次随机数据量的独立实验中，不同观察成本值或不同伯努利实验成功率的情况下，各项命中率指标也各不相同。

表4**：不同观察成本值不同伯努利实验成功率各项命中率指标的实验值**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| c | p | 最大值命中率 | 前命中率 | 前命中率 |
| 0.0001 | 0.1 | 9.9% | 12.4% | 20..5% |
| 0.0001 | 0.5 | 4.4% | 8.9% | 17.1% |
| 0.001 | 0.8 | 5.1% | 23.9% | 30.4% |
| 0.004 | 0.1 | 44.9% | 95.7% | 98.6% |
| 0.004 | 0.5 | 20.7% | 92% | 94.4% |
| 0.004 | 0.8 | 10% | 61% | 71.5% |
| 0.01 | 0.1 | 47.1% | 94.8% | 97.6% |
| 0.3 | 0.5 | 27.4% | 98.3% | 100% |
| 0.8 | 0.8 | 30.1% | 76.8% | 100% |

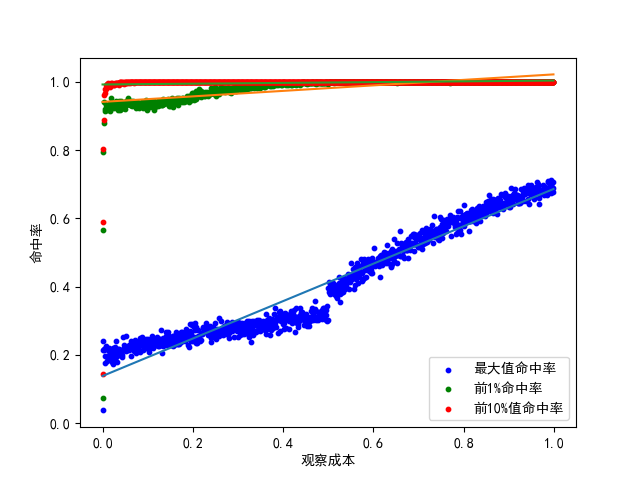


图4.7 时各项命中率指标-观察成本值曲线

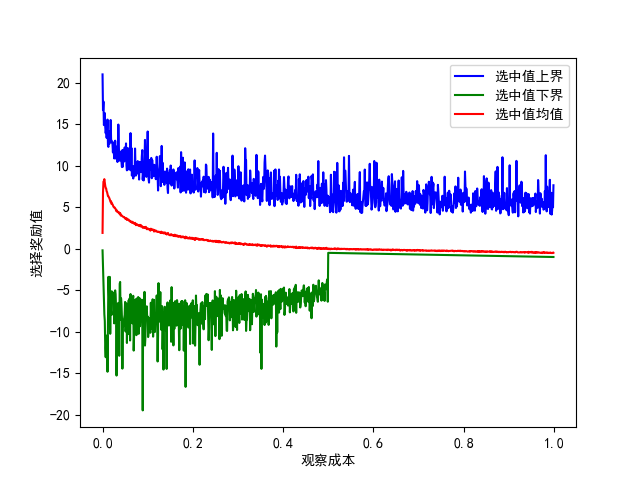


图4.8 时选择奖励值上界、下界、均值-观察成本曲线

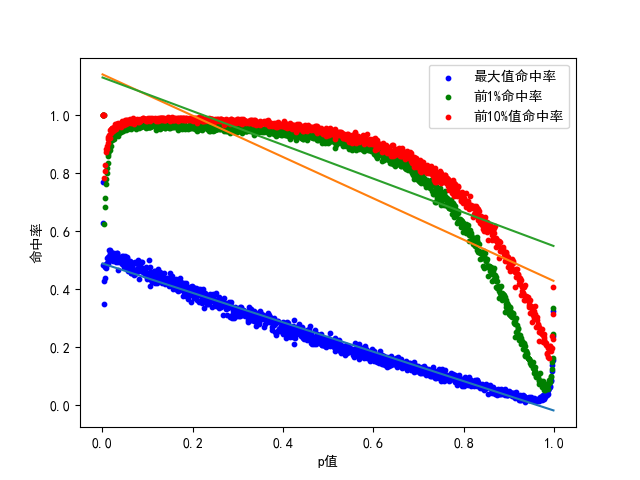


图4.9 时各项命中率指标-伯努利实验成功率曲线

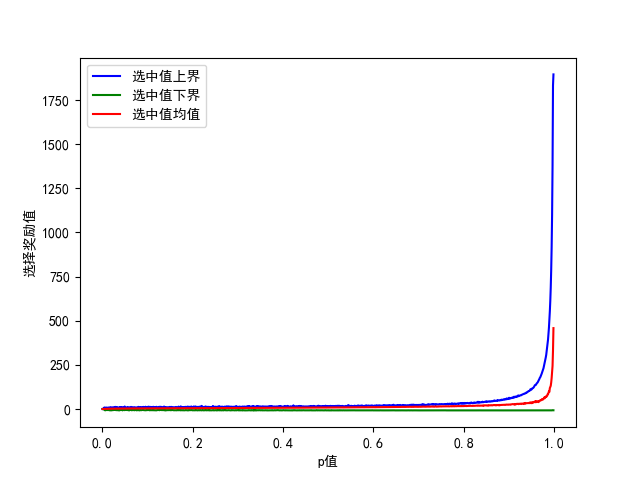


图4.10 时选择奖励值上界、下界、均值-伯努利实验成功率曲线

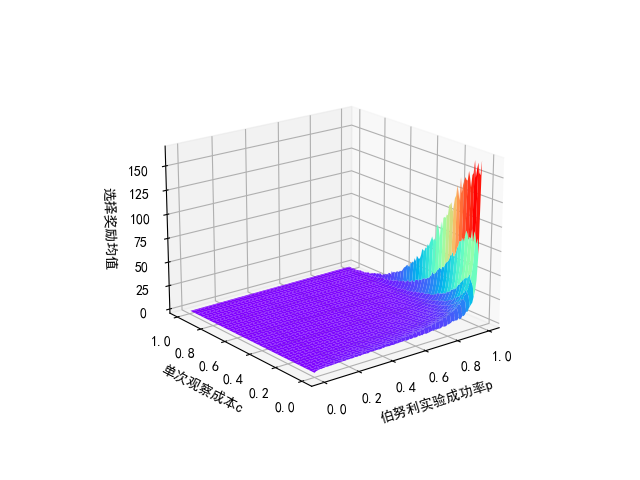


图4.11 选择奖励均值-伯努利实验成功率-观察成本关系图

总结在观察对象服从指数分布时的实验结果得：1、该选择算法的最大值命中率随观察成本与伯努利实验成功率变化而变化，且与观察成本呈正相关关系，与伯努利实验成功率呈负相关关系。2、条件较弱的前命中率与前命中率随观察成本与伯努利实验成功率的变化趋势与最大值命中率相同，但在时命中率随增大而下降的速率增大。3、当一定时，该选择算法下选择奖励值的均值随观察成本先增大后变小，并且在附近取得极大值，在时奖励均值大约为，在时奖励均值随观察成本增大而下降的速率越来越小。4、当一定时，该选择算法下选择奖励值的上界方差较大，随观察成本增大呈减小趋势，选择奖励值的下界在时方差较大，随观察成本呈先减小后增大趋势，在时约等于，方差极小且随增大而缓慢减小。5、当一定时，该算法下的选择奖励均值与上界方差极小，随伯努利实验成功率增大而增大，增长速率也随值增大而增大，且当趋近于时奖励均值与上界趋近与无穷大，且当时，该算法下的选择奖励下界值始终约等于0。

# 结论（黑体小二号）

**（黑体小二号字体，居中）**

结论是毕业论文的总结，是整篇论文的归宿。应精炼、准确、完整。着重阐述自己的创造性成果及其在本研究领域中的意义、作用，还可进一步提出需要讨论的问题和建议。

**（格式：宋体，Times New Roman小四号字，两边对齐，首行缩进2个字符，行距23磅，字符间距为“标准”）**

# 参考文献（黑体小二号）

**（黑体小二号字体，居中）**

**（各篇文献按正文部分标注的序号依次列出）（格式：宋体+新罗马小四号字，两端对齐，行距23磅，悬挂缩进1.27 cm。）**

**范例：**

1. Weissmüller J. Alloy effects in nanostructures[J]. Nanostructured Materials, 1993, 3(1-6): 261-272.
2. Mazzolani F. Aluminium alloy structures[M]. CRC Press, 1994.
3. Fu S H, Dong J X, Zhang M C, et al. Alloy design and development of INCONEL718 type alloy[J]. Materials Science and Engineering: A, 2009, 499(1-2): 215-220.

**要求：**

按论文中出现的顺序列出直接引用的主要参考文献，其注释一律采用尾注形式，以阿拉伯数字顺序编码，数字外加方括号“[]”。可列于参考文献表的文献类型包括图书、期刊、会议论文集和学位论文等。作者引用的图书、期刊等名称要完整、清楚，不提倡缩写。要求按顺序准确标明：作者，书（篇）名，出版地，出版社，出版时间和页码。参考文献的著录应符合国家标准（以GB/T 7714-2015为标准），参考文献的序号左顶格，并用数字加方括号表示，与正文中的引文标示一致，如[1]，[2]……。每一条参考文献著录均以“.”结束。具体各类参考文献的编排格式**(注意：只是参考文献格式要求，具体排版仿照上方范例)**如下：

1. **著作**

著录格式：[序号] 编著者.书名[M].版本，出版地：出版社，出版年.

例：[1] 张伯伟.全唐五代诗格会考[M].南京：江苏古籍出版社，2002.

[2] CRAWFPRD W, GORMAN M. Future libraries: dreams, madness, & reality[M]. Chicago: American Library Association, 1995.

1. **论文集、会议录**

著录格式：[序号] 编著者.题名[G/C].出版地：出版社，出版年.

例：[1] 中国职工教育研究会.职工教育研究论文集[G].北京：人民教育出版社，1985.

[2] BABU B V, NAGAR A K, DEEP K, et al. Proceedings of the Second International Conference on Soft Computing for Problem Solving, December 28-30, 2012[C]. New Delhi: Springer, 2014.

1. **报告**

著录格式：[序号] 报告撰写者.报告题名及编号[R/OL].出版地：出版社，出版年.

例：[1] U.S. Department of Transportation Federal Highway Administration. Guidelines for handling excavated acid-producing material: PB 91-194001[R]. Springfield: U.S. Department of Commerce National Information Service, 1990.

1. **学位论文**

著录格式：[序号] 学位论文撰写者.题名[D].地点：单位，年.

例：[1] CALMS R B. Infrared spectroscopic studies on solid oxygen[D].Berkeley: Univ. of California, 1965.

1. **专利**

著录格式：[序号] 专利申请者.专利题名：专利号[P].公告日.

例：[1] 张凯军.轨道火车及高速轨道火车紧急安全制动辅助装置：201220158825.2[P].2012-04-05.

1. **标准文献**

著录格式：[序号] 标准提出者：标准编号[S].出版地：出版社，出版年.

例：[1] 全国信息与文献标准化技术委员会.文献著录：第4部分 非书资料：GB/T 3792.4—2009[S]. 北京：中国标准出版社，2010.

1. **专著中析出的文献**

著录格式：[序号] 析出文献的著者.书名[M]//版本.出版地：出版社，出版年：页码.

例：[1] 白农书. 植物开花研究[M]//李承森.植物科学发展.北京：高等教育出版社，1998：146-163.

1. **期刊中析出的文献**

著录格式：[序号] 析出文献的著者.题名[J].刊名，年，卷（期）：页码.

例：[1] 杨洪升.四库馆私家抄校书考略[J].文献，2013(1)：56-75.

[2] SAITO M, MIYAZAKI K. Jadeite-bearing metagabbro in serpentinite mélange of the “Kurosegawa Belt” in Izumi Town, Yatsushiro City, Kumamoto Prefecture, central Kyushu[J]. Bulletin of the geological survey of Japan, 2006, 57(5/6):169-176.

1. **报纸中析出的文献**

著录格式：[序号] 析出文献的著者.文献题名[N].报纸名，出版日期(版次).

例：[1] 丁文祥. 数字革命与竞争国际化[N].中国青年报，2000-11-20（15）.

1. **电子资源**

著录格式：[序号] 编著者.电子资源题名[EB/OL].（发表日期）[更新日期].获取和访问路径.数字对象唯一标识符.

例：[1]萧钰.出版业信息化迈入快车道[EB/OL].（2001-12-19）[2002-04-15].http://www.creader.com/news/20011219/200112190019.html.

# 附 录（黑体小二号）

**（黑体小二号字体，居中）**

论文附录依次用大写字母“附录A、附录B、附录C……”表示，附录内的分级序号可采用“附A1、附A1.1、附A1.1.1”等表示，图、表、公式均依此类推为“图A1、表A1、式（A1）”等。包含以下内容：

（1）代码、图表、标准、手册等数据；

（2）未发表过的一手文献；

（3）公式推导与证明、调查表等；

（4）辅助性教学工具或表格；

（5）其他需要展示或说明的内容

……

（标题黑体小二号，内容Times New Roman/宋体，小四号，行距20磅）

# 致 谢

时值此论文完成之际，首先非常感谢戴东波老师给予我从论文定题之初到论文写作之时孜孜不倦的指导。戴老师一直以来对学术研究执着求真的态度、丰富的学科知识、辩证的思维方式与负责的工作态度使我受益匪浅。在整个论文的定题、行文的思路以及论文修改过程中提供了十分有用的帮助。我将牢记戴老师的教诲，在以后的学习生活中认真负责、努力拼搏、究其原理，超越之前的自己。

其次我要感谢我的家人，感谢我的家人对我二十多年的悉心照顾，让我健康快乐的长大。感谢家人对我无微不至的关心，让我即使远在千里之外也能时时感受家的温暖。感谢家人尊重和支持我的每一个决定，让我在做出转专业、考研、找工作的决心时都有莫大的底气。感谢家人的安慰与鼓励，在我陷入焦虑、彷徨、自我怀疑之时分担我的忧愁，给予我鼓励，使我重获信心与动力。没有家人们背后的支持和帮助，我就不能实现任何想实现的目标。以后我会更加努力，用成绩回报家人的付出。

同时我还要感谢我的朋友沈颖，感谢她陪我度过了本科生涯中最难过的一段时光。大四的上半学年是我本科生涯中最迷茫，最困惑，最焦虑的时期。我站在人生的岔路踌躇不知该将脚步迈向何方。人生中重大的决定常常只源于最简单草率的原因，仅是年少的爱慕，落在秤托上，便让我内心的天平完全倾倒。我就这样在秋招的开始之际，纵身跃进了考研的苦海当中。但我很幸运，备考的过程并非我独自孤军奋战，因为一直有沈颖在陪伴。当我压抑焦虑难以喘息之时，她会耐心听我倾倒苦水并给予我鼓励。当我做出更换报考院校的决断时，她的一句“我觉得你做的这个决定挺好的”让我的颤抖的内心多了一分底气。当我对未来感到困惑陷入自我怀疑时，她会分享她的见解看法，认真分享我的处境情况，帮助我一步步认识自己，引导我逐步自我解惑。半年的时间只是本科生涯的八分之一，但却让我发生了翻天覆地的变化。我要感谢沈颖，感谢她在这段时光中的陪伴，感谢她帮助我在感情生活中逐步走向成熟，感谢她帮助我理解读研的于自己而言的深层意义，感谢她帮助我更加深刻的认清自我、完善自我。

祖睿

上海大学图书馆

2024年05月10日

