经济学视角下的随机服务系统

孙　微　李世勇　著

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内 容 简 介

本书对经济学视角下的随机服务系统这门学科进行了初步介绍，研究了经济学视角下的随机服务系统中有限理性主体的博弈和决策行为。本书重点介绍了几类经济学视角下的随机服务系统模型的建模求解过程及结果和意义，包括具有启动-关闭策略的排队系统、具有多重休假策略的排队系统、具有工作休假策略的排队系统、具有中途退出行为的排队系统、具有部分服务信息的排队系统和具有随机服务参数的排队系统。

本书可供行为运筹及管理科学领域的研究人员、高校教师和相关专业研究生使用，也可供随机服务企业的管理人员进行阅读参考之用。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

经济学视角下的随机服务系统/孙微，李世勇著. —北京：电子工业出版社，2017.1
ISBN 978-7-121-30425-5

Ⅰ. ①经… Ⅱ. ①孙… ②李… Ⅲ. ①排队论—研究 Ⅳ. ①O226

中国版本图书馆 CIP 数据核字（2016）第 286791 号

策划编辑：王二华
责任编辑：王二华
印　　刷：
装　　订：
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　邮编 100036
开　　本：720×1000　1/16　印张：16.75　字数：420 千字
版　　次：2017 年 1 月第 1 版
印　　次：2017 年 1 月第 1 次印刷
定　　价：49.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：（010）88254532。

作 者 简 介

孙微，博士，副教授，燕山大学经济管理学院教师，先后主持国家自然科学基金青年项目1项，教育部人文社会科学项目1项，中国博士后科学基金面上项目1项（一等资助），河北省自然科学基金优秀青年基金1项，河北省自然科学基金青年基金1项，以第一作者/通信作者身份在多个国际高水平学术期刊上发表学术论文40余篇，包括 *European Journal of Operational Research*、*Asia-Pacific Journal of Operational Research*、*Performance Evaluation*、*International Journal of Systems Science*、*Electronic Commerce Research*、*Applied Mathematics and Computation*、*Top*、*Applied Mathematical Modelling*、*4OR: A Quarterly Journal of Operations Research*、*Central European Journal of Operations Research*、*Quality Technology & Quantitative Management* 等国际学术期刊，入选河北省第二批青年拔尖人才支持计划及河北省高等学校青年拔尖人才支持计划。目前担任中国运筹学会行为运筹与管理分会理事。

李世勇，博士，副教授，燕山大学经济管理学院教师，先后主持国家自然科学基金面上项目1项，国家自然科学基金青年项目1项，中国博士后科学基金特别资助项目1项，教育部高等学校博士学科点专项科研基金1项，教育部人文社会科学项目1项，中国博士后科学基金面上项目1项，河北省自然科学基金1项，以第一作者/通信作者身份在多个国际高水平学术期刊上发表学术论文50余篇，包括 *Electronic Commerce Research*、*Performance Evaluation*、*International Journal of Systems Science*、*International Journal of Communication Systems*、*Informatica*、*Asia-Pacific Journal of Operational Research*、*Applied Mathematics and Computation*、*Applied Mathematical Modelling*、*Top*、*Quality Technology & Quantitative Management* 等国际学术期刊，入选河北省第二批青年拔尖人才支持计划、河北省高等学校青年拔尖人才支持计划、河北省“三三三人才工程”第三层次人选及河北省宣传文化系统“四个一批”人才工程。

前　言

自从 1909 年丹麦工程师 Erlang 开创了排队论学科以来，经典排队论已经经历了逾百年的历史，并已发展成为运筹学和应用概率论中的重要分支学科。排队论又名随机服务系统理论，经典随机服务系统理论到目前为止已经基本确立了成熟的理论体系，并在生产、交通、服务、通信、管理和军事等领域得到了广泛的应用。

然而，在经典随机服务系统理论的研究过程中，顾客和服务员在排队和提供服务的过程中往往被视为完全理性主体，而不被视为主观能动的有限理性主体，不能通过权衡自身的利益得失来主动地做出各类决策。然而，现实之中系统主体的偏好类型及信息环境等都会影响他们的决策行为。当随机服务系统中的顾客和服务员被视为有限理性主体时，出于自身利益最大化的目的，他们内部之间势必会产生非合作博弈行为，这种博弈行为的结果就是各自进行均衡决策。因此，从合理性和实用性的角度看，对有限理性主体的博弈和决策行为进行分析的重要性不言而喻。

20 世纪 60 年代末，Naor（1969）发表在 *Econometrica* 上的一篇文章开创了该领域的研究先河。后来，为了和经典随机服务系统理论相区分，该领域被命名为 Economics of Queues，可译为排队经济学，也可称为排队博弈论。在本书中称其为经济学随机服务系统理论。经济学随机服务系统理论的主要研究内容是排队系统中有限理性顾客与有限理性服务员的博弈和决策行为，以及由他们相互或内部之间的博弈行为所产生的经济学问题。

目前，经济学随机服务系统理论被归为行为运筹学的一个分支，并正作为一个蓬勃发展的研究领域受到排队论研究者、经济学工作者和服务企业管理者等群体的高度关注。然而，该领域的理论体系还远远没有完善，若干典型的排队系统的研究工作还仅仅停留在经典随机服务系统理论的层面上。

作者撰写本书的目的：让读者通过阅读，能够对经济学随机服务系统理论这个学科的研究背景及意义有一个初步的认识，了解它的发展现状，熟悉该领域问题的建模及求解过程，并激励正在或准备在该领域展开研究的学者从中获取启发和灵感，以对未来课题的研究起到抛砖引玉的作用。

作者从 2008 年开始对经济学随机服务系统理论进行研究，在国内属于较早对该领域展开研究的学者之一，到目前已经积累了一些研究成果。本书中除“绪论”之外均为作者自己的研究工作。而且，在每章的最后都设置了“相关文献评述”一节，给出本章内容和相关工作的出处，便于读者阅读和进一步的文献研究。

阅读本书需要高等数学、微观经济学、概率论、经典排队论、博弈论、随机过程和矩阵分析等的基本知识，因此一些基础的预备知识在本书中不加赘述。

2012 年至 2016 年间，我们关于经济学随机服务系统理论及其应用的研究得到了国家自然科学基金青年项目、面上项目的连续资助。与本书内容直接相关的研究成果，主要来自于编号为 71101124、71301139 和 71671159 三个项目的子课题。

在此，我们对国家自然科学基金委表示感谢，对第一作者的博士生导师田乃硕教授在课题研究和生活方面的无私帮助表示衷心的感谢，对在书稿撰写过程中提出宝贵意见和建议的金顺福教授、马占友教授和徐秀丽教授等表示感谢，对在课题研究过程中做了大量工作的毛炳蔚副教授、鄂成国博士等表示感谢。

孙微　李世勇

2016 年 10 月

于秦皇岛　燕山大学

目　录

第 1 章　绪　论……1
1.1　经典随机服务系统理论……1
1.2　经济学随机服务系统理论……2
1.2.1　问题的提出……2
1.2.2　均衡决策行为……3
1.2.3　系统信息环境……4
1.2.4　顾客偏好类型……4
1.2.5　系统最优控制……5
1.3　预备知识……5
1.3.1　博弈及其构成……6
1.3.2　非合作博弈……7
1.3.3　顾客剩余效用……9
1.3.4　单位社会福利……9
第 2 章　具有启动-关闭策略的排队系统……10
2.1　启动-关闭策略……10
2.2　可中断启动-关闭策略……10
2.2.1　模型描述……10
2.2.2　完全可视排队系统……11
2.2.3　完全不可视排队系统……15
2.3　可跳跃启动-关闭策略……19
2.3.1　完全可视排队系统……20
2.3.2　完全不可视排队系统……25
2.4　非中断启动-关闭策略……29

2.4.1 完全可视排队系统 …… 29
2.4.2 完全不可视排队系统 …… 35
2.5 数值比较 …… 39
2.6 相关文献评述 …… 42
第 3 章 具有多重休假策略的排队系统 …… 44
3.1 多重休假策略 …… 44
3.2 具有不同决策准则的排队系统 …… 44
3.2.1 模型描述 …… 44
3.2.2 线性等待损耗 …… 45
3.2.3 二次等待损耗 …… 55
3.2.4 指数等待损耗 …… 62
3.3 具有阈值多重休假策略的排队系统 …… 72
3.3.1 模型描述 …… 72
3.3.2 完全可视排队系统 …… 73
3.3.3 完全不可视排队系统 …… 83
3.4 相关文献评述 …… 90
第 4 章 具有工作休假策略的排队系统 …… 92
4.1 工作休假策略 …… 92
4.2 具有多重工作休假策略的排队系统 …… 92
4.2.1 模型描述 …… 92
4.2.2 完全可视排队系统 …… 93
4.2.3 几乎不可视排队系统 …… 100
4.2.4 完全不可视排队系统 …… 107
4.3 具有双阶段工作休假策略的排队系统 …… 109
4.3.1 模型描述 …… 109
4.3.2 完全可视排队系统 …… 109
4.3.3 几乎不可视排队系统 …… 119
4.3.4 完全不可视排队系统 …… 130
4.4 具有双级工作休假策略的排队系统 …… 135
4.4.1 模型描述 …… 135
4.4.2 几乎不可视排队系统 …… 135

4.4.3　完全不可视排队系统 …… 143
4.5　具有阈值工作休假策略的排队系统 …… 150
4.5.1　模型描述 …… 150
4.5.2　队长与忙循环 …… 151
4.5.3　数值比较 …… 158
4.6　相关文献评述 …… 163
第 5 章　具有中途退出行为的排队系统 …… 164
5.1　中途退出行为 …… 164
5.2　具有不可靠服务员与维修员的排队系统 …… 164
5.2.1　模型描述 …… 164
5.2.2　非清场排队系统 …… 165
5.2.3　清场排队系统 …… 171
5.2.4　数值比较 …… 177
5.3　具有顺序退出行为与休假策略的排队系统 …… 181
5.3.1　模型描述 …… 181
5.3.2　完全可视排队系统 …… 181
5.3.3　几乎可视排队系统 …… 184
5.3.4　几乎不可视排队系统 …… 187
5.3.5　完全不可视排队系统 …… 191
5.3.6　数值比较 …… 193
5.4　相关文献评述 …… 196
第 6 章　具有部分服务信息的排队系统 …… 198
6.1　最大熵原理 …… 198
6.2　模型描述 …… 199
6.3　风险中立顾客 …… 200
6.3.1　纳什均衡 …… 200
6.3.2　社会最优 …… 202
6.3.3　利润最大 …… 206
6.4　风险规避顾客 …… 209
6.4.1　纳什均衡 …… 211
6.4.2　社会最优 …… 213

6.4.3 利润最大……220
6.5 相关文献评述……223
第 7 章 具有随机服务参数的排队系统……225
7.1 服务参数的不确定性……225
7.2 模型描述……225
7.3 可变服务速率……226
7.3.1 服务速率信息封闭……227
7.3.2 服务速率信息公开……227
7.3.3 可变服务速率对利润的影响……234
7.4 可变等待损耗……235
7.4.1 等待损耗信息封闭……236
7.4.2 等待损耗信息公开……236
7.4.3 可变等待损耗对利润的影响……242
7.5 可变服务收益……243
7.5.1 服务收益信息封闭……243
7.5.2 服务收益信息公开……243
7.5.3 可变服务收益对利润的影响……248
7.6 比较结果总结……250
7.7 相关文献评述……251
参考文献……252

第1章 绪 论

1.1 经典随机服务系统理论

排队是我们在日常生活中经常遇到的现象，如顾客到商店购买商品形成的排队；上下班坐公共汽车，等待公共汽车的排队；文件等待打印和发送；电话局的占线；故障机器的停机待修；水库的存贮调节，等等。由于顾客到达间隔时间和服务时间具有随机性，使排队现象不可避免。一般来说，排队是一种分配稀缺商品（服务资源）的办法。在资源配置不足的情况下，往往不得不引入排队机制来解决分配问题，因为排队可最大限度地节约顾客的时间成本，使资源得到相对的优化配置。为了解决排队现象对顾客个人和社会带来的不利影响以及相应的优化设计问题，此方面的研究工作就逐渐发展成了一门学科——排队论（Queueing Theory），即随机服务系统理论（Random Service System Theory）。

经典的随机服务系统理论是研究服务系统在运行过程中所产生的排队等待现象的一门数学理论，是运筹学的一个重要分支。具体地说，它是在研究各种排队系统概率规律性的基础上，通过研究各种服务系统在排队中的概率特性，得到队长、等待时间等数量指标的变化规律，解决相应排队系统的最优设计和最优控制问题。

经典随机服务系统理论起源于丹麦工程师 Erlang[1]关于电话交换机使用情况的研究。一直到第二次世界大战之后，经典随机服务系统理论才得以迅猛发展。在 20 世纪 50 年代，英国统计学家 Kendall[2]给出了排队系统的符号表示。从此，学者们对随机服务系统展开了大量的研究，研究对象主要集中在连续时间排队系统。然而，随着计算机通信网络的普及与发展，学者们发现离散时间排队系统正是刻画计算机通信网络结构与原理的最理想模型，且最初由 Meisling[3]对其展开了研究。平行于连续时间经典随机服务系统理论，离散时间经典随机服务系统理论也已经发展得比较完善，请见参考文献[4]与[5]。

从 20 世纪七八十年代开始，随着计算机网络、柔性制造系统和异步转移模式等高新技术领域的发展，以往的关于无休假排队系统的研究结果在处理这些新领域提出的新问题时表现出了局限性。Levy 和 Yechiali[6]从有效利用系统闲期的观点出发，首次研究了具有休假策略的 M/G/1 型排队系统，并引入了“休假”和“休假策略”等术语。关于经典休假排队系统的研究工作可参见参考文献[7]与[8]。在 2002 年，Servi 和 Finn[9]又引入了一类半休假策略，指的是在休假期间服务员将以较低的速度接待顾客，而不是完全停止对顾客的服务，又称为工作休假（Working Vacation，WV）策略。关于各种工作休假策略的介绍及相关文献请见参考文献[10]与[11]。

从经典随机服务系统理论的研究历史来看，研究目标主要是导出系统各类数量指标并进行优化控制。在此过程中，研究对象均是完全理性主体，即顾客或服务员在排队或提供服务的过程中不能主动地做出各类排队决策，即他们不具有主观能动性。

1.2 经济学随机服务系统理论

与经典随机服务系统理论不同，经济学随机服务系统理论中将研究对象定义为有限理性主体，即顾客或服务员在排队或提供服务的过程中能够根据自身的利益得失而主动地做出各类排队决策。在现实之中，这是更加合乎常理的，因为系统主体的偏好类型及信息环境等都会影响他们的决策行为。

1.2.1 问题的提出

当随机服务系统中的顾客和服务员被视为有限理性主体时，出于自身利益最大化的目的，他们内部之间势必会产生非合作博弈行为，这种博弈行为的结果就是导致一个均衡格局的形成，这种均衡格局在经济学和博弈论中被称为“纳什均衡”或“非合作博弈均衡”（Nash Equilibrium）[12–16]。然而，这个均衡从整个全局或社会角度看很可能并不是最优的（经典的案例就是纳什提出的“囚徒的困境”问题），由此大量经济学和排队论学者针对这一经济学现象在排队论和博弈论背景和框架下开始了研究工作。

从学科归属来看，该研究方向属于行为运筹学（Behavioral Operations Research）的一个分支，是经济学、经典排队论、博弈论、决策论和心理学等学科

相结合的一门交叉学科。从名称上来说，按照国外知名学者对该方向的归纳命名，可译为排队经济学（Economics of Queues）[17]。简而言之，排队经济学的主要研究内容是排队系统中有限理性顾客与有限理性服务员的博弈和决策行为，以及由他们相互或内部之间的博弈行为所产生的经济学问题。对排队经济学系统中的博弈和决策行为进行研究，即是对顾客或服务员的均衡决策过程和均衡策略选择进行研究。

为了便于排队论学者和其他读者能够正确认识排队经济学与经典随机服务系统理论的联系和区别，我们在本书中用经济学随机服务系统理论的名称来代替排队经济学的名称，以与经典随机服务系统理论相呼应。最早的该领域的研究成果要从 Naor[18]在 1969 年发表在 *Econometrica* 上的一篇文章算起。此后，有关经济学随机服务系统理论的研究成果不断涌现。

1.2.2　均衡决策行为

对随机服务系统主体的均衡决策行为进行研究是经济学随机服务系统理论的主要目的。对于顾客而言，主要的决策行为包括均衡止步策略与均衡退出策略的选择，对于优先权排队或重试排队来说，是否购买优先权或重试频率的选择等也是顾客面临的决策行为之一[19]。我们将在后续章节的相关文献评述部分对文献内容给予介绍。而对于服务员来说，均衡决策行为包括服务价格与服务能力的确定，对于多服务员排队来说，服务顺序及出价顺序等的确定也是服务员需要面临的决策行为。

从服务员利润最大化的角度考虑，Naor[18]就已经在他的模型中提出了为使利润最大化而对顾客收取的费用要高于社会所希望收取的费用的观点。此后，又陆续出现了许多关于服务员定价的研究成果[20–27]，但相比之下，服务能力选择和竞争方面的研究工作并不多见[28,29]。至于服务顺序和出价顺序决策行为方面的研究，Hassin[30]讨论了具有两个服务员的排队系统，其服务顺序分别为先到先服务（First Come First Service，FCFS）与随机服务（Random Service，RS），数值结果表明，均衡状态下，具有先到先服务顺序的服务员将占有大部分的市场份额。Guo 和 Hassin[31]同样讨论了两个服务员排队系统，在出价顺序方面，当服务需求较大时，先出价者将具有先动优势（First-mover Advantage），否则，当服务需求较小时，先动优势将不复存在。

1.2.3 系统信息环境

在经济学随机服务系统中，系统信息环境对于主体决策选择往往起到至关重要的作用。对于顾客来说，系统信息一般包括顾客到达时的队列长度、服务率、服务价格、服务员所处状态等。对于服务员来说，系统信息一般包括顾客到达率、队列长度等。当然，如果是多服务员排队系统，其他服务员的服务率和服务价格等信息对于其中任意一个服务员的决策行为也是具有重要意义的。

一般地，我们用信息水平或信息精度来度量系统信息量的多少，且当主体能观察到某类信息时，称为可视情形（Observable Case），否则，称为不可视情形（Unobservable Case）。一般地，当主体能观察到有关系统状态的所有信息时，称为完全可视情形（Fully Observable Case），当主体能观察到有关系统状态的部分主要信息时，称为几乎可视情形（Almost Observable Case）；否则，当主体只能观察到有关系统状态的部分次要信息时，称为几乎不可视情形（Almost Unobservable Case），当主体完全不能观察到有关系统状态的所有信息时，称为完全不可视情形（Fully Unobservable Case）。

经济学随机服务系统理论的奠基人 Naor[18]首次研究了怎样管理和控制队长可视 M/M/1 排队系统的问题。他发现在队长可视排队系统中，个体顾客的决策往往背离整个社会的利益偏好。这种差异是由于个体最优行为所产生的负面作用引起的。首先发现基本的队长不可视排队系统性质的是 Edelson 和 Hildebrand[32]，他们对 Naor[18]的模型进行了改进，取消了队长可视的条件而假定队长不可视。与队长可视排队一样，得出了个体顾客为了追求自身利益最大化而导致系统过度拥塞的结论。针对队长可视和队长不可视两类排队系统的异同，Hassin[33]最早对它们中的社会福利与服务员利润进行了比较。读者还可参阅其他一些有关系统信息环境的代表性文献与综述[34–37]。

1.2.4 顾客偏好类型

与系统信息环境一样，顾客偏好类型对顾客的决策行为同样起着决定性的作用。在经济学随机服务系统中，顾客偏好类型一般可分为风险规避（Risk Averse）、风险中立（Risk Neutral）和风险偏好（Risk Appetite）三种类型，且顾客的偏好类型体现在其对排队延迟损耗的敏感度上。在同一系统中，到达顾客可以全部为同类顾客（Homogeneous Customers），也可以为有限类顾客，甚至彼此可均

为异类顾客（Heterogeneous Customers）。顾客偏好类型的不同，可致使两种可能现象的发生：拥挤偏好（Follow the Crowd, FTC）和拥挤厌恶（Avoid the Crowd, ATC）。在 ATC 情形下，最多只有一个均衡点，而在 FTC 情形下，可能存在多个均衡点。

Guo 和 Hassin[38]首先研究了具有阈值休假策略的队长可视与队长不可视排队系统中同类顾客的均衡和最优止步策略，且在队长不可视情形下系统最多具有三个均衡点。随后，Guo 和 Hassin[39]又研究了具有阈值休假策略的完全可视与完全不可视排队系统中异类顾客（包括两类顾客和无穷类顾客）的均衡和最优止步策略，且在队长不可视情形下系统可具有多个均衡点。更多关于异类顾客的研究工作可参阅 Hassin 和 Haviv 的综述[36]。

1.2.5　系统最优控制

由于顾客的个体均衡决策行为往往背离整个社会的整体利益偏好，因此，对于社会计划者（Social Planner）而言，希望通过调整服务规则（出售优先权）或调节服务费用（调高收费标准）等措施，来达到控制顾客均衡决策行为的目的，从而使得顾客的均衡和社会最优决策行为尽量保持一致。

Naor[18]在他的模型中就提出，要想控制顾客的均衡行为，就需要对顾客收取一定的费用。Hassin[33]对 Naor[18]的模型进行了改进，提出了 LCFS-PR 模型并对此进行了初步研究。他提出了一种新的方法，在此方法下，即使不对顾客收取费用，也能达到社会福利最优。最早对优先权排队系统进行纳什均衡分析的学者是 Balachandran[40]，他讨论了一个队长可视 M/M/1 排队系统，服务顺序不再采用先到先服务规则，而是根据顾客的付款额度高低来对其服务顺序进行排序。Hassin[41]则首先建议用优先权拍卖的方法来解决不可视排队系统中均衡状态下顾客的过度拥塞问题。

1.3　预备知识

本节将简要介绍本书中涉及的基本概念及术语[17,19,42]，主要包括博弈及其构成、非合作博弈、顾客剩余效用与单位社会福利。

1.3.1 博弈及其构成

博弈，即一些个人、队组或其他组织，面对一定的环境条件，在一定的规则下，同时或先后，一次或多次，从各自允许选择的行为或策略中进行选择并加以实施，各自取得相应结果或收益的过程。

从上述定义中可以看出，规定或定义一个博弈需要设定下列五个方面。

1. 博弈的参加者（Players）

即在所定义的博弈中究竟有哪几个独立决策、独立承担结果的个人或组织。对我们来说，只要在一个博弈中统一决策、统一行动、统一承担结果，不管一个组织有多大，哪怕是一个国家，甚至是许多国家组成的联合国，都可以作为博弈中的一个参加者。并且，在博弈的规则确定之后，各参加者都是平等的，大家都必须严格按照规则办事。只有两个参与者的博弈现象称为“两人博弈”，而多于两个参与者的博弈现象称为“多人博弈”。

2. 各参与者选择的全部策略（Strategies）或行为（Actions）的集合

即规定每个参与者在进行决策时，可以选择的方法、做法或经济活动的水平、量值等。在不同博弈中可供参与者选择的策略或行为的数量很不相同，在同一个博弈中，不同参与者的可选策略或行为的内容和数量也常不同，有时只有有限的几种，甚至只有一种，而有时又可能有许多种，甚至无限多种可选策略或行为。

3. 博弈信息（Information）

即参与者所掌握的对选择策略有帮助的情报资料或知识，特别是有关自然的选择，其他参与者的特征和行动的知识。根据一个参与者是否对其他参与者（包括虚拟人自然）的行为选择有准确的了解，可将博弈进行分化。

4. 进行博弈的次序（Orders）

在现实的各种博弈活动中，当存在多个独立参与者进行决策时，有时候需要这些参与者同时做出选择，因为这样能保证公平合理，而很多时候各参与者的决策又有先后之分，并且有时一个参与者还要做不止一次的决策选择，这就会产生次序问题。因此规定一个博弈必须规定其中的次序，即使博弈的其他方面都相同，次序不同往往会导致不同的博弈结果。

5. 参与者的收益（Payoffs）

对应于各参与者的每一组可能的决策选择，都应有一个结果表示该策略组合下各参与者的得失。绝大多数的博弈本身都有数量的结果或可以量化为数量的结

果，如收入、利润、损失、个人效用和社会效用、社会福利等。博弈中的这些可能结果的量化数值，称为各参与者在相应情况下的“ 收益 ”。规定一个博弈必须对收益做出规定，收益可以是正值，也可以是负值，它们是分析博弈模型的标准和基础。值得注意的是，虽然各参与者在各种情况下的收益应该是客观存在的，但这并不意味着各参与者都了解各方的收益情况。

确定了上述五个方面就确定了一个博弈。博弈论就是系统研究用上述方法定义的各种博弈问题的理论和方法，在各参与者具有充分或者有限理性、能力的条件下，寻求合理的策略选择和合理选择策略时博弈的结果，并分析这些结果的经济意义和效率意义。

我们把所有参与者同时或可看作同时选择策略的博弈称为“ 静态博弈 ”（Static Game），把决策有先后顺序的博弈称为“ 动态博弈 ”（Dynamic Game）。动态博弈中，在决策之前对博弈的进程完全了解的参与者，称为具有“ 完美信息 ”（Perfect Information）的参与者，如果动态博弈的所有参与者都具有完美信息，则称为“ 完美信息动态博弈 ”。动态博弈中，在决策之前不完全了解此前全部博弈进程的参与者，称为具有“ 不完美信息 ”（Imperfect Information）的参与者，具有该类参与者的动态博弈则称为“ 不完美信息动态博弈 ”。

1.3.2　非合作博弈

一般地，我们将允许存在有约束力协议的博弈称为“ 合作博弈 ”（Cooperative Game）。与此相对的是，我们将不允许存在有约束力协议的博弈称为“ 非合作博弈 ”（Uncooperative Game）。由于在合作博弈和非合作博弈两类博弈中，参与者的基本行为逻辑和研究方法有很大差异，因此它们是两类很不相同的博弈。事实上，合作博弈理论和非合作博弈理论正是博弈论最基本的一个分类，它们在产生和发展的路径，在经济学中的地位、作用和影响等许多方面都有很大的差异。现代博弈论的研究重点，也是研究和应用较多较广泛的，主要是非合作博弈理论。

1.3.2.1　纯策略与混合策略

我们常用 G 表示一个非合作博弈。如果 G 有 n 个参与者，每个参与者全部可选策略的集合称为策略空间，分别用 S_i（i =1, 2, ⋯, n）表示。

一般地，我们将各参与者都完全了解所有参与者各种情况下的博弈称为完全信息（Complete Information）博弈，而将至少有部分参与者不完全了解其他参与

者收益情况的博弈称为不完全信息（Incomplete Information）博弈或贝叶斯博弈（Bayesian Game）。

纯策略（Pure Strategy）：在完全信息博弈中，如果在每个给定信息下，只能选择一种特定策略，这个策略称为纯策略。

混合策略（Mixed Strategy）：如果在每个给定信息下只以某种概率选择不同策略，称为混合策略。混合策略是纯策略在空间上的概率分布，纯策略是混合策略的特例。

收益函数（Payoff Function）：每个参与者在参与博弈时依据其所属类型和选择的行动可获得的收益。

记 $s = (s_1, \cdots, s_n)$为一个策略组合，$s_i \in S_i$ 表示第 i 个参与者的策略，每个参与者的收益函数为 $F_i(s)$。用 s_{-i} 表示除了第 i 个参与者，其他所有参与者的策略组合。假设第 i 个参与者有两种行为选择，s_i^1 和 s_i^2，如果 s_i 是一个混合策略，以概率 α 选择 s_i^1，以概率 $1-\alpha$ 选择 s_i^2，那么第 i 个参与者在该混合策略下的收益就是

$$F_i(s_i, s_{-i}) = \alpha F_i(s_i^1, s_{-i}) + (1-\alpha) F_i(s_i^2, s_{-i})$$

策略 s_i^1 被称为比策略 s_i^2 “弱占优”（Weakly Dominant）的策略，指的是对于任意其他所有参与者的策略组合 s_{-i} 都有 $F_i(s_i^1, s_{-i}) \geq F_i(s_i^2, s_{-i})$ 成立，其中至少有一个 s_{-i} 使得“>”成立。如果称 s_i 是一个弱占优的策略，那么它对其他所有策略都是弱占优的。

如果

$$s_i^* \in \arg\max_{\{s_i \in S_i\}} F_i(s_i, s_{-i})$$

那么策略 s_i^* 被称为是相对于 s_{-i} 的最佳策略。

1.3.2.2 纳什均衡

在博弈 G 中，如果由各个参与者的各一个策略组成的某个策略组合 $s = (s_1, \cdots, s_n)$中，任一参与者 i 的策略 s_i^e，都是对其余参与者策略组合 s_{-i}^e 的最佳策略，也即

$$F_i(s_i^e, s_{-i}^e) \geq F_i(s_i, s_{-i}^e)$$

对任意的 $s_i \in S_i$ 都成立，则称 $s^e = (s_1^e, \cdots, s_n^e)$ 为 G 的一个纳什均衡。本书中提到的“均衡”都表示的是“纳什均衡”。

如果各参与者都是同类型的，那么得到的是对称均衡（Symmetric Equilibrium）。由一个动态博弈第一阶段以外的某阶段开始的后续博弈阶段构成的，有初始信息集和进行博弈所需要的全部信息，能够自成一个博弈的原博弈的一部分，称为原

动态博弈的一个子博弈（Subgame）。

如果在一个完美信息动态博弈中，各参与者的策略构成的一个策略组合满足在整个动态博弈及它的所有子博弈中都构成纳什均衡的条件，那么这个策略组合称为该动态博弈的一个子博弈精炼纳什均衡（Subgame Perfect Equilibrium）。简单地说，这组策略必须在每一个子博弈中都是最优的。子博弈精炼纳什均衡是稳定的，是比纳什均衡更强的均衡（排除均衡策略中不可置信的威胁或承诺）。如果不是子博弈精炼纳什均衡，可能在某些子博弈中不符合参与者的自身利益。

1.3.3　顾客剩余效用

在随机服务系统中，顾客在接受服务之后会得到相应的回报收益（Reward），而在排队等待的过程中会产生等待损耗（Waiting Cost），需要支付等待费用。等待费用为顾客等待时间的函数，且函数形式与顾客风险偏好类型有关。除了间接的等待费用之外，可能在服务之前还需要直接支付一定的服务费用（Service Payment）或优先权购买费用等。把顾客直接和间接的花费合在一起称为总费用，那么顾客剩余效用（Customer Residual Utility）就是回报收益与总费用之差。在非合作博弈中，顾客的个人目标是让自己的剩余效用达到最大，而不考虑其他人的收益情况。

1.3.4　单位社会福利

如果从整个社会的角度来考虑一个随机服务系统，那么其目标就是使得单位社会福利（Social Welfare Per Time Unit）达到最大。这里的单位社会福利指的是单位时间内排队系统中所有成员（包括顾客和服务员）积累的总平均剩余效用，通常也称为单位平均社会福利或简称为社会福利。如果考虑系统运行时产生的各种成本费用，还需要在单位社会福利中将其剔除掉。需要注意的是，顾客和服务员之间的利益转换（如服务员向顾客收取的服务费用或优先权购买费用等）对单位社会福利不产生任何影响。

第 2 章　具有启动-关闭策略的排队系统

2.1　启动-关闭策略

本章考虑三种启动-关闭策略：可中断、可跳跃和非中断启动-关闭策略。当系统空竭后，服务员开始一个关闭期。若关闭期内没有顾客到达，转入空闲状态。一旦有顾客到达，则转入启动期。对于具有可中断启动-关闭策略的排队系统，如果关闭期内有顾客到达则立即停止关闭期进入忙期；对于具有可跳跃启动-关闭策略的排队系统，如果关闭期内有顾客到达则在关闭期结束之后立即进入忙期；而对于具有非中断启动-关闭策略的排队系统，即使关闭期内有顾客到达也必须经历一个启动期才能转入忙期。

对于分别具有这三类启动-关闭策略的排队系统，本章考虑了完全可视和完全不可视两种情形。对于完全可视情形，导出了顾客的均衡阈值止步策略，分析了系统的稳态行为。另一方面，对于完全不可视情形，分别得到了顾客的均衡和社会最优混合止步策略并找到了最大社会福利。为了使得顾客的均衡混合止步策略与最优混合止步策略保持一致，得到了使服务员利润最大化的定价策略。

本章内容为作者研究成果[43,44]的进一步完善和补充。

2.2　可中断启动-关闭策略

2.2.1　模型描述

首先考虑具有可中断启动-关闭策略的完全可视和完全不可视排队系统。顾客到达过程服从参数为 λ 的 Poisson 过程，且服务时间服从参数为 μ 的指数分布。

一旦系统空竭，服务员进入一个关闭期。如果期间有顾客到达，则服务员立即进入忙期。否则，服务员进入空闲状态，直到有顾客到达为止。该到达顾客将

使服务员进入启动期，启动期内顾客可继续到达并等候启动期结束。关闭期和启动期均服从指数分布，参数分别为 ξ 和 θ。假设到达间隔、服务时间、关闭期和启动期相互独立。

在此，用符号$(N(t),\ I(t))$表示时刻 t 系统所处的状态，其中 $N(t)$表示时刻 t 系统中的顾客数（即队长），$I(t)$表示时刻 t 服务员所处的状态，且

$$I(t)=\begin{cases}0,\ \text{服务员处于忙期}\\ 1,\ \text{服务员处于关闭期}\\ 2,\ \text{服务员处于启动期}\end{cases}$$

因此，在完全可视情形下，到达顾客在决定是否进入系统前可观测到系统的状态 $N(t)$和 $I(t)$。相反，在完全不可视情形下，到达顾客不能观测到系统的状态 $N(t)$和 $I(t)$。

在分析完全可视情形时，均衡状态下一个顾客在状态 i $(i = 0, 1, 2)$进入系统的平均逗留时间可表示为 $s_e(i)$，且相应的平均剩余效用可表示为 $b_e(i)$。显然，对于具有可中断启动-关闭策略的完全可视排队系统，过程$\{(N(t), I(t)) : t\geq 0\}$是一个具有状态空间 $S = \{(n, 0) : n\geq 1\}\cup\{(n, 1) : n = 0\}\cup\{(n, 2) : n\geq 0\}$的连续时间 Markov 链。

本章的核心目的是为了研究在均衡和社会最优条件下顾客是如何做出是否进入队列的决策问题的，即研究顾客的均衡和最优止步策略。为了对此决策过程进行建模，假设每个顾客服务完成后获得的收益为 R，在其排队过程中每单位时间需要承受的损耗为 C，此顾客需交纳的服务费用为 p。均衡状态下，顾客风险中立且使得自身的平均剩余效用最大。但在社会最优条件下，顾客将使得单位社会福利最大化。假设

$$R>\frac{C}{\mu}+\frac{C}{\theta}+p \tag{2-1}$$

这个条件保证了在一个到达顾客发现服务员处于启动期并决定进入系统的情况下，接受服务所得的收益大于排队过程中的平均损耗。否则，一旦某个关闭期内没有顾客到达，以后将不会有顾客再进入系统。假设决定不进入系统而离开的顾客将永远不再回来，且进入系统的顾客不允许中途退出。

2.2.2　完全可视排队系统

在完全可视情形下，顾客在到达时刻 t 既可观察到服务员的状态 $I(t)$又可观察到队列长度 $N(t)$。假设一个顾客观察到系统状态为(n, i)而决定进入系统，则他的逗留时间为

$$(s_e(0), s_e(1), s_e(2)) = \left(\frac{n+1}{\mu}, \frac{1}{\mu}, \frac{n+1}{\mu} + \frac{1}{\theta} \right)$$

因此他的平均剩余效用为

$$(b_e(0), b_e(1), b_e(2)) = \left(R - \frac{C(n+1)}{\mu} - p, R - \frac{C}{\mu} - p, R - \frac{C(n+1)}{\mu} - \frac{C}{\theta} - p \right)$$

如果剩余效用为正值，则此顾客一定会进入系统。

定理 2.2.1 在具有可中断启动-关闭策略的完全可视 M/M/1 排队系统中，如果时刻 t 到达的顾客观测到系统状态为$(N(t), I(t))$，则存在阈值

$$(n_e(0), n_e(1), n_e(2)) = \left(\left\lfloor \frac{(R-p)\mu}{C} \right\rfloor - 1, 0, \left\lfloor \frac{(R-p)\mu}{C} - \frac{\mu}{\theta} \right\rfloor - 1 \right) \tag{2-2}$$

且“ $N(t) \leq n_e(I(t))$则进入系统，否则放弃离开 ”是唯一的均衡止步策略。

对于系统的稳态分析，注意到如果所有的顾客遵从式（2-2）给出的均衡阈值止步策略，则系统过程是一个状态空间为 $S_{ob}=\{(n, 0) \mid 1 \leq n \leq n_e(0)+1\} \cup \{(0, 1)\} \cup \{(n, 2) \mid 0 \leq n \leq n_e(2)+1\}$的 Markov 链，且状态转移图如图 2-1 所示。

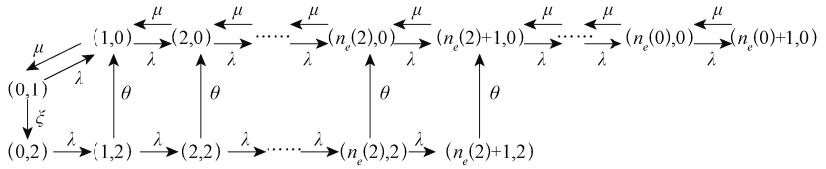

图 2-1 可中断启动-关闭策略和完全可视情形下的状态转移图

相应的系统稳态队长分布$\{p_{ob}(n, i) : (n, i) \in S_{ob}\}$为以下均衡方程组

$$p(0, 1)(\lambda + \xi) = p(1, 0)\mu \tag{2-3}$$

$$p(0, 2)\lambda = p(0, 1)\xi \tag{2-4}$$

$$p(1, 0)(\lambda + \mu) = p(0, 1)\lambda + p(1, 2)\theta + p(2, 0)\mu \tag{2-5}$$

$$p(n, 0)(\lambda + \mu) = p(n-1, 0)\lambda + p(n, 2)\theta + p(n+1, 0)\mu,\ 2 \leq n \leq n_e(2)+1 \tag{2-6}$$

$$p(n, 0)(\lambda + \mu) = p(n-1, 0)\lambda + p(n+1, 0)\mu,\ n_e(2)+2 \leq n \leq n_e(0) \tag{2-7}$$

$$p(n_e(0)+1, 0)\mu = p(n_e(0), 0)\lambda \tag{2-8}$$

$$p(n, 2)(\lambda + \theta) = p(n-1, 2)\lambda,\ 1 \leq n \leq n_e(2) \tag{2-9}$$

$$p(n_e(2)+1, 2)\theta = p(n_e(2), 2)\lambda \tag{2-10}$$

的唯一正则解。定义

$$\rho = \frac{\lambda}{\mu},\ \sigma = \frac{\lambda}{\lambda + \theta}$$

迭代式（2-7）和式（2-9），并同时考虑式（2-4）、式（2-8）和式（2-10），得到

$$p(n,2)=\frac{\xi}{\lambda}\sigma^n p(0,1),\ 0\le n\le n_e(2) \tag{2-11}$$

$$p(n_e(2)+1,2)=\frac{\xi}{\theta}\sigma^{n_e(2)}p(0,1) \tag{2-12}$$

$$p(n,0)=\rho^{n-n_e(2)-1}p(n_e(2)+1,0),\ n_e(2)+1\le n\le n_e(0)+1 \tag{2-13}$$

$$p(n_e(2)+1,0)\mu=p(n_e(2),0)\lambda+p(n_e(2)+1,2)\theta \tag{2-14}$$

从式（2-6）和式（2-14），可知$\{p(n,1):1\le n\le n_e(2)+1\}$是非齐次线性差分方程

$$\begin{aligned}\mu x_{n+1}-(\lambda+\mu)x_n+\lambda x_{n-1}&=-\theta p(n,2)\\&=-\frac{\theta\xi}{\lambda}\sigma^n p(0,1),\ 2\le n\le n_e(2)\end{aligned} \tag{2-15}$$

的解。考虑其相应的特征方程

$$\mu x^2-(\lambda+\mu)x+\lambda=0$$

此方程具有两个实根 :1 和 ρ, 所以式 (2-15) 的齐次形式的通解为 $x_n^{hom}=A1^n+B\rho^n$（假设 $\rho\neq1$），而式（2-15）的通解为 $x_n^{gen}=x_n^{hom}+x_n^{spec}$，其中 x_n^{spec} 是式（2-15）的一个特解。由于式（2-15）的非齐次部分 $-\frac{\theta\xi}{\lambda}\sigma^n p(0,1)$ 是参数 σ 的几何形式，因此可得到形为 $C\sigma^n$ 的特解（假设 $\sigma\neq1$，$\sigma\neq\rho$）。将 $x_n^{spec}=C\sigma^n$ 代入式（2-15），得到

$$C=\frac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}p(0,1) \tag{2-16}$$

因此，式（2-15）的通解为

$$x_n^{gen}=A1^n+B\rho^n+C\sigma^n,\quad 1\le n\le n_e(2)+1 \tag{2-17}$$

其中 A, B 为未知系数。从式（2-17）可知，当 $n=1$ 时，有

$$A+\rho B+\frac{\xi}{\mu-\lambda-\theta}p(0,1)=\left(\rho+\frac{\xi}{\mu}\right)p(0,1) \tag{2-18}$$

当 $n=2$ 时，将式（2-17）代入式（2-5），得到

$$A+\rho^2B+\frac{\lambda\xi}{(\mu-\lambda-\theta)(\lambda+\theta)}p(0,1)=\rho\left(\rho+\frac{\xi}{\mu}+\frac{\xi}{\lambda+\theta}\right)p(0,1) \tag{2-19}$$

求解方程组（2-18）和（2-19），得到

$$\begin{cases}A=0\\B=\left(1-\dfrac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}\right)p(0,1)\end{cases}$$

因此，根据式（2-17），得到

$$p(n,0)=\left(\rho^n+\frac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}(\sigma^n-\rho^n)\right)p(0,1),\quad 1\le n\le n_e(2)+1 \tag{2-20}$$

最后，状态概率 $p(0,1)$可通过平衡方程

$$\sum_{n=1}^{n_e(0)+1}p(n,0)+p(0,1)+\sum_{n=0}^{n_e(2)+1}p(n,2)=1$$

求得。经简化后，以下定理给出了系统稳态队长分布。

定理 2.2.2 在具有可中断启动-关闭策略的完全可视 M/M/1 排队系统中，如果顾客均采用如式（2-2）所示的均衡阈值止步策略，且 $\sigma\neq 1\neq\rho\neq\sigma$，则系统稳态队长分布 $\{p_{ob}(n,i):(n,i)\in S_{ob}\}$为

$$\begin{aligned}p_{ob}(0,1)=&\left[\frac{\xi}{\mu}\left(\frac{1}{\rho(1-\rho)(1-\sigma)}-\frac{\sigma^{n_e(2)+1}}{(1-\rho)(1-\sigma)}\right.\right.\\&\left.\left.+\frac{1}{(1-\rho)(\sigma-\rho)}\rho^{n_e(0)+2}\left(1-\left(\frac{\sigma}{\rho}\right)^{n_e(2)+1}\right)\right)+\frac{1-\rho^{n_e(0)+2}}{1-\rho}\right]^{-1}\end{aligned} \tag{2-21}$$

$$p_{ob}(n,2)=\frac{\xi}{\lambda}\sigma^n p_{ob}(0,1)=\frac{1}{\rho}\frac{\xi}{\mu}\sigma^n p_{ob}(0,1),\quad 0\le n\le n_e(2) \tag{2-22}$$

$$p_{ob}(n_e(2)+1,2)=\frac{\xi}{\theta}\sigma^{n_e(2)}p_{ob}(0,1)=\frac{1}{\rho(1-\sigma)}\frac{\xi}{\mu}\sigma^{n_e(2)+1}p_{ob}(0,1) \tag{2-23}$$

$$\begin{aligned}p_{ob}(n,0)&=\left(\rho^n+\frac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}(\sigma^n-\rho^n)\right)p_{ob}(0,1)\\&=\left(\rho^n+\frac{1}{\sigma-\rho}\frac{\xi}{\mu}(\sigma^n-\rho^n)\right)p_{ob}(0,1),\quad 1\le n\le n_e(2)\end{aligned} \tag{2-24}$$

$$\begin{aligned}p_{ob}(n,0)&=\rho^n\left(1+\frac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}\left(\left(\frac{\sigma}{\rho}\right)^{n_e(2)+1}-1\right)\right)p_{ob}(0,1)\\&=\left(\rho^n+\frac{1}{\sigma-\rho}\frac{\xi}{\mu}(\sigma^{n_e(2)+1}-\rho^{n_e(2)+1})\rho^{n-n_e(2)-1}\right)p_{ob}(0,1)\\&\quad n_e(2)+1\le n\le n_e(0)+1\end{aligned} \tag{2-25}$$

由于顾客的止步状态为$(n_e(0)+1,\ 0)$或$(n_e(2)+1,\ 2)$，因此，均衡状态下，顾客的单位社会福利为

$$\begin{aligned}U_{ob}=&(R-p)\lambda(1-p_{ob}(n_e(0)+1,0)-p_{ob}(n_e(2)+1,2))\\&-C\left(\sum_{n=1}^{n_e(0)+1}np_{ob}(n,0)+\sum_{n=0}^{n_e(2)+1}np_{ob}(n,2)\right)\end{aligned} \tag{2-26}$$

注释 2.2.1 对于完全可视情形下顾客的最优止步策略来说，定义顾客的最

优止步阈值为$(n^*(0), n^*(1), n^*(2))$ $(n^*(1)=0)$，则社会最优条件下的单位社会福利为

$$SW_{ob}(n^*(0), n^*(2)) = R\lambda(1-p_{ob}(n^*(0)+1,0)-p_{ob}(n^*(2)+1,2)) - C\left(\sum_{n=1}^{n^*(0)+1} np_{ob}(n,0)+\sum_{n=0}^{n^*(2)+1} np_{ob}(n,2)\right)$$

由于服务员在各个状态的最优止步阈值的显式无法得到，所以通过数值算例来佐证社会最优止步阈值策略的存在性，如图 2-2 所示。可以明显地观察到只要服务费用不是太高，一般情况下，$n_e(0) > n^*(0)$和 $n_e(2) > n^*(2)$是成立的。

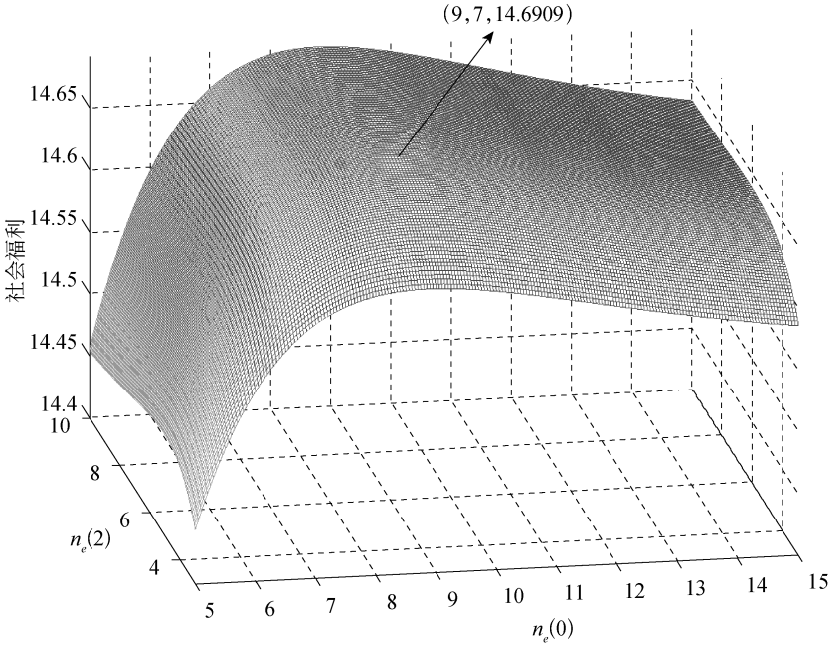

图 2-2　可中断启动-关闭策略和完全可视情形下的单位社会福利

2.2.3　完全不可视排队系统

2.2.3.1　纳什均衡

接下来，开始考虑完全不可视情形下顾客的均衡止步行为。在完全不可视情形下，有两种纯策略可供顾客选择：进入系统和止步离开。一个纯策略或者一个混合策略可以由进入概率 $q(0 \le q \le 1)$来描述。假设顾客潜在到达率为 Λ，均衡进入概率为 q_e，均衡混合止步策略下的有效到达率为 λ_e，且显然有 $\lambda_e = \Lambda q_e \le \Lambda < \mu$。状态转移图如图 2-3 所示。

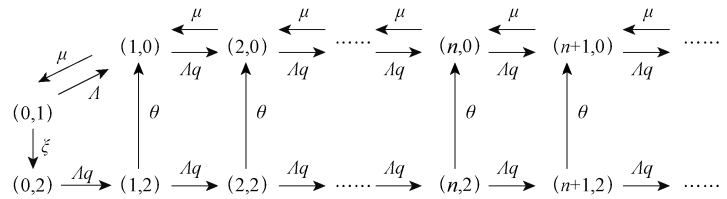

图 2-3　可中断启动-关闭策略和完全不可视情形下的状态转移图

定理 2.2.3　在具有可中断启动-关闭策略的完全不可视 M/M/1 排队系统中，假设 $\Lambda<\mu$，则“以概率 q_e 进入系统”为唯一的均衡混合止步策略，且

$$q_e=\begin{cases}\dfrac{1}{\Lambda}\left(\mu-\dfrac{1}{\dfrac{R-p}{C}-\dfrac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}}}\right), & R\in\left[\dfrac{C}{\mu}+\dfrac{C}{\theta}+p,\dfrac{C}{\mu-\Lambda}+\dfrac{C\left(1+\frac{\xi}{\theta^2}\right)}{1+\frac{\xi}{\theta}}+p\right)\\ 1, & R\in\left[\dfrac{C}{\mu-\Lambda}+\dfrac{C\left(1+\frac{\xi}{\theta^2}\right)}{1+\frac{\xi}{\theta}}+p,+\infty\right)\end{cases}\tag{2-27}$$

证明　考虑一个标记顾客，对于具有可中断启动-关闭策略的完全不可视 M/M/1 排队系统来说，顾客到达率为 $\Lambda q(\Lambda<\mu)$，首先导出此顾客的平均逗留时间。定义忙期开始时系统中的顾客数为 Q_b。以关闭期内是否有顾客到达为条件，得到 Q_b 的分布律为

$$b_j=\begin{cases}1-c^*(\Lambda q)+c^*(\Lambda q)u_0, & j=1\\ c^*(\Lambda q)u_{j-1}, & j>1\end{cases}\tag{2-28}$$

其中 $c^*(\Lambda q)$表示关闭期内没有顾客到达的概率，$u_j(j\geq 0)$表示有 j 个顾客于随后的启动期内到达的概率。由于启动期和关闭期都服从指数分布，所以有

$$c^*(\Lambda q)=\frac{\xi}{\Lambda q+\xi}\quad u_j=\int_0^{+\infty}\frac{(\Lambda q t)^j}{j!}\theta e^{-(\Lambda q+\theta)t}\mathrm{d}t$$

因此，得到忙期开始时系统中的平均顾客数 $E[Q_b]$为

$$E[Q_b]=1-c^*(\Lambda q)+\Lambda q c^*(\Lambda q)\frac{1}{\theta}\tag{2-29}$$

此顾客的平均附加延迟 $E[W_d]$为

$$E[W_d]=\frac{1-c^*(\Lambda q)}{E[Q_b]}+\frac{\Lambda q c^*(\Lambda q)}{\theta E[Q_b]}\frac{1}{\theta}\tag{2-30}$$

也就是说，附加延迟以概率$(1-c^*(\Lambda q))/E[Q_b]$等于零，以概率$(\Lambda q c^*(\Lambda q))/(\theta E[Q_b])$等于剩余的启动期。因此，此顾客的平均逗留时间为

$$E[W]=E[W_0]+E[W_d]=W_{un}(q)=\frac{1}{\mu-\Lambda q}+\frac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}} \tag{2-31}$$

其中 $E[W_0]=1/(\mu-\Lambda q)$为经典 M/M/1 排队系统的平均逗留时间，其平均剩余效用为

$$R-C\left(\frac{1}{\mu-\Lambda q}+\frac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}}\right)-p \tag{2-32}$$

当

$$R\in\left[\frac{C}{\mu}+\frac{C}{\theta}+p,\frac{C}{\mu-\Lambda}+\frac{C\left(1+\frac{\xi}{\theta^2}\right)}{1+\frac{\xi}{\theta}}+p\right)$$

时，式（2-32）在区间(0, 1)内具有唯一零点，即为式（2-27）的前半支。当

$$R\in\left[\frac{C}{\mu-\Lambda}+\frac{C\left(1+\frac{\xi}{\theta^2}\right)}{1+\frac{\xi}{\theta}}+p,+\infty\right)$$

时，式（2-32）对于任一 q 都为正数，因此 $q_e=1$，即式（2-27）的后半支。

因此，均衡状态下，单位社会福利为

$$U_{un}=\Lambda q_e\left(R-C\left(\frac{1}{\mu-\Lambda q_e}+\frac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}}\right)-p\right) \tag{2-33}$$

2.2.3.2　社会最优

继续考虑完全不可视情形下顾客的社会最优止步策略。定义顾客的社会最优混合止步策略为 q^*。对于社会计划者来说，决策的过程就是选择最优的进入概率 q^* 使得社会福利 $SW_{un}(q)$达到最大，其中

$$SW_{un}(q)=\Lambda q\left(R-C\left(\frac{1}{\mu-\Lambda q}+\frac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}}\right)\right) \tag{2-34}$$

下列引理给出了社会福利函数的特征。

引理 2.2.1　如果 $\rho=\Lambda/\mu<1$，则社会福利函数（2-34）单峰且在区间[0,1]上

为严格凹函数。

证明 重写 $SW_{un}(q)$，得到

$$SW_{un}(q)=\Lambda q(R-CWun(q))$$

其一阶和二阶导数分别为

$$SW'_{un}(q)=\Lambda(R-CW_{un}(q))-\Lambda CqW'_{un}(q)$$

$$SW''_{un}(q)=-2C\Lambda W'_{un}(q)-\Lambda CqW''_{un}(q)$$

因为 $W_{un}(q)$为严格单增的凸函数，所以在区间[0,1]上，有

$$W'_{un}(q)=\frac{\Lambda}{(\mu-\Lambda q)^2}>0,\quad W''_{un}(q)=\frac{2\Lambda^2}{(\mu-\Lambda q)^3}>0$$

即在区间[0,1]上，$SW''_{un}(q)<0$。

式（2-34）的一阶最优条件为

$$SW'_{un}(q)=\Lambda\left(R-C\left(\frac{1}{\mu-\Lambda q}+\frac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}}\right)\right)-\frac{\Lambda^2Cq}{(\mu-\Lambda q)^2} \tag{2-35}$$

解方程 $SW'_{un}(q)=0$ 得到唯一的最优解为

$$q^*_{fd}=\frac{\mu(C\theta^2-R\theta^2+C\xi-R\theta\xi)+\sqrt{-C\theta\mu(\theta+\xi)(C\theta^2-R\theta^2+C\xi-R\theta\xi)}}{\Lambda(-R\theta\xi+C\xi+C\theta^2-R\theta^2)} \tag{2-36}$$

定理 2.2.4 （1）如果 $0<q^*_{fd}<1$，则存在唯一的最优到达概率 $q^*=q^*_{fd}$；

（2）如果 $q^*_{fd}\geq 1$，则存在唯一的最优到达概率 $q^*=1$。

对于具有可中断启动-关闭策略的完全不可视排队系统来说，图 2-4 给出了单位社会福利关于顾客进入概率的敏感度，且观察到最优进入概率分别为 $q^*=0.9145$ 和 $q^*=1$，相应的最优社会福利分别为 $SW^*(q^*)=4.7668$ 和 $SW^*(q^*)=6.6333$。

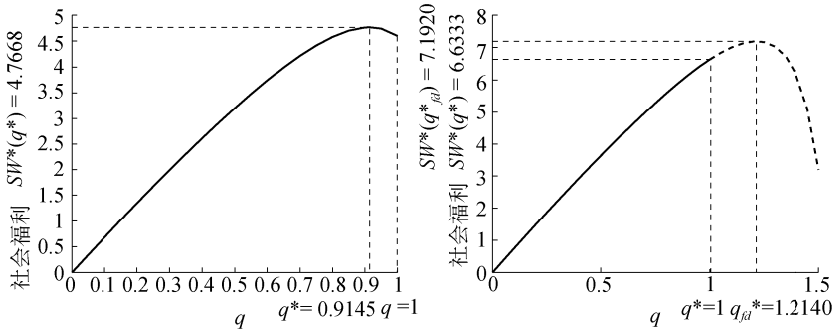

图 2-4 可中断启动-关闭策略下的最优止步策略和最大社会福利

2.2.3.3　利润最大

最后考虑垄断情形。显然，垄断者会对顾客收取使其利润最大化的服务费用 p_m，因此 $p_m+CW_{un}(q)=R$。垄断者的目的就是使 $\Lambda q p_m$ 达到最大且

$$p_m = R - C\left(\frac{1}{\mu-\Lambda q}+\frac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}}\right) \tag{2-37}$$

由于垄断者与社会计划者的目标一致，社会最优到达率 $\lambda^*=\Lambda q^*$ 可以通过对顾客收取费用得到，所以此费用也可以使得垄断者的利润达到最大，该费用为

$$p_m = p^* = R - CW(\lambda_{fd}^*) = R - C\left(\frac{1}{\mu-\lambda_{fd}^*}+\frac{1+\frac{\xi}{\theta^2}}{1+\frac{\xi}{\theta}}\right) \tag{2-38}$$

其中 $\lambda_{fd}^*=\Lambda q_{fd}^*$。

图 2-5 说明垄断服务费用 p_m 单调非增。因此，需求增加可导致垄断价格下降。这是由于到达率的增加会导致顾客等待时间的增长，所以随着到达率的增加，顾客不愿意支付较高的服务费用，即服务质量和服务需求成反比。

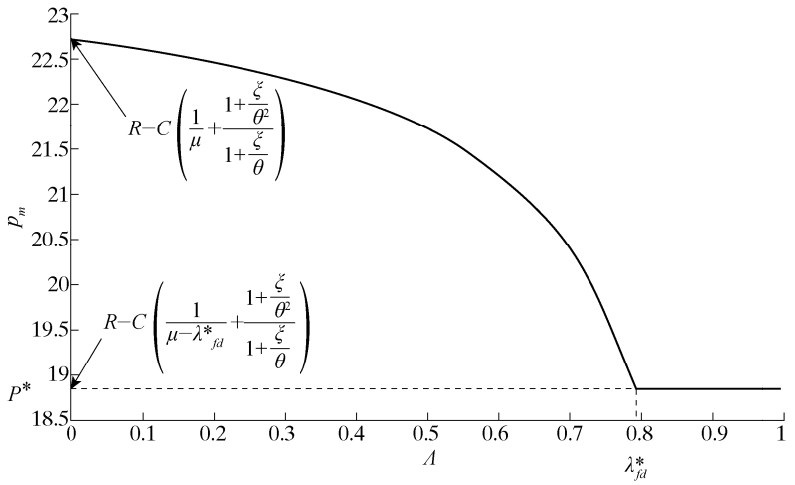

图 2-5　可中断启动-关闭策略下的垄断服务费用

2.3　可跳跃启动-关闭策略

类似于以上关于具有可中断启动-关闭策略的排队系统的分析，接下来开始

考虑具有可跳跃启动-关闭策略的排队系统中顾客的止步行为，并得到与其平行的结果。

不同于具有可中断启动-关闭策略的排队系统，关闭期内到达的顾客（如果有到达）不能被马上服务而必须在关闭期完成之后才能开始一个忙期，同时随后的启动期将被跳过。如果关闭期内没有顾客到达，服务员进入空闲期且期间一旦有顾客到达则其进入启动期，启动期结束后开始进入忙期。在此系统中，假设 $R>\frac{C}{\mu}+\frac{C}{\theta}+p$。否则，系统第一次空竭后就不会有顾客再进入系统。

2.3.1 完全可视排队系统

在具有可跳跃启动-关闭策略的排队系统中，如果均衡状态下到达顾客在观察到系统状态为(n, i)后决定进入系统，他的平均逗留时间为

$$(s_e(0), s_e(1), s_e(2))=\left(\frac{n+1}{\mu}, \frac{n+1}{\mu}+\frac{1}{\xi}, \frac{n+1}{\mu}+\frac{1}{\theta}\right)$$

平均剩余效用为

$$\begin{aligned}&(b_e(0), b_e(1), b_e(2))\\=&\left(R-\frac{C(n+1)}{\mu}-p, R-\frac{C(n+1)}{\mu}-\frac{C}{\xi}-p, R-\frac{C(n+1)}{\mu}-\frac{C}{\theta}-p\right)\end{aligned}$$

如果剩余效用为正值，则此顾客一定会进入系统。

定理 2.3.1 在具有可跳跃启动-关闭策略的完全可视 M/M/1 排队系统中，如果时刻 t 到达的顾客观测到系统状态为$(N(t), I(t))$，则存在阈值

$$\begin{aligned}&(n_e(0), n_e(1), n_e(2))\\=&\left(\left\lfloor\frac{(R-p)\mu}{C}\right\rfloor-1, \left\lfloor\frac{(R-p)\mu}{C}-\frac{\mu}{\xi}\right\rfloor-1, \left\lfloor\frac{(R-p)\mu}{C}-\frac{\mu}{\theta}\right\rfloor-1\right)\end{aligned}\tag{2-39}$$

且“$N(t)\leq n_e(I(t))$则进入系统，否则放弃离开”是唯一的均衡止步策略。

如果所有的顾客遵从式（2-39）给出的均衡阈值止步策略，则系统过程是一个状态空间为 $S_{ob}=\{(n, 0)|1\leq n\leq n_e(0)+1\}\cup\{(n, 1)|0\leq n\leq n_e(1)+1\}\cup\{(n, 2)|0\leq n\leq n_e(2)+1\}$的 Markov 链。状态转移图如图 2-6 所示。

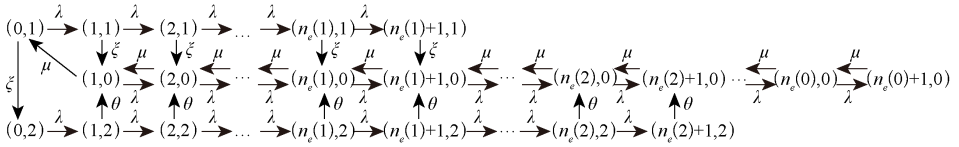

图 2-6 可跳跃启动-关闭策略和完全可视情形下的状态转移图

相应的系统稳态队长分布 $\{p_{ob}(n,i):(n,i)\in S_{ob}\}$ 为以下均衡方程组

$$p(0,1)(\lambda+\xi)=p(1,0)\mu \tag{2-40}$$

$$p(0,2)\lambda=p(0,1)\xi \tag{2-41}$$

$$p(1,0)(\lambda+\mu)=p(1,2)\theta+p(2,0)\mu+p(1,1)\xi \tag{2-42}$$

$$p(n,0)(\lambda+\mu)=p(n-1,0)\lambda+p(n,2)\theta+p(n,1)\xi+p(n+1,0)\mu,\quad 2\le n\le n_e(1)+1 \tag{2-43}$$

$$p(n,0)(\lambda+\mu)=p(n-1,0)\lambda+p(n,2)\theta+p(n+1,0)\mu,\quad n_e(1)+2\le n\le n_e(2)+1 \tag{2-44}$$

$$p(n,0)(\lambda+\mu)=p(n-1,0)\lambda+p(n+1,0)\mu,\ n_e(2)+2\le n\le n_e(0) \tag{2-45}$$

$$p(n_e(0)+1,0)\mu=p(n_e(0),0)\lambda \tag{2-46}$$

$$p(n,2)(\lambda+\theta)=p(n-1,2)\lambda,\ 1\le n\le n_e(2) \tag{2-47}$$

$$p(n_e(2)+1,2)\theta=p(n_e(2),2)\lambda \tag{2-48}$$

$$p(n,1)(\lambda+\xi)=p(n-1,1)\lambda,\ 1\le n\le n_e(1) \tag{2-49}$$

$$p(n_e(1)+1,1)\xi=p(n_e(1),1)\lambda \tag{2-50}$$

的唯一正则解。定义

$$\delta=\frac{\lambda}{\lambda+\xi}$$

迭代式 (2-45)、式(2-47)、式(2-49)，并同时考虑式 (2-41)、式(2-46)、式 (2-48)、式 (2-50)，得到

$$p(n,1)=\delta^n p(0,1),\ 0\le n\le n_e(1) \tag{2-51}$$

$$p(n_e(1)+1,1)=\frac{\lambda}{\xi}\delta^{n_e(1)}p(0,1) \tag{2-52}$$

$$p(n,0)=\rho^{n-n_e(2)-1}p(n_e(2)+1,0),\ n_e(2)+1\le n\le n_e(0)+1 \tag{2-53}$$

$$p(n_e(2)+1,0)\mu=p(n_e(2),0)\lambda+\xi\sigma^{n_e(2)}p(0,1) \tag{2-54}$$

$$p(n,2)=\frac{\xi}{\lambda}\sigma^n p(0,1),\ 1\le n\le n_e(2) \tag{2-55}$$

$$p(n_e(2)+1,2)=\frac{\xi}{\theta}\sigma^{n_e(2)}p(0,1) \tag{2-56}$$

鉴于式 (2-51) 至式(2-56)，首先考虑概率$\{p(n,0):1\le n\le n_e(1)+1\}$。从式 (2-43) 观察到它们是非齐次线性差分方程

$$\begin{aligned}\mu x_{n+1}-(\lambda+\mu)x_n+\lambda x_{n-1}&=-\theta p(n,2)-\xi p(n,1)\\&=-\left(\frac{\xi}{\lambda}\sigma^n\theta+\xi\delta^n\right)p(0,1),\ 2\le n\le n_e(1)\end{aligned} \tag{2-57}$$

的解。由于式（2-57）的非齐次部分是参数σ和δ的几何形式，所以它的特解x_n^{spec}可设形为$C_1\sigma^n + D_1\delta^n$。将$x_n^{spec} = C_1\sigma^n + D_1\delta^n$代入式（2-57），得到

$$\begin{cases} C_1 = \dfrac{-\dfrac{\xi\theta}{\lambda}\sigma}{\mu\sigma^2 - (\lambda+\mu)\sigma + \lambda} = \dfrac{\xi(\lambda+\theta)p(0,1)}{\lambda(\mu-\lambda-\theta)} \\ D_1 = \dfrac{-\xi\delta}{\mu\delta^2 - (\lambda+\mu)\delta + \lambda} = \dfrac{(\lambda+\xi)p(0,1)}{\mu-\lambda-\xi} \end{cases} \tag{2-58}$$

因此，式（2-57）的通解为

$$x_n^{gen} = A_1 1^n + B_1\rho^n + C_1\sigma^n + D_1\delta^n,\ \ 1 \le n \le n_e(1)+2 \tag{2-59}$$

其中A_1和B_1是未知系数。结合式（2-40），当$n=1$时，从式（2-59）得到

$$A_1 + \rho B_1 + C_1\sigma + D_1\delta = \frac{\lambda+\xi}{\mu}p(0,1) \tag{2-60}$$

而且当$n=2$时，将式（2-59）代入式（2-57）中并同时考虑式（2-40）、式（2-51）、式（2-55），得到

$$A_1 + \rho^2 B_1 + C_1\sigma^2 + D_1\delta^2 = \frac{1}{\mu}\left(\frac{(\lambda+\mu)(\lambda+\xi)}{\mu} - \xi(1-\sigma+\delta)\right)p(0,1) \tag{2-61}$$

解方程组（2-60）和（2-61）得到$A_1=0$,

$$B_1 = \left(\frac{\mu\xi - (\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)} - \frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)p(0,1)$$

因此，由式（2-59）得到

$$\begin{aligned} p(n,0) = &\left(\left(\frac{\mu\xi - (\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)} - \frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^n + \frac{\xi}{\mu-\lambda-\theta}\sigma^{n-1}\right. \\ &\left. + \frac{\lambda}{\mu-\lambda-\xi}\delta^{n-1}\right)p(0,1),\quad 1 \le n \le n_e(1)+2 \end{aligned} \tag{2-62}$$

类似地，导出稳态概率$\{p(n,0): n_e(1)+2 \le n \le n_e(2)\}$。它们是差分方程

$$\mu x_{n+1} - (\lambda+\mu)x_n + \lambda x_{n-1} = -\theta p(n,2) = -\frac{\theta\xi}{\lambda}\sigma^n p(0,1) \tag{2-63}$$

的解，其中$n_e(1)+1 \le n \le n_e(2)$。因为式（2-63）的非齐次部分是参数$\sigma$的几何形式，所以考虑形为$C_2\sigma^n$的特解。将$x_n^{spec} = C_2\sigma^n$代入式（2-63）中，得到

$$C_2 = C_1 = \frac{\xi(\lambda+\theta)p(0,1)}{\lambda(\mu-\lambda-\theta)} \tag{2-64}$$

因此，式（2-63）的通解为

$$x_n^{gen} = A_2 1^n + B_2\rho^n + C_2\sigma^n,\ \ n_e(1)+1 \le n \le n_e(2) \tag{2-65}$$

其中A_2和B_2为未知系数。当$n = n_e(1)+1$和$n = n_e(1)+2$时，将式（2-65）代入

式（2-62）中，得到

$$\begin{aligned}&A_2+B_2\rho^{n_e(1)+1}+\frac{\xi(\lambda+\theta)p(0,1)}{\lambda(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}\\&=\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^{n_e(1)+1}\right.\\&\left.+\frac{\xi}{\mu-\lambda-\theta}\sigma^{n_e(1)}+\frac{\lambda}{\mu-\lambda-\xi}\delta^{n_e(1)}\right)p(0,1)\end{aligned}\tag{2-66}$$

$$\begin{aligned}&A_2+B_2\rho^{n_e(1)+2}+\frac{\xi(\lambda+\theta)p(0,1)}{\lambda(\mu-\lambda-\theta)}\sigma^{n_e(1)+2}\\&=\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\mu-\lambda-\xi}-\frac{\mu\xi}{\mu-\lambda-\theta}\right)\rho^{n_e(1)+1}\right.\\&\left.+\frac{\mu\xi}{\mu-\lambda-\theta}\sigma^{n_e(1)+1}+\frac{\lambda(\lambda+\xi)}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\right)p(0,1)\end{aligned}\tag{2-67}$$

解方程组（2-66）和（2-67），得到

$$\begin{aligned}A_2=&\frac{\lambda-\rho}{1-\rho}\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^{n_e(1)+1}\right.\\&\left.+\frac{\xi}{\rho(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\right)p(0,1)\end{aligned}\tag{2-68}$$

$$\begin{aligned}B_2=&\frac{1-\lambda}{(1-\rho)\rho^{n_e(1)+1}}\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^{n_e(1)+1}\right.\\&\left.+\frac{\xi(1-\mu)}{(1-\lambda)(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\right)p(0,1)\end{aligned}\tag{2-69}$$

所以由式（2-65）可得

$$\begin{aligned}&p(n,0)\\&=\left(\frac{\lambda-\rho}{1-\rho}\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^{n_e(1)+1}+\frac{\xi}{\rho(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}\right.\right.\\&\left.+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\right)+\frac{1-\lambda}{(1-\rho)\rho^{n_e(1)+1}}\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^{n_e(1)+1}\right.\\&\left.\left.+\frac{\xi(1-\mu)}{(1-\lambda)(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\right)\rho^n+\frac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}\sigma^n\right)p(0,1)\\&n_e(1)+1\le n\le n_e(2)\end{aligned}\tag{2-70}$$

根据式（2-54），得到

$$\begin{aligned} p(n_e(2)+1,0) &= \rho p(n_e(2),0)+\frac{\xi}{\mu}\sigma^{n_e(2)}p(0,1) \\ &= \rho(A_2+B_2\rho^{n_e(2)}+C_2\sigma^{n_e(2)})+\frac{\xi}{\mu}\sigma^{n_e(2)}p(0,1) \end{aligned} \tag{2-71}$$

又根据式（2-53），有

$$\begin{aligned} p(n,0) &= \rho^{n-n_e(2)-1}\left(\rho(A_2+B_2\rho^{n_e(2)}+C_2\sigma^{n_e(2)})+\frac{\xi}{\mu}\sigma^{n_e(2)}p(0,1)\right), \\ & n_e(2)+2\le n\le n_e(0)+1 \end{aligned} \tag{2-72}$$

这样,就把所有的稳态概率表示成了关于$p(0,1)$的形式,且由式(2-51)、式(2-52)、式（2-55）、式（2-56）、式（2-62）、式（2-70）至式（2-72），以及平衡方程

$$\sum_{n=1}^{n_e(0)+1}p(n,0)+\sum_{n=0}^{n_e(1)+1}p(n,1)+\sum_{n=0}^{n_e(2)+1}p(n,2)=1 \tag{2-73}$$

可得到稳态概率$p(0,1)$的表达式。以下定理给出了系统稳态队长分布。

定理 2.3.2 在具有可跳跃启动-关闭策略的完全可视排队系统中,如果顾客均采用如式（2-39）所示的均衡阈值止步策略，且$\sigma\neq1\neq\rho\neq\sigma$，$\delta\neq1\neq\rho\neq\delta$，则系统稳态队长分布$\{p_{ob}(n,i):(n,i)\in S_{ob}\}$为

$$p_{ob}(n,1)=\delta^n p_{ob}(0,1),\ 1\le n\le n_e(1) \tag{2-74}$$

$$p_{ob}(n_e(1)+1,1)=\frac{\lambda}{\xi}\delta^{n_e(1)}p_{ob}(0,1) \tag{2-75}$$

$$p_{ob}(n,2)=\frac{\xi}{\lambda}\sigma^n p_{ob}(0,1),\ 1\le n\le n_e(2) \tag{2-76}$$

$$p_{ob}(n_e(2)+1,2)=\frac{\xi}{\theta}\sigma^{n_e(2)}p_{ob}(0,1) \tag{2-77}$$

$$\begin{aligned} p_{ob}(n,0)=&\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^n+\frac{\xi}{\mu-\lambda-\theta}\sigma^{n-1}\right. \\ &\left.+\frac{\lambda}{\mu-\lambda-\xi}\delta^{n-1}\right)p_{ob}(0,1),\ 1\le n\le n_e(1) \end{aligned} \tag{2-78}$$

$$\begin{aligned} p_{ob}(n,0)=&\frac{\lambda-\rho}{1-\rho}\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^{n_e(1)+1}+\frac{\xi}{\rho(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}\right. \\ &+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}+\frac{1-\lambda}{(1-\rho)\rho^{n_e(1)+1}}\left(\left(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\right)\rho^{n_e(1)+1}\right. \\ &\left.\left.+\frac{\xi(1-\mu)}{(1-\lambda)(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\right)\rho^n+\frac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}\sigma^n\right)p_{ob}(0,1), \\ & n_e(1)+1\le n\le n_e(2) \end{aligned} \tag{2-79}$$

$$
\begin{aligned}
p_{ob}(n,0)=\Bigg(&\rho\Bigg(\frac{\lambda-\rho}{1-\rho}\Bigg(\Bigg(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\Bigg)\rho^{n_e(1)+1}+\frac{\xi}{\rho(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}\\
&+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\Bigg)+\frac{1-\lambda}{(1-\rho)\rho^{n_e(1)+1}}\Bigg(\Bigg(\frac{\mu\xi-(\lambda+\xi)^2}{\lambda(\mu-\lambda-\xi)}-\frac{\mu\xi}{\lambda(\mu-\lambda-\theta)}\Bigg)\rho^{n_e(1)+1}\\
&+\frac{\xi(1-\mu)}{(1-\lambda)(\mu-\lambda-\theta)}\sigma^{n_e(1)+1}+\frac{\lambda+\xi}{\mu-\lambda-\xi}\delta^{n_e(1)+1}\Bigg)\rho^{n_e(2)}+\frac{\xi(\lambda+\theta)}{\lambda(\mu-\lambda-\theta)}\sigma^{n_e(2)}\Bigg)\\
&+\frac{\xi}{\mu}\sigma^{n_e(2)}\Bigg)\rho^{n-n_e(2)-1}p_{ob}(0,1),\ n_e(2)+1\le n\le n_e(0)+1
\end{aligned}
\tag{2-80}
$$

其中 $p_{ob}(0, 1)$可以由平衡方程（2-73）得到。

因为顾客的止步状态为$(n_e(0)+1, 0)$、$(n_e(1)+1, 1)$或$(n_e(2)+1, 2)$，因此，均衡状态下，单位社会福利为

$$
\begin{aligned}
U_{ob}=&(R-p)\lambda(1-p_{ob}(n_e(0)+1,0)-p_{ob}(n_e(1)+1,1)-p_{ob}(n_e(2)+1,2))\\
&-C\Bigg(\sum_{n=1}^{n_e(0)+1}np_{ob}(n,0)+\sum_{n=0}^{n_e(1)+1}np_{ob}(n,1)+\sum_{n=0}^{n_e(2)+1}np_{ob}(n,2)\Bigg)
\end{aligned}
\tag{2-81}
$$

2.3.2　完全不可视排队系统

2.3.2.1　纳什均衡

首先, 考虑完全不可视情形下顾客的均衡止步行为, 状态转移图如图 2-7 所示。

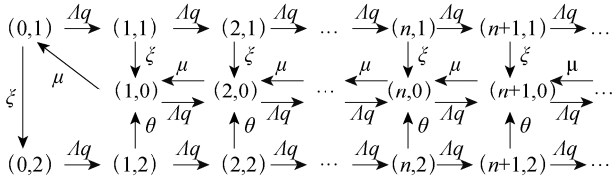

图 2-7　可跳跃启动-关闭策略和完全不可视情形下的状态转移图

定理 2.3.3　在具有可跳跃启动-关闭策略的完全不可视 M/M/1 排队系统中，假设 $\Lambda<\mu$，则“以概率 q_e 进入系统”为唯一的均衡混合止步策略且

$$
q_e=\begin{cases}\dfrac{1}{\Lambda}x^*, & R\in\left[\dfrac{C}{\mu}+\dfrac{C}{\theta}+p,\dfrac{C}{\mu-\Lambda}+\dfrac{C(\Lambda(\Lambda+\xi)\theta^2+\xi^3(\Lambda+\theta))}{\Lambda(\Lambda+\xi)\theta^2\xi+\theta\xi^3(\Lambda+\theta)}+p\right)\\ 1, & R\in\left[\dfrac{C}{\mu-\Lambda}+\dfrac{C(\Lambda(\Lambda+\xi)\theta^2+\xi^3(\Lambda+\theta))}{\Lambda(\Lambda+\xi)\theta^2\xi+\theta\xi^3(\Lambda+\theta)}+p,+\infty\right)\end{cases}
\tag{2-82}
$$

其中 x^*是下列方程

$$
R-C\left(\frac{1}{\mu-x}+\frac{x(x+\xi)\theta^2+\xi^3(x+\theta)}{x(x+\xi)\theta^2\xi+\theta\xi^3(x+\theta)}\right)-p=0
\tag{2-83}
$$

的唯一非负可行解。

证明 标记一个顾客。由于顾客到达率为 $\Lambda q(\Lambda<\mu)$，导出此顾客的平均逗留时间。以关闭期内是否有顾客到达为条件，得到忙期开始时的顾客数 Q_b 的分布律为

$$b_j = c_j + c^*(\Lambda q)u_{j-1},\quad j\geq 1 \tag{2-84}$$

其中

$$c_j = \int_0^{+\infty}\frac{(\Lambda q t)^j}{j!}\theta e^{-(\Lambda q+\xi)t}\mathrm{d}t$$

表示关闭期内有 j 个顾客到达的概率。因此，得到忙期开始时系统中的平均顾客数 $E[Q_b]$为

$$E[Q_b] = \Lambda q\frac{1}{\xi} + c^*(\Lambda q)\left(\Lambda q\frac{1}{\theta}+1\right) \tag{2-85}$$

因此平均附加延迟 $E[W_d]$为

$$E[W_d] = \frac{\frac{\Lambda q}{\xi}}{E[Q_b]}\frac{1}{\xi} + \frac{c^*(\Lambda q)(\Lambda q\frac{1}{\theta}+1)}{E[Q_b]}\frac{1}{\theta} \tag{2-86}$$

也就是说，附加延迟以概率 $\Lambda q/(\xi E[Q_b])$ 等于剩余关闭期，以概率 $c^*(\Lambda q)/E[Q_b]$ 等于整个启动期，并以概率 $\Lambda q c^*(\Lambda q)/(\theta E[Q_b])$ 等于剩余启动期。因此，平均逗留时间为

$$E[W] = E[W_0] + E[W_d] = W_{un}(q) = \frac{1}{\mu-\Lambda q} + \frac{\Lambda q(\Lambda q+\xi)\theta^2+\xi^3(\Lambda q+\theta)}{\Lambda q(\Lambda q+\xi)\theta^2\xi+\theta\xi^3(\Lambda q+\theta)} \tag{2-87}$$

且平均剩余效用为

$$R - C\left(\frac{1}{\mu-\Lambda q} + \frac{\Lambda q(\Lambda q+\xi)\theta^2+\xi^3(\Lambda q+\theta)}{\Lambda q(\Lambda q+\xi)\theta^2\xi+\theta\xi^3(\Lambda q+\theta)}\right) - p \tag{2-88}$$

当

$$R\in\left[\frac{C}{\mu}+\frac{C}{\theta}+p, \frac{C}{\mu-\Lambda}+\frac{C(\Lambda(\Lambda+\xi)\theta^2+\xi^3(\Lambda+\theta))}{\Lambda(\Lambda+\xi)\theta^2\xi+\theta\xi^3(\Lambda+\theta)}+p\right)$$

时，式（2-88）在区间（0, 1）内有唯一零点，即为式（2-82）的前半支。当

$$R\in\left[\frac{C}{\mu-\Lambda}+\frac{C(\Lambda(\Lambda+\xi)\theta^2+\xi^3(\Lambda+\theta))}{\Lambda(\Lambda+\xi)\theta^2\xi+\theta\xi^3(\Lambda+\theta)}+p, +\infty\right)$$

时，式（2-88）的值对于任何一个 q 都为正值，因此唯一的均衡进入概率为 $q_e=1$，即为式（2-82）的后半支。

因此，均衡状态下，单位社会福利为

$$U_{un}=\Lambda q_e\left(R-C\left(\frac{1}{\mu-\Lambda q_e}+\frac{\Lambda q_e(\Lambda q_e+\xi)\theta^2+\xi^3(\Lambda q_e+\theta)}{\Lambda q_e(\Lambda q_e+\xi)\theta^2\xi+\theta\xi^3(\Lambda q_e+\theta)}\right)-p\right) \tag{2-89}$$

2.3.2.2　社会最优

可跳跃启动-关闭策略下的社会福利函数 $SW_{un}(q)$为

$$SW_{un}(q)=\Lambda q\left(R-C\left(\frac{1}{\mu-\Lambda q}+\frac{\Lambda q(\Lambda q+\xi)\theta^2+\xi^3(\Lambda q+\theta)}{\Lambda q(\Lambda q+\xi)\theta^2\xi+\theta\xi^3(\Lambda q+\theta)}\right)\right) \tag{2-90}$$

它的一阶导数为

$$\begin{aligned}SW'_{un}(q)=&\Lambda\left(R-C\left(\frac{1}{\mu-\Lambda q}+\frac{\Lambda q(\Lambda q+\xi)\theta^2+\xi^3(\Lambda q+\theta)}{\Lambda q(\Lambda q+\xi)\theta^2\xi+\theta\xi^3(\Lambda q+\theta)}\right)\right)\\&-\Lambda qC\left(\frac{\Lambda(\Lambda q+\xi)\theta^2+\Lambda^2q\theta^2+\xi^3\Lambda}{\Lambda q(\Lambda q+\xi)\theta^2\xi+\theta\xi^3(\Lambda q+\theta)}+\frac{\Lambda}{(\mu-\Lambda q)^2}\right.\\&\left.-\frac{\Lambda q(\Lambda q+\xi)\theta^2+\xi^3(\Lambda q+\theta)}{(\Lambda q(\Lambda q+\xi)\theta^2\xi+\theta\xi^3(\Lambda q+\theta))^2}(\Lambda(\Lambda q+\xi)\theta^2\xi+\Lambda^2q\theta^2\xi+\theta\xi^3\Lambda)\right)\end{aligned} \tag{2-91}$$

因此最优进入概率 q^*和方程 $SW'_{un}(q)=0$ 的解 q^*_{fd} 存在以下关系。

定理 2.3.4　（1）如果 $0<q^*_{fd}<1$ 和 $SW''_{un}(q^*_{fd})\le 0$，存在唯一 $q^*=q^*_{fd}$；

（2）如果 $0<q^*_{fd}<1$ 和 $SW''_{un}(q^*_{fd})>0$ 或者 $q^*_{fd}\ge 1$，存在唯一 $q^*=1$。

对于具有可跳跃启动-关闭策略的完全不可视排队系统来说，图 2-8 给出了单位社会福利关于顾客进入概率的敏感度，且观察到最优进入概率分别为 $q^*=0.8626$ 和 $q^*=1$，相应的最优社会福利分别为 $SW^*(q^*)=12.3322$ 和 $SW^*(q^*)=11.9444$。

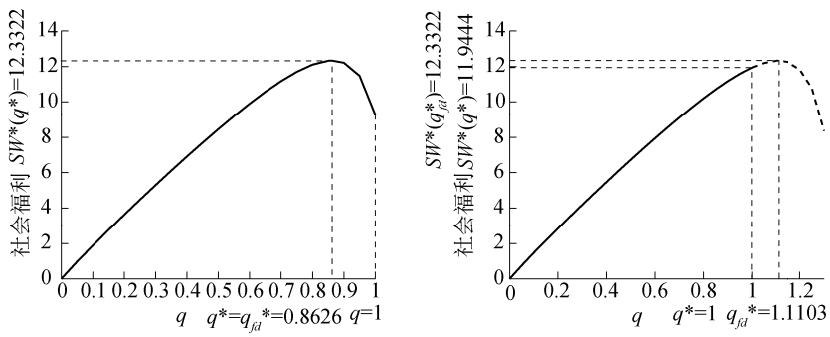

图 2-8　可跳跃启动-关闭策略下的最优止步策略和最大社会福利

现在，给出一个具体的充分条件来保证社会福利函数（2-90）单峰且为严格凹函数，并分析在此条件下的最优进入概率。

引理 2.3.1　如果 $\rho=\Lambda/\mu<1$ 和 $\theta=\xi$，则社会福利函数（2-90）在区间[0, 1]内单峰且为严格凹函数。

证明 标记 $\lambda_q = \Lambda q$ 为 q 的线性函数。只需证明如果 $\theta = \xi$，则 $SW''_{un}(\lambda_q) < 0$。因为 $W_{un}(\lambda_q)$ 的一阶和二阶导数分别为

$$\begin{aligned} W'_{un}(\lambda_q) &= \frac{1}{(\mu-\lambda_q)^2} + \frac{\xi(\theta-\xi)(\theta\xi+2\lambda_q\theta+\lambda_q^2)}{(\lambda_q^2\theta+\lambda_q\theta\xi+\xi^2\lambda_q+\xi^2\theta)^2} \\ W''_{un}(\lambda_q) &= \frac{2}{(\mu-\lambda_q)^3} + \frac{2\theta\xi(\xi-\theta)(\lambda_q+\xi)(\lambda_q^2+3\lambda_q\theta-\lambda_q\xi+\xi^2)}{(\lambda_q^2\theta+\lambda_q\theta\xi+\xi^2\lambda_q+\xi^2\theta)^3} \end{aligned} \tag{2-92}$$

所以得到

$$\begin{cases} W'_{un}(\lambda_q) > 0, & \text{如果}\theta \geq \xi \\ W'_{un}(\lambda_q) < (>=) 0, & \text{如果}\theta < \xi \end{cases}$$

$$\begin{cases} W''_{un}(\lambda_q) > 0, & \text{如果}\theta \leq \xi \\ W''_{un}(\lambda_q) < (>=) 0, & \text{如果}\theta > \xi \end{cases}$$

可见 $\theta = \xi$ 是 $W'_{un}(\lambda_q) > 0$ 和 $W''_{un}(\lambda_q) > 0$ 的充分条件。基于

$$\begin{aligned} SW''_{un}(\lambda_q) &= -2CW'_{un}(\lambda_q) - \lambda_q CW''_{un}(\lambda_q) \\ &= -\frac{2C\mu}{(\mu-\lambda_q)^3} - \frac{2C\xi(\theta-\xi)(\theta^2\xi^3+(\lambda_q^3+3\lambda_q^2\theta+3\lambda_q\theta^2)\xi^2+\lambda_q^3\theta(\xi-\theta))}{((\lambda_q+\theta)\xi^2+\xi\lambda_q\theta+\lambda_q^2\theta)^3} \end{aligned}$$

可知，如果 $\theta = \xi$，则 $SW''_{un}(\lambda_q) < 0$。

当条件 $\theta = \xi$ 满足时，式（2-90）的一阶最优条件为

$$SW'_{un}(q) = \frac{\Lambda(R\theta\mu^2 - 2R\theta\mu\Lambda q + R\theta\Lambda^2 q^2 - C\theta\mu + C\mu^2 + 2C\mu\Lambda q - \Lambda^2 q^2 C)}{\theta(\mu-\Lambda q)^2} \tag{2-93}$$

解方程 $SW'_{un}(q) = 0$ 得到唯一的最优可行解

$$q^*_{fd} = \frac{-R\theta\mu + C\mu + \sqrt{R\theta^2 C\mu - C^2\theta\mu}}{\Lambda(C-R\theta)} \tag{2-94}$$

定理 2.3.5 （1）如果 $0 < q^*_{fd} < 1$，存在唯一的最优进入概率 $q^* = q^*_{fd}$；

（2）如果 $q^*_{fd} \geq 1$，存在唯一的最优进入概率 $q^* = 1$。

图 2-9 给出了当 $\theta = \xi$ 时单位社会福利关于顾客进入概率的敏感度。最优进入概率分别为 $q^* = 0.7799$ 和 $q^* = 1$，最优社会福利分别为 $SW^*(q^*) = 4.0668$ 和 $SW^*(q^*) = 6.3$。

2.3.2.3 利润最大

现考虑垄断情形。对于具有可跳跃启动-关闭策略的完全不可视排队系统来说，垄断者的目的就是使得 $\Lambda q p_m$ 达到最大，且

$$p_m = p^* = R - C\left(\frac{1}{\mu-\lambda^*_{fd}} + \frac{\lambda^*_{fd}(\lambda^*_{fd}+\xi)\theta^2+\xi^3(\lambda^*_{fd}+\theta)}{\lambda^*_{fd}(\lambda^*_{fd}+\xi)\theta^2\xi+\theta\xi^3(\lambda^*_{fd}+\theta)}\right) \tag{2-95}$$

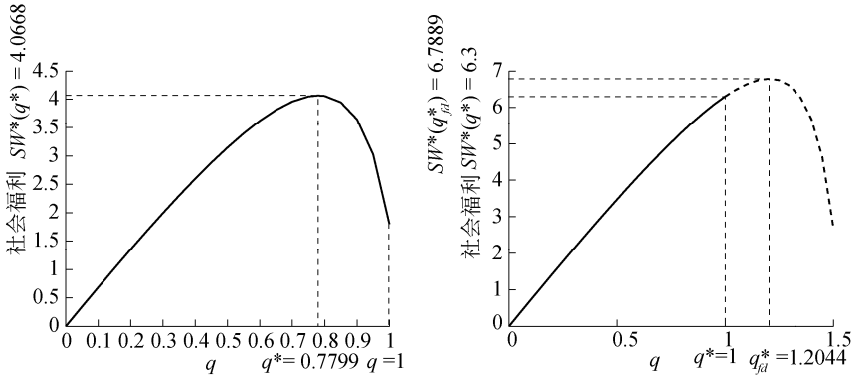

图 2-9　可跳跃启动-关闭策略下的最优止步策略和最大社会福利（$\theta=\xi$）

其中 $\lambda_{fd}^{*}=\Lambda q_{fd}^{*}$。图 2-10 说明垄断服务费用 p_m 单调非增。

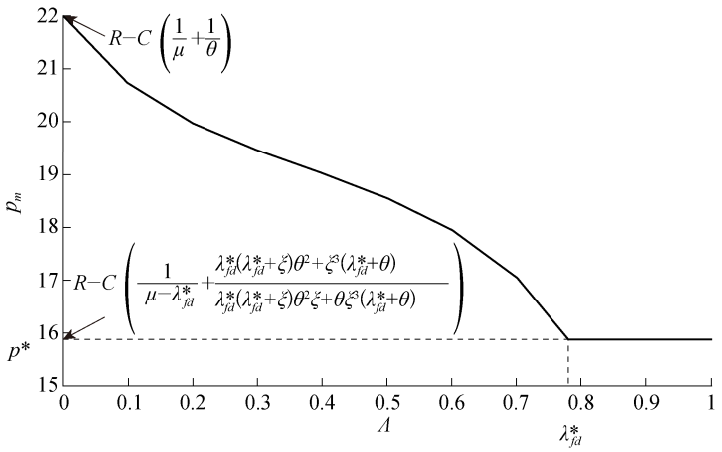

图 2-10　可跳跃启动-关闭策略下的垄断服务费用

2.4　非中断启动-关闭策略

最后，考虑具有非中断启动-关闭策略的排队系统。一旦系统进入关闭期，服务员则不可能中断此关闭期及随后的启动期。即使关闭期内有顾客到达，也必须等到启动期结束后才能对其进行服务。

2.4.1　完全可视排队系统

在具有非中断启动-关闭策略的排队系统中，如果均衡状态下到达顾客在观

察到系统状态为(n, i)后决定进入系统，他的平均逗留时间为

$$(s_e(0), s_e(1), s_e(2)) = \left(\frac{n+1}{\mu}, \frac{n+1}{\mu} + \frac{1}{\theta} + \frac{1}{\xi}, \frac{n+1}{\mu} + \frac{1}{\theta} \right)$$

则平均剩余效用分别为

$$\begin{aligned} &(b_e(0), b_e(1), b_e(2)) \\ &= \left(R - \frac{C(n+1)}{\mu} - p, R - \frac{C(n+1)}{\mu} - C\left(\frac{1}{\theta} + \frac{1}{\xi} \right) - p, R - \frac{C(n+1)}{\mu} - \frac{C}{\theta} - p \right) \end{aligned}$$

如果剩余效用为正值，则此顾客一定会进入系统。

定理 2.4.1 在具有非中断启动-关闭策略的完全可视 M/M/1 排队系统中，如果时刻 t 到达的顾客观测到系统状态为$(N(t), I(t))$，则存在阈值

$$\begin{aligned} &(n_e(0), n_e(1), n_e(2)) \\ &= \left(\left\lfloor \frac{(R-p)\mu}{C} \right\rfloor - 1, \left\lfloor \frac{(R-p)\mu}{C} - \mu\left(\frac{1}{\theta} + \frac{1}{\xi} \right) \right\rfloor - 1, \left\lfloor \frac{(R-p)\mu}{C} - \frac{\mu}{\theta} \right\rfloor - 1 \right) \end{aligned} \tag{2-96}$$

且“$N(t) \le n_e(I(t))$则进入系统，否则放弃离开”是唯一的均衡止步策略。

如果所有的顾客遵从式（2-96）给出的均衡阈值止步策略，则系统过程是一个状态空间为 $S_{ob}=\{(n, 0) \mid 1 \le n \le n_e(0)+1\} \cup \{(n, 1) \mid 0 \le n \le n_e(1)+1\} \cup \{(n, 2) \mid 0 \le n \le n_e(2)+1\}$的 Markov 链，且状态转移图如图 2-11 所示。

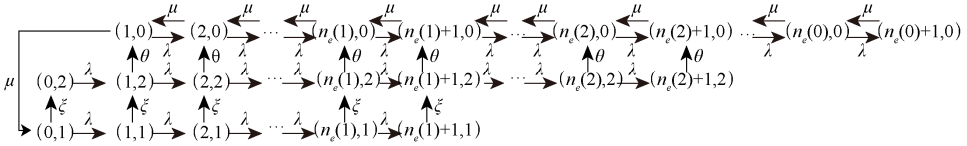

图 2-11 非中断启动-关闭策略和完全可视情形下的状态转移图

相应的系统稳态分布$\{p_{ob}(n, i) : (n, i) \in S_{ob}\}$为以下均衡方程组

$$p(0,1)(\lambda + \xi) = p(1,0)\mu \tag{2-97}$$

$$p(0,2)\lambda = p(0,1)\xi \tag{2-98}$$

$$p(1,0)(\lambda + \mu) = p(1,2)\theta + p(2,0)\mu \tag{2-99}$$

$$\begin{aligned} p(n,0)(\lambda + \mu) = p(n-1,0)\lambda + p(n,2)\theta + p(n+1,0)\mu, \\ 2 \le n \le n_e(2)+1 \end{aligned} \tag{2-100}$$

$$p(n,0)(\lambda + \mu) = p(n-1,0)\lambda + p(n+1,0)\mu,\ n_e(2)+2 \le n \le n_e(0) \tag{2-101}$$

$$p(n_e(0)+1,0)\mu = p(n_e(0),0)\lambda \tag{2-102}$$

$$p(n,2)(\lambda + \theta) = p(n-1,2)\lambda + p(n,1)\xi, 1 \le n \le n_e(1)+1 \tag{2-103}$$

$$p(n,2)(\lambda + \theta) = p(n-1,2)\lambda,\ n_e(1)+2 \le n \le n_e(2) \tag{2-104}$$

$$p(n_e(2)+1,2)\theta = p(n_e(2),2)\lambda \tag{2-105}$$

$$p(n,1)(\lambda+\xi) = p(n-1,1)\lambda, 1 \le n \le n_e(1) \tag{2-106}$$

$$p(n_e(1)+1,1)\xi = p(n_e(1),1)\lambda \tag{2-107}$$

的唯一正则解。迭代式（2-101）、式（2-104）和式（2-106），并同时考虑式（2-102）、式（2-105）、式（2-107），得到

$$p(n,1) = \delta^n p(0,1),\ 0 \le n \le n_e(1) \tag{2-108}$$

$$p(n_e(1)+1,1) = \frac{\lambda}{\xi}\delta^{n_e(1)} p(0,1) \tag{2-109}$$

$$p(n,0) = \rho^{n-n_e(2)-1} p(n_e(2)+1,0),\ n_e(2)+1 \le n \le n_e(0)+1 \tag{2-110}$$

$$p(n_e(2)+1,0)\mu = p(n_e(2),0)\lambda + p(n_e(2)+1,2)\theta \tag{2-111}$$

$$p(n,2) = \sigma^{n-n_e(1)-1} p(n_e(1)+1,2),\ n_e(1)+1 \le n \le n_e(2) \tag{2-112}$$

$$p(n_e(2)+1,2) = \frac{\lambda}{\theta}\sigma^{n_e(2)-n_e(1)-1} p(n_e(1)+1,2) \tag{2-113}$$

然后，考虑式（2-108）并迭代式（2-103），得到

$$\begin{aligned} p(n,2) &= p(n-1,2)\sigma + \frac{\xi}{\lambda+\theta}\delta^n p(0,1) \\ &= \sigma\left(p(n-1,2) + \frac{\xi}{\lambda}\delta^n p(0,1) \right) \\ &= \frac{\xi}{\lambda}\left(\sigma^n + \frac{\sigma\delta(\delta^n-\sigma^n)}{\delta-\sigma} \right) p(0,1),\ 1 \le n \le n_e(1)+1 \end{aligned} \tag{2-114}$$

根据式（2-112）至式（2-114）有

$$p(n,2) = \frac{\xi}{\lambda}\left(1 + \frac{\sigma\delta\left(\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1} - 1 \right)}{\delta-\sigma} \right)\sigma^n p(0,1),\ \ n_e(1)+1 \le n \le n_e(2) \tag{2-115}$$

$$p(n_e(2)+1,2) = \frac{\xi}{\theta}\left(1 + \frac{\sigma\delta\left(\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1} - 1 \right)}{\delta-\sigma} \right)\sigma^{n_e(2)} p(0,1) \tag{2-116}$$

鉴于式（2-114）至式（2-116），考虑$\{p\ (n,\ 0) : 1 \le n \le n_e\ (1)\}$。它们是非齐次线性差分方程

$$\mu x_{n+1} - (\lambda+\mu)x_n + \lambda x_{n-1} = \frac{\theta\xi}{\lambda}\left(\sigma^n + \frac{\sigma\delta(\delta^n-\sigma^n)}{\delta-\sigma} \right) p(0,1) \tag{2-117}$$

的解，其中$2 \le n \le n_e(1)+1$。因为式（2-117）的非齐次部分是σ和δ的几何形式，所以考虑形为$C_1\sigma^n + D_1\delta^n$的特解。将$x_n^{spec} = C_1\sigma^n + D_1\delta^n$代入式（2-117）中，得到

$$\begin{cases} C_1 = \dfrac{\dfrac{-\theta\xi}{\lambda}\left(1-\dfrac{\sigma\delta}{\delta-\sigma}\right)\sigma p(0,1)}{\mu\sigma^2-(\lambda+\mu)\sigma+\lambda} = \dfrac{\xi(\lambda+\theta)(\lambda+\xi-\theta)p(0,1)}{\lambda(\theta-\xi)(\lambda+\theta-\mu)} \\ D_1 = \dfrac{\dfrac{-\theta\xi}{\lambda}\dfrac{\sigma\delta}{\delta-\sigma}\delta p(0,1)}{\mu\delta^2-(\lambda+\mu)\delta+\lambda} = -\dfrac{\theta(\lambda+\xi)p(0,1)}{(\theta-\xi)(\lambda+\xi-\mu)} \end{cases} \tag{2-118}$$

因此，式（2-117）的通解为

$$x_n^{gen} = A_1 1^n + B_1\rho^n + C_1\sigma^n + D_1\delta^n, \quad 1 \le n \le n_e(1)+2 \tag{2-119}$$

其中 A_1 和 B_1 为未知系数。当 $n=1$ 时，从式（2-119）得到

$$A_1 + B_1\rho + C_1\sigma + D_1\delta = \frac{\lambda+\xi}{\mu}p(0,1) \tag{2-120}$$

当 $n=2$ 时，将式（2-119）代入式（2-99）中，并考虑式（2-97）和式（2-114），得到

$$A_1 + B_1\rho^2 + C_1\sigma^2 + D_1\delta^2 = \frac{1}{\mu}\left(\frac{(\lambda+\mu)(\lambda+\xi)}{\mu} - \frac{\xi\sigma\theta(1+\delta)}{\lambda}\right)p(0,1) \tag{2-121}$$

解方程组（2-120）和（2-121），得到 $A_1 = 0$,

$$B_1 = \left(1+\frac{\mu}{\theta-\xi}+\frac{\theta\xi}{\lambda(\lambda+\theta-\mu)}+\frac{\xi(\lambda+\xi)}{(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)p(0,1)$$

因此，由式（2-119）得到

$$\begin{aligned} p(n,0) = &\left(\left(1+\frac{\mu}{\theta-\xi}+\frac{\theta\xi}{\lambda(\lambda+\theta-\mu)}+\frac{\xi(\lambda+\xi)}{(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)\rho^n\right. \\ &\left.+\frac{\xi(\lambda+\xi-\theta)}{(\theta-\xi)(\lambda+\theta-\mu)}\sigma^{n-1} - \frac{\theta\lambda}{(\theta-\xi)(\lambda+\xi-\mu)}\delta^{n-1}\right)p(0,1), \\ &1 \le n \le n_e(1)+2 \end{aligned} \tag{2-122}$$

类似地，开始推导稳态概率 $\{p(n,0): n_e(1)+1 \le n \le n_e(2)\}$。它们是非齐次线性差分方程

$$\mu x_{n+1} - (\lambda+\mu)x_n + \lambda x_{n-1} = -\frac{\theta\xi}{\lambda}\left(1+\frac{\sigma\delta\left(\left(\dfrac{\delta}{\sigma}\right)^{n_e(1)+1}-1\right)}{\delta-\sigma}\right)\sigma^n p(0,1) \tag{2-123}$$

的解，其中 $n_e(1)+1 \le n \le n_e(2)$。因为式（2-133）的非齐次部分是参数 σ 的几何形式，考虑形为 $C_2\sigma^n$ 的特解，将 $x_n^{spec} = C_2\sigma^n$ 代入式（2-123），得到

$$C_2 = \left(\frac{\xi(\lambda+\theta)(\lambda+\xi-\theta)}{\lambda(\theta-\xi)(\lambda+\theta-\mu)} - \frac{\xi(\lambda+\theta)}{(\theta-\xi)(\lambda+\theta-\mu)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1}\right)p(0,1) \tag{2-124}$$

因此，式（2-123）的通解为

$$x_n^{gen} = A_2 1^n + B_2\rho^n + C_2\sigma^n, \quad n_e(1)+1 \le n \le n_e(2) \tag{2-125}$$

其中 A_2 和 B_2 为未知系数。当 $n = n_e(1)+1$ 和 $n = n_e(1)+2$ 时，将式（2-125）代入式（2-123），得到

$$A_2 + B_2\rho^{n_e(1)+1} + C_2\sigma^{n_e(1)+1} = B_1\rho^{n_e(1)+1} + C_1\sigma^{n_e(1)+1} + D_1\delta^{n_e(1)+1} \tag{2-126}$$

$$A_2 + B_2\rho^{n_e(1)+2} + C_2\sigma^{n_e(1)+2} = B_1\rho^{n_e(1)+2} + C_1\sigma^{n_e(1)+2} + D_1\delta^{n_e(1)+2} \tag{2-127}$$

解方程组（2-126）和（2-127），得到

$$A_2 = \frac{\rho}{1-\rho}\frac{\theta}{\theta-\xi}\left(1-\frac{\xi(\lambda+\xi-\mu)}{\theta(\lambda+\theta-\mu)}\right)\delta^{n_e(1)+1}p(0,1) \tag{2-128}$$

$$\begin{aligned} B_2 = &\left(\left(1+\frac{\mu}{\theta-\xi}+\frac{\theta\xi}{\lambda(\lambda+\theta-\mu)}+\frac{\xi(\lambda+\xi)}{(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)\right. \\ &\left.-\frac{1}{\rho(1-\rho)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)}\frac{\lambda\theta\xi}{(\lambda+\xi)(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)p(0,1) \end{aligned} \tag{2-129}$$

因此，由式（2-125）得到

$$\begin{aligned} &p(n,0) \\ &=\left(\frac{\rho}{1-\rho}\frac{\theta}{\theta-\xi}\left(1-\frac{\xi(\lambda+\xi-\mu)}{\theta(\lambda+\theta-\mu)}\right)\delta^{n_e(1)+1}+\left(\left(1+\frac{\mu}{\theta-\xi}+\frac{\theta\xi}{\lambda(\lambda+\theta-\mu)}\right.\right.\right. \\ &\left.\left.\left.+\frac{\xi(\lambda+\xi)}{(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)-\frac{1}{\rho(1-\rho)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)}\frac{\lambda\theta\xi}{(\lambda+\xi)(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)\rho^n\right. \\ &\left.+\left(\frac{\xi(\lambda+\xi-\theta)}{(\theta-\xi)(\lambda+\theta-\mu)}-\frac{\xi\lambda}{(\theta-\xi)(\lambda+\theta-\mu)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1}\right)\sigma^{n-1}\right)p(0,1), \\ &n_e(1)+1 \le n \le n_e(2) \end{aligned} \tag{2-130}$$

根据式（2-110）、式（2-111）和式（2-116），有

$$\begin{aligned} &p(n,0) \\ &=\rho^{n-n_e(2)}\left(A_2+B_2\rho^{n_e(2)}+C_2\sigma^{n_e(2)}+\frac{\xi}{\lambda}\left(1+\frac{\sigma\delta\left(\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1}-1\right)}{\delta-\sigma}\right)\sigma^{n_e(2)}p(0,1)\right), \\ &n_e(2)+1 \le n \le n_e(0)+1 \end{aligned} \tag{2-131}$$

由式（2-108）至式（2-109）、式（2-114）至式（2-116）、式（2-122）、式（2-130）至式（2-131），以及平衡方程

$$\sum_{n=1}^{n_e(0)+1} p(n,0) + \sum_{n=0}^{n_e(1)+1} p(n,1) + \sum_{n=0}^{n_e(2)+1} p(n,2) = 1 \tag{2-132}$$

可得到稳态概率 $p(0,1)$的表达式。以下定理给出了系统稳态队长分布。

定理 2.4.2　在具有非中断启动-关闭策略的完全可视排队系统中，如果顾客

均采用如式（2-96）所示的均衡阈值止步策略，且$\sigma \neq 1 \neq \rho \neq \sigma$，$\delta \neq 1 \neq \rho \neq \delta$，则系统稳态队长分布$\{p_{ob}(n,i):(n,i)\in S_{ob}\}$为

$$p_{ob}(n,1)=\sigma^{n}p_{ob}(0,1),\ \ 1\le n\le n_e(1) \tag{2-133}$$

$$p_{ob}(n_e(1)+1,1)=\frac{\lambda}{\xi}\sigma^{n_e(1)}p_{ob}(0,1) \tag{2-134}$$

$$p_{ob}(n,2)=\frac{1}{\rho}\frac{\xi}{\mu}\left(\sigma^{n}+\frac{\sigma\delta(\delta^{n}-\sigma^{n})}{\delta-\sigma}\right)p_{ob}(0,1),1\le n\le n_e(1) \tag{2-135}$$

$$\begin{aligned}p_{ob}(n,2)=&\frac{1}{\rho}\frac{\xi}{\mu}\left(1+\frac{\sigma\delta\left((\frac{\delta}{\sigma})^{n}-1\right)}{\delta-\sigma}\right)\sigma^{n}p_{ob}(0,1),\\&n_e(1)+1\le n\le n_e(2)\end{aligned} \tag{2-136}$$

$$p_{ob}(n_e(2)+1,2)=\frac{1}{\rho(1-\sigma)}\frac{\xi}{\mu}\left(1+\frac{\sigma\delta\left(\left(\frac{\delta}{\sigma}\right)^{n}-1\right)}{\delta-\sigma}\right)\sigma^{n_e(2)}p_{ob}(0,1) \tag{2-137}$$

$$\begin{aligned}p_{ob}(n,0)=&\left(\left(1+\frac{\mu}{\theta-\xi}+\frac{\theta\xi}{\lambda(\lambda+\theta-\mu)}+\frac{\xi(\lambda+\xi)}{(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)\rho^{n}\right.\\&+\frac{\xi(\lambda+\xi-\theta)}{(\theta-\xi)(\lambda+\theta-\mu)}\sigma^{n-1}-\left.\frac{\theta\lambda}{(\theta-\xi)(\lambda+\xi-\mu)\delta^{n-1}}\right)p_{ob}(0,1),\\&1\le n\le n_e(1)\end{aligned} \tag{2-138}$$

$$\begin{aligned}&p_{ob}(n,0)\\=&\left(\frac{\rho}{1-\rho}\frac{\theta}{\theta-\xi}\left(1-\frac{\xi(\lambda+\xi-\mu)}{\theta(\lambda+\theta-\mu)}\right)\delta^{n_e(1)+1}+\left(\left(1+\frac{\mu}{\theta-\xi}+\frac{\theta\xi}{\lambda(\lambda+\theta-\mu)}\right.\right.\right.\\&\left.\left.+\frac{\xi(\lambda+\xi)}{(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)-\frac{1}{\rho(1-\rho)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)}\frac{\lambda\theta\xi}{(\lambda+\xi)(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)\rho^{n}\\&\left.+\left(\frac{\xi(\lambda+\xi-\theta)}{(\theta-\xi)(\lambda+\theta-\mu)}-\frac{\xi\lambda}{(\theta-\xi)(\lambda+\theta-\mu)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1}\right)\sigma^{n-1}\right)p_{ob}(0,1),\\&n_e(1)+1\le n\le n_e(2)\end{aligned} \tag{2-139}$$

$$\begin{aligned}&p_{ob}(n,0)\\=&\left(\rho\left(\left(\frac{\rho}{1-\rho}\frac{\theta}{\theta-\xi}\left(1-\frac{\xi(\lambda+\xi-\mu)}{\theta(\lambda+\theta-\mu)}\right)\delta^{n_e(1)+1}+\rho^{n_e(2)}\left(1+\frac{\mu}{\theta-\xi}+\frac{\theta\xi}{\lambda(\lambda+\theta-\mu)}\right.\right.\right.\right.\\&\left.\left.\left.+\frac{\xi(\lambda+\xi)}{(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)-\frac{1}{\rho(1-\rho)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)}\frac{\lambda\theta\xi}{(\lambda+\xi)(\lambda+\theta-\mu)(\lambda+\xi-\mu)}\right)\end{aligned}$$

$$+\left(\frac{\xi(\lambda+\theta)(\lambda+\xi-\theta)}{\lambda(\theta-\xi)(\lambda+\theta-\mu)}-\frac{\xi(\lambda+\theta)}{(\theta-\xi)(\lambda+\theta-\mu)}\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1}\right)\sigma^{n_e(2)}+\frac{\xi}{\mu}\sigma^{n_e(2)}$$
$$\times\left(1+\frac{\sigma\delta\left(\left(\frac{\delta}{\sigma}\right)^{n_e(1)+1}-1\right)}{\delta-\sigma}\right)\rho^{n-n_e(2)-1}p_{ob}(0,1),\ n_e(2)+1\le n\le n_e(0)+1 \tag{2-140}$$

其中 $p_{ob}(0,1)$可以由平衡方程（2-132）得到。

因为顾客的止步状态为 $(n_e(0)+1,0)$、$(n_e(1)+1,1)$ 和 $(n_e(2)+1,2)$，因此，均衡状态下，单位社会福利为

$$U_{ob}=(R-p)\lambda(1-p_{ob}(n_e(0)+1,0)-p_{ob}(n_e(1)+1,1)-p_{ob}(n_e(2)+1,2))$$
$$-C\left(\sum_{n=1}^{n_e(0)+1}np_{ob}(n,0)+\sum_{n=0}^{n_e(1)+1}np_{ob}(n,1)+\sum_{n=0}^{n_e(2)+1}np_{ob}(n,2)\right) \tag{2-141}$$

2.4.2　完全不可视排队系统

2.4.2.1　纳什均衡

考虑完全不可视情形下顾客的均衡止步行为，且状态转移图如图 2-12 所示。顾客的均衡行为由以下定理给出。

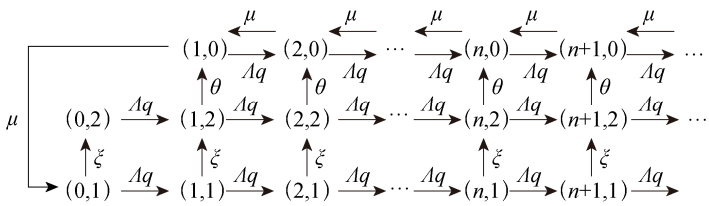

图 2-12　非中断启动-关闭策略和完全不可视情形下的状态转移图

定理 2.4.3　在具有非中断启动-关闭策略的完全不可视 M/M/1 排队系统中，假设 $\Lambda<\mu$，则“以概率 q_e 进入系统”为唯一的均衡混合止步策略且

$$q_e=\begin{cases}\dfrac{1}{\Lambda}x^*, & R\in\left[\dfrac{C}{\mu}+\dfrac{C}{\theta}+p,\ \dfrac{C}{\mu-\Lambda}+\dfrac{C\left(\Lambda(\Lambda+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}\right)}{\Lambda(\Lambda+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}+p\right)\\ 1, & R\in\left[\dfrac{C}{\mu-\Lambda}+\dfrac{C\left(\Lambda(\Lambda+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}\right)}{\Lambda(\Lambda+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}+p,+\infty\right)\end{cases} \tag{2-142}$$

其中 x^*是下列方程

$$R-C\left(\frac{1}{\mu-x}+\frac{x(x+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{x(x+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}\right)-p=0 \tag{2-143}$$

的唯一非负可行解。

证明 标记一个顾客，首先导出此顾客的平均逗留时间。由于顾客到达率为 $\Lambda q(\Lambda<\mu)$，以关闭期内是否有顾客到达为条件，得到忙期开始时系统中的顾客数 Q_b 的分布律为

$$b_j=\sum_{i=1}^{j}c_i u_{j-i}+c^*(\Lambda q)u_{j-1},\quad j\geq 1 \tag{2-144}$$

所以忙期开始时系统中的平均顾客数 $E[Q_b]$为

$$E[Q_b]=\Lambda q\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+c^*(\Lambda q) \tag{2-145}$$

因此，平均附加延迟 $E[W_d]$为

$$E[W_d]=\frac{\frac{\Lambda q}{\xi}}{E[Q_b]}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{\frac{\Lambda q}{\theta}}{E[Q_b]}\frac{1}{\theta}+\frac{c^*(\Lambda q)}{E[Q_b]}\frac{1}{\theta} \tag{2-146}$$

即附加延迟以概率 $\Lambda q/(\xi E[Q_b])$ 为剩余关闭期和一个启动期，以概率 $c^*(\Lambda q)/E[Q_b]$ 为一个启动期，以概率 $\Lambda q/(\theta E[Q_b])$ 为剩余启动期。因此，

$$E[W]=W_{un}(q)=\frac{1}{\mu-\Lambda q}+\frac{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi} \tag{2-147}$$

且平均剩余效用为

$$R-C\left(\frac{1}{\mu-\Lambda q}+\frac{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}\right)-p \tag{2-148}$$

当

$$R\in\left[\frac{C}{\mu}+\frac{C}{\theta}+p,\ \frac{C}{\mu-\Lambda}+\frac{C\left(\Lambda(\Lambda+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}\right)}{\Lambda(\Lambda+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}+p\right)$$

时，式（2-148）在（0, 1）内有唯一解，即为式（2-142）的前半支。否则，当 R 不小于上区间的右端点时，式（2-148）对任一 q 都为正，因此 $q_e=1$，即式（2-142）的后半支。

因此，均衡状态下，单位社会福利为

$$U_{un}=\Lambda q_e\left(R-C\left(\frac{1}{\mu-\Lambda q_e}+\frac{\Lambda q_e(\Lambda q_e+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{\Lambda q_e(\Lambda q_e+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}\right)-p\right) \tag{2-149}$$

2.4.2.2　社会最优

非中断启动-关闭策略下的社会福利函数为

$$SW_{un}(q)=\Lambda q\left(R-C\left(\frac{1}{\mu-\Lambda q}+\frac{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}\right)\right) \tag{2-150}$$

它的一阶导数为

$$\begin{aligned}SW'_{un}(q)=\Lambda&\left(R-C\left(\frac{1}{\mu-\Lambda q}+\frac{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\xi}\right)\right)-\Lambda qC\left(\frac{\Lambda}{(\mu-\Lambda q)^2}\right.\\&+\frac{\Lambda(\Lambda q+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\frac{1}{\theta^2}\right)+\Lambda^2 q\left(\frac{1}{\xi}\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\frac{1}{\theta^2}\right)}{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\xi}\\&\left.-\frac{\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{\left(\Lambda q(\Lambda q+\xi)\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\xi\right)^2}\times\left(\Lambda(\Lambda q+\xi)\left(\frac{1}{\xi}+\frac{1}{\theta}\right)+\Lambda^2 q\left(\frac{1}{\xi}+\frac{1}{\theta}\right)\right)\right)\end{aligned} \tag{2-151}$$

所以社会最优进入概率 q^* 与方程 $SW'_{un}(q)=0$ 的解 q^*_{fd} 存在以下关系。

定理 2.4.4　（1）如果 $0<q_1^*<1$ 和 $SW''_{un}(q_1^*)\le 0$，存在唯一 $q^*=q_1^*$；

（2）如果 $0<q_1^*<1$ 和 $SW''_{un}(q_1^*)>0$ 或者 $q_1^*\ge 1$，存在唯一 $q^*=1$。

同样给出一个充分条件来保证社会福利函数（2-150）单峰且为严格凹函数。

引理 2.4.1　如果 $\rho=\lambda/\mu<1$ 且进入概率 q 的取值范围 $[\underline{q},\overline{q}]$ 非空并满足

$$[\underline{q},\overline{q}]=[0,1]\cap\{q\mid\Lambda^3(\xi+\theta)q^3-3\Lambda\theta\xi^2q-\xi^3\theta\le 0\}$$

则社会福利方程（2-150）在区间 $[\underline{q},\overline{q}]$ 内单峰且为严格凹函数。

证明 如果 $3\Lambda\theta\xi^2+\xi^3\theta\geq\Lambda^3(\xi+\theta)$，只需证明 $W''_{un}(\lambda_q)<0$。因为 $W_{un}(\lambda_q)$ 的一阶和二阶导数分别为

$$W'_{un}(\lambda_q)=\frac{1}{(\mu-\lambda_q)^2}+\frac{(2\lambda_q+\xi)\theta^2\xi}{((\lambda_q+\theta)\xi^2+(\lambda_q^2+\lambda_q\theta)\xi+\lambda_q^2\theta)^2}>0$$

和

$$W''_{un}(\lambda_q)=\frac{2}{(\mu-\lambda_q)^3}-\frac{2\theta^2\xi(\xi^3+3\lambda_q\xi^2+(3\lambda_q^2+3\lambda_q\theta)\xi+3\lambda_q^2\theta)}{((\lambda_q+\theta)\xi^2+(\lambda_q^2+\lambda_q\theta)\xi+\lambda_q^2\theta)^3}$$

虽然 $W'_{un}(\lambda_q)>0$ 恒成立，但 $W''_{un}(\lambda_q)>0$ 成立的充分条件并不能从 $W''_{un}(\lambda_q)$ 的表达式中明显观察到。基于

$$\begin{aligned}SW''_{un}(\lambda_q)&=-2CW'_{un}(\lambda_q)-\lambda_qCW''_{un}(\lambda_q)\\&=-\frac{2C\mu}{(\mu-\lambda_q)^3}-\frac{2C\theta^2\xi(3\lambda_q\theta\xi^2-\lambda_q^3(\xi+\theta)+\xi^3\theta)}{((\lambda_q+\theta)\xi^2+(\lambda_q^2+\lambda_q\theta)\xi+\lambda_q^2\theta)^3}\end{aligned}$$

如果 $\lambda_q^3(\xi+\theta)-3\lambda_q\theta\xi^2-\xi^3\theta\leq 0$，则 $SW''_{un}(\lambda_q)<0$，也就是说，如果 $\Lambda^3(\xi+\theta)q^3-3\Lambda\theta\xi^2q-\xi^3\theta\leq 0$，则 $SW''_{un}(q)<0$。

对于具有非中断启动-关闭策略的完全不可视排队系统来说，图 2-13 给出了单位社会福利关于顾客进入概率的敏感度。观察到最优进入概率分别为 $q^*=0.8945$ 和 $q^*=0.9555$，相应的最优社会福利分别为 $SW^*(q^*)=4.3053$ 和 $SW^*(q^*)=5.4225$。

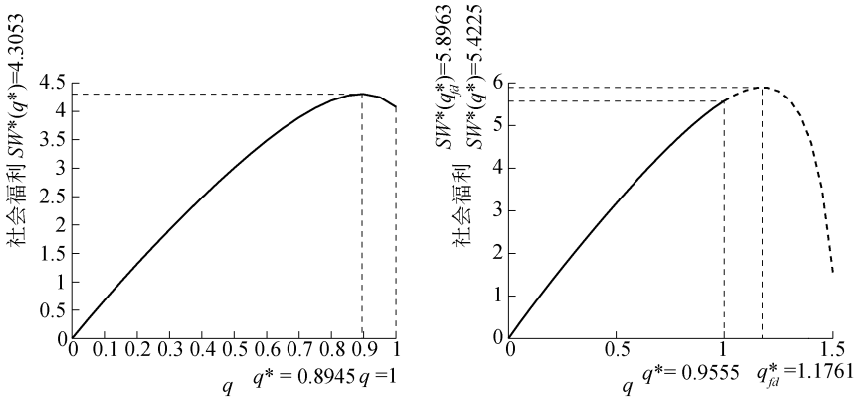

图 2-13 非中断启动-关闭策略下的最优止步策略和最大社会福利

2.4.2.3 利润最大

现考虑垄断情形。对于具有非中断启动-关闭策略的完全不可视排队系统来说，垄断者的目的就是使得 Λqp_m 达到最大，且

$$p_m = p^* = R - C\left(\frac{1}{\mu-\lambda_{fd}^*} + \frac{\lambda_{fd}^*(\lambda_{fd}^*+\xi)\left(\frac{1}{\xi}\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\frac{1}{\theta^2}\right)+\frac{\xi}{\theta}}{\lambda_{fd}^*(\lambda_{fd}^*+\xi)\left(\frac{1}{\theta}+\frac{1}{\xi}\right)+\xi}\right) \tag{2-152}$$

其中 $\lambda_{fd}^* = \Lambda q_{fd}^*$。图 2-14 说明垄断服务费用 p_m 同样是单调非增。

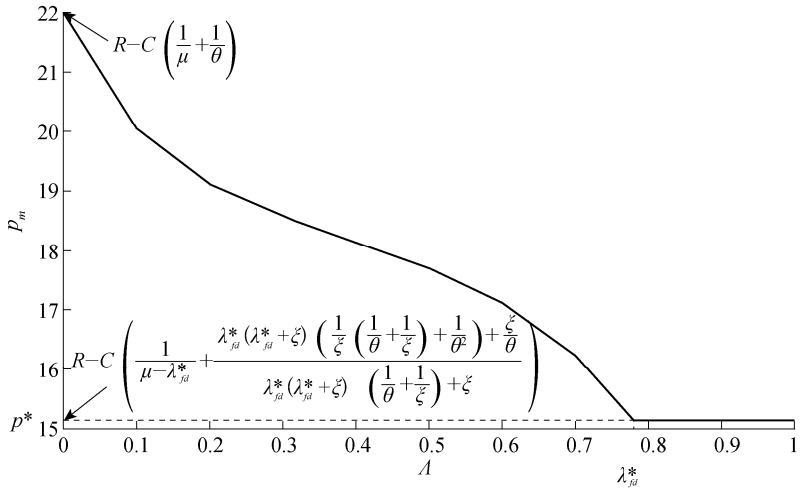

图 2-14　非中断启动-关闭策略下的垄断服务费用

2.5　数值比较

本节将数值比较分别具有三种启动-关闭策略的完全不可视排队系统中顾客的均衡和社会最优止步行为。首先，对于每种策略，对均衡进入概率和社会最优进入概率进行内部比较。然后，针对三种策略，对均衡进入概率和社会最优进入概率进行纵向比较。

对于具有可中断启动-关闭策略的排队系统来说，当实际服务费用分别高于或低于垄断服务费用的情况下，图 2-15 比较了均衡进入概率 q_e 和社会最优进入概率 q^*。观察到，当潜在到达率较小时，q_e 和 q^* 都等于 1。然而，当 Λ 超过某个阈值后两者均递减。对于所有可能的 $\Lambda(<\mu)$，得到结论：当 $p < p^*$ 时，$q_e \geq q^*$，且当 $p > p^*$ 时，$q_e \leq q^*$。

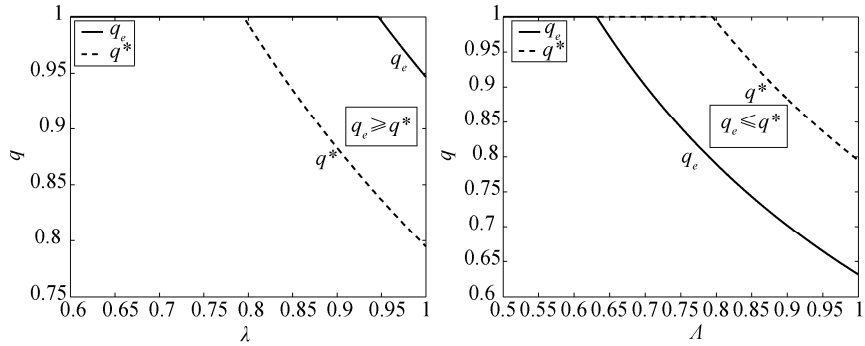

图 2-15　可中断启动-关闭策略下 q_e 和 q^* 的比较

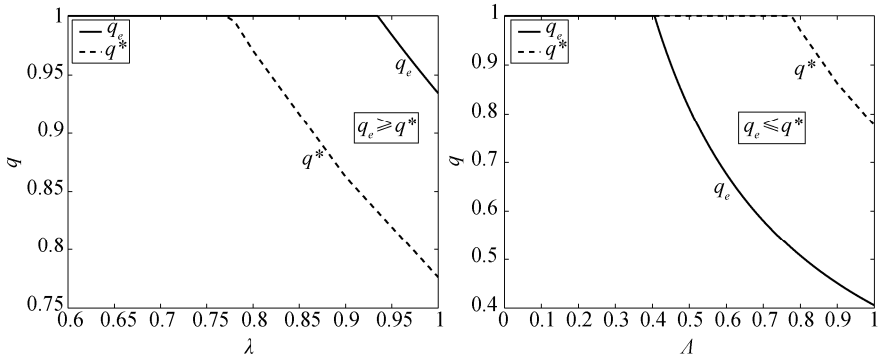

图 2-16　可跳跃启动-关闭策略下 q_e 和 q^* 的比较

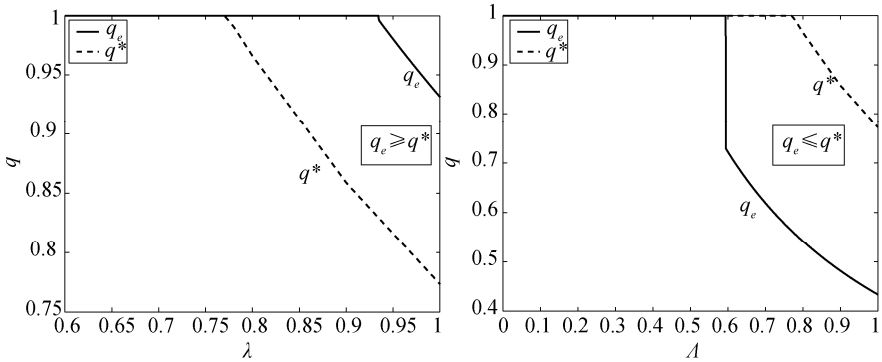

图 2-17　非中断启动-关闭策略下 q_e 和 q^* 的比较

当然，当 $p = p^*$ 时，$q_e = q^*$。平行于图 2-15，图 2-16 和图 2-17 分别比较了具有可跳跃启动-关闭策略和具有非中断启动-关闭策略的排队系统中的均衡进入概率和社会最优进入概率。

图 2-18 和图 2-19 则分别比较了三种启动-关闭策略下的均衡进入概率和社会最优进入概率。容易观察到

$$q_{e\,\text{type 1}} > q_{e\,\text{type 2}} > q_{e\,\text{type 3}} \text{ 且 } q^*_{\text{type 1}} > q^*_{\text{type 2}} > q^*_{\text{type 3}}$$

这是因为顾客不愿意承担由于服务员休假（包括启动期和关闭期）而导致的附加延迟。

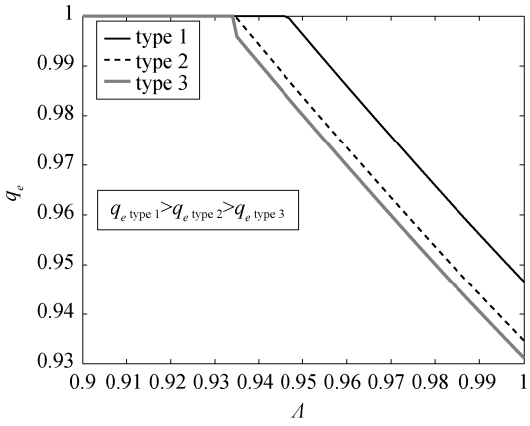

图 2-18　三种启动-关闭策略下 q_e 的比较

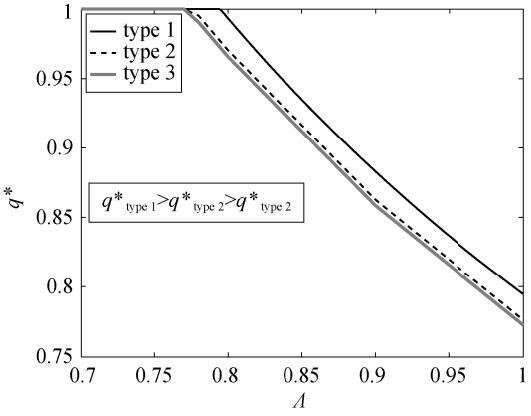

图 2-19　三种启动-关闭策略下 q^* 的比较

接下来，继续数值比较三种启动-关闭策略下单位时间内顾客的均衡剩余效用, 即均衡状态下的单位社会福利, 并观察它们随某些参数的变化规律, 如图 2-20 所示。

图 2-20 表明，无论采用哪一种启动-关闭策略，完全可视情形下单位时间内顾客的均衡剩余效用总是比完全不可视情形下的要大。这说明系统信息量的多少

确实影响顾客的均衡剩余效用。

至于三种启动-关闭策略下顾客均衡剩余效用关于其他一些参数的敏感度，观察到它们均随 R 和 θ 递增。然而，可中断启动-关闭策略下单位时间内顾客的均衡剩余效用随 ξ 递减，这与可跳跃和非中断启动-关闭策略下的递增趋势相反。这说明在可中断启动-关闭策略下，关闭期越长，对顾客越有利。原因在于随着关闭期的缩短，到达顾客发现服务员处于关闭期状态的概率越小，因此平均延迟增加而导致均衡剩余效用降低。对于 λ 来说，单位时间内顾客的均衡剩余效用均可达到最大值。

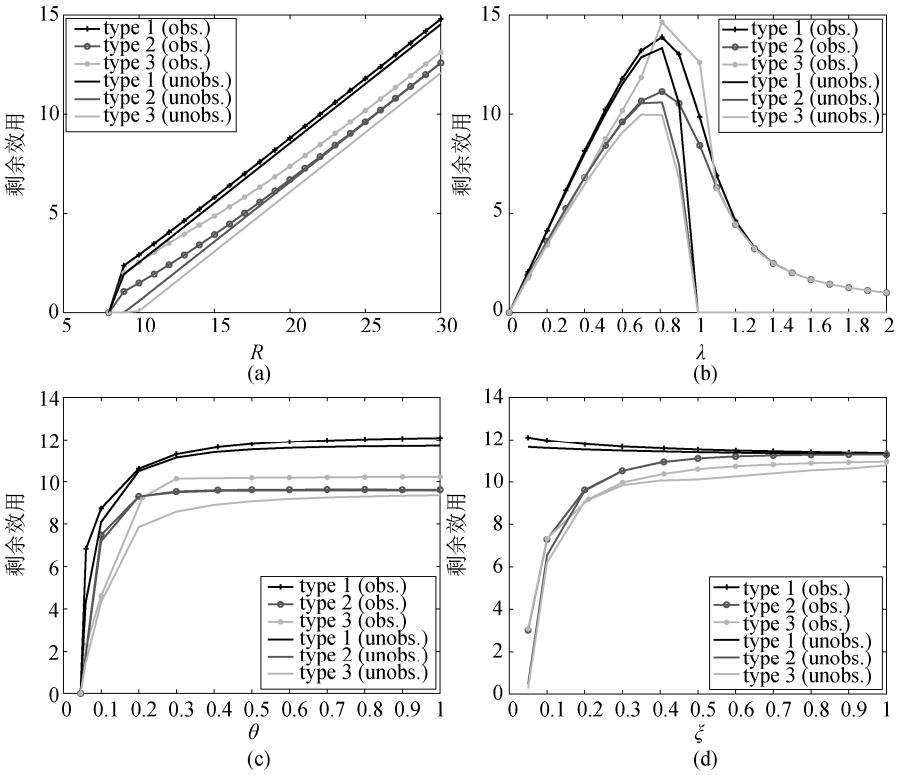

图 2-20　完全可视和完全不可视情形及三种启动-关闭策略下的顾客均衡剩余效用

2.6　相关文献评述

从经典随机服务系统的研究角度，Doshi[45]最早研究了一个具有启动期的 GI/G/1 排队系统。Krishna 和 Reddy 等[46]分析了一个具有多重休假和启动期的批

到达排队系统。随后，Choudhury[47,48]与 Bischof[49]考虑了各类具有启动期的单服务员排队系统，Ke[50]研究了具有各类休假策略与启动期的 M/G/1 排队系统中的优化控制问题。然而，Artalejo 和 Economou 等[51] 则讨论了多类更复杂的具有启动期的多服务员排队系统。

除了启动期之外，Takagi[7]又最早提出了关闭期的概念，并研究了一个具有启动关闭策略的 M/G/1 排队系统。Arumuganathan 和 Jeyakumar[52]分析了一个具有多重休假和关闭期的 M^X/G(a,b)/1 排队系统。随后，他们[53]又考虑了一个同类型的排队系统，并将 *N* 策略与启动期加入模型假设之中。Ke[54,55]则研究了一个具有服务员失效与启动关闭策略的 M/G/1 排队系统。

以上文献均是从经典随机服务系统的研究角度去导出系统的主要性能指标，而有关顾客均衡止步策略的研究工作，最早要追溯到 Burnetas 和 Economou[35]于 2007 年发表的一篇文章。他们在文中研究了具有启动策略的 Markov 单服务员排队系统中顾客的均衡止步策略。根据顾客已知的系统信息量的多少，分别讨论了完全可视、几乎可视、几乎不可视与完全不可视四类信息量水平。而后，孙微等[43,44]则讨论了具有三类启动关闭策略的排队系统：可中断启动-关闭策略、可跳跃启动-关闭策略和非中断启动-关闭策略，研究了顾客的均衡和最优止步策略及服务员的利润最大化策略。

第 3 章　具有多重休假策略的排队系统

3.1　多重休假策略

多重休假策略是经典休假策略的一种常见形式。休假规则为：一旦排队系统空竭，服务员则进入休假状态，休假期间不对顾客提供服务。当一次休假结束后，若服务员发现系统中没有顾客等待，将立即开始下一次休假，直到某次休假结束时，系统中至少有一位顾客等待服务为止。否则，服务员立即开始一个新的忙期。

3.2　具有不同决策准则的排队系统

本节将研究具有多重休假策略的队长不可视 Markov 排队系统，导出并比较具有五类决策标准［最大最小（maximin）准则、最大最大（maximax）准则、Hurwicz 准则、最小最大后悔值（minimax regret）准则和 Laplace 准则］与三类等待费用函数（线性、二次与指数形式）的同类顾客的均衡阈值/混合止步策略。

本节内容为作者研究成果[56]的进一步完善和补充。

3.2.1　模型描述

我们假设同类顾客到达过程服从 Poisson 分布，顾客潜在到达率为 Λ，服务员服务率服从参数为 μ 的指数分布。假定队长不可视，且到达顾客只知道服务时间服从指数分布，但参数 μ 的具体取值并不被告知。由于队长不可视且顾客类型相同，他们将采取同样的均衡策略。将均衡状态下顾客的进入概率记为 $q_e(0 \leq q_e \leq 1)$，因此，有效到达率为 $\lambda_e = \Lambda q_e$。

一旦系统空竭，服务员将采取多重休假策略。假定随机长度的单次休假期 V 服从一般分布，记 V 的前三阶原点矩分别为 $E[V] = v$, $E[V^2] = v^{(2)}$和 $E[V^3] = v^{(3)}$。

因此，服务员的状态可分为

$$\text{服务员状态}=\begin{cases}\mathbf{0}, & \text{服务员处于休假期}\\ \mathbf{1}, & \text{服务员处于忙期}\end{cases}$$

就像队长和服务时间参数取值一样，假定服务员所处状态信息也对到达顾客是封闭的。

单个顾客服务完成后的服务收益为 R，等待 w 个时间单位（包括服务时间）后的等待损耗为 $c(w)$。因此，单个顾客的平均剩余效用记为 U，为 $U=R-c(w)$。函数 $c(\cdot)$的不同形式代表着顾客对潜在风险的不同的敏感度。假定所有顾客具有相同的等待损耗函数和延迟敏感参数，且服务顺序为先到先服务。

3.2.1.1　两者取一

首先假定有关服务时间参数 μ 的部分信息为 μ_1 和 μ_2 两者中的某一个，且 $0<\mu_1<\mu_2$。将该信息和服务员可能的两个状态相结合，我们可以确定存在四类可能的系统状态，分别为

$$\text{系统状态}=\begin{cases}(\mu_1,\mathbf{1}), & \text{服务员处于忙期且}\mu=\mu_1\\ (\mu_1,\mathbf{0}), & \text{服务员处于假期且}\mu=\mu_1\\ (\mu_2,\mathbf{1}), & \text{服务员处于忙期且}\mu=\mu_2\\ (\mu_2,\mathbf{0}), & \text{服务员处于假期且}\mu=\mu_2\end{cases}$$

3.2.1.2　取值范围

然后假定有关服务时间参数 μ 的部分信息为 μ 处于闭区间$[\underline{\mu},\overline{\mu}]$内，且 $0<\underline{\mu}<\overline{\mu}$。在此情形下，将存在无穷多个可能的系统状态，即为

$$\text{系统状态}=\begin{cases}(\underline{\mu},\mathbf{1}), & \text{服务员处于忙期且}\mu=\mu_1\\ (\underline{\mu},\mathbf{0}), & \text{服务员处于假期且}\mu=\mu_1\\ (\forall\mu\in(\underline{\mu},\overline{\mu}),\mathbf{1}), & \text{服务员处于忙期且}\mu\in(\underline{\mu},\overline{\mu})\\ (\forall\mu\in(\underline{\mu},\overline{\mu}),\mathbf{0}), & \text{服务员处于假期且}\mu\in(\underline{\mu},\overline{\mu})\\ (\overline{\mu},\mathbf{1}), & \text{服务员处于忙期且}\mu=\mu_2\\ (\overline{\mu},\mathbf{0}), & \text{服务员处于假期且}\mu=\mu_2\end{cases}$$

以下我们将看到具有不同决策准则的顾客会得到完全不同的系统状态的估计结果。

3.2.2　线性等待损耗

首先，假定所有顾客的等待损耗均为等待时间的线性函数，即 $c(w)=cw$，其

中 c 为常数且 $c>0$。显然，这类顾客为风险中立顾客。

3.2.2.1 两者取一

为了得到各种决策准则下顾客的均衡混合策略，我们应该首先找到顾客在各个系统状态下的平均剩余效用。

如果系统状态$(\mu_1, \mathbf{1})$最终实现，顾客完成服务之后的平均剩余效用为

$$U_{\text{join}}^{(\mu_1,1)} = R - \frac{c}{\mu_1 - \lambda}$$ ①

其中 λ 是有效到达率。

如果系统状态$(\mu_1, \mathbf{0})$最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_1,0)} = R - \frac{c}{\mu_1 - \lambda} - \frac{cv^{(2)}}{2v}$$

其中 $cv^{(2)}/(2v)$为顾客到达时的剩余休假期。

如果系统状态$(\mu_2, \mathbf{1})$最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_2,1)} = R - \frac{c}{\mu_2 - \lambda}$$

如果系统状态$(\mu_2, \mathbf{0})$最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_2,0)} = R - \frac{c}{\mu_2 - \lambda} - \frac{cv^{(2)}}{2v}$$

接下来，我们开始分别分析在五类决策准则下，风险中立顾客的均衡混合策略。

1. 最大最小准则

最大最小准则又称为悲观准则，到达顾客总是假定系统状态为最不利的$(\mu_1,\mathbf{0})$。因此，只有当$U_{\text{join}}^{(\mu_1,0)} > 0$时，该类顾客才会进入，否则离开。求解方程

$$U_{\text{join}}^{(\mu_1,0)} = R - \frac{c}{\mu_1 - \lambda} - \frac{cv^{(2)}}{2v} = 0$$

得到如下命题中的结果。

命题 3.2.1 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且顾客的决策准则为最大最小准则，则具有线性等待损耗 $c(w) = cw$ 的顾客的均衡有效到达率为

① 顾客等待时间的一阶矩和二阶矩，以及具有指数损耗函数的顾客平均等待时间的推导过程请见参考文献[56]。

$$\lambda_{e(cw)}^{\mathrm{pes}_a}=\begin{cases}\Lambda, & \Lambda\le\mu_1-\dfrac{c}{R-\dfrac{cv^{(2)}}{2v}}\\ \mu_1-\dfrac{c}{R-\dfrac{cv^{(2)}}{2v}}, & \Lambda>\mu_1-\dfrac{c}{R-\dfrac{cv^{(2)}}{2v}}\end{cases}\tag{3-1}$$

以及相应的均衡进入概率为 $q_{e(cw)}^{\mathrm{pes}_a}=\lambda_{e(cw)}^{\mathrm{pes}_a}/\Lambda$。

2. 最大最大准则

和悲观准则相反，最大最大准则又称为乐观准则，即到达顾客总是假定系统状态为最有利的$(\mu_2, \mathbf{1})$。求解方程

$$U_{\mathrm{join}}^{(\mu_2,1)}=R-\frac{c}{\mu_2-\lambda}=0$$

得到如下命题中的结果。

命题 3.2.2　如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且顾客的决策准则为最大最大准则，则具有线性等待损耗 $c(w)=cw$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw)}^{\mathrm{opt}_a}=\begin{cases}\Lambda, & \Lambda\le\mu_2-\dfrac{c}{R}\\ \mu_2-\dfrac{c}{R}, & \Lambda>\mu_2-\dfrac{c}{R}\end{cases}\tag{3-2}$$

以及相应的均衡进入概率为 $q_{e(cw)}^{\mathrm{opt}_a}=\lambda_{e(cw)}^{\mathrm{opt}_a}/\Lambda$。

3. Hurwicz 准则

定理 3.2.1　如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，以下方程

$$R-(1-\alpha)\frac{cv^{(2)}}{2v}-\left(\alpha\frac{c}{\mu_2-\lambda}+(1-\alpha)\frac{c}{\mu_1-\lambda}\right)=0\tag{3-3}$$

的解为 λ^*，且顾客的决策准则为 Hurwicz 准则，则具有线性等待损耗 $c(w)=cw$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw)}^{\mathrm{Hur}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases}\tag{3-4}$$

以及相应的均衡进入概率为 $q_{e(cw)}^{\mathrm{Hur}_a}=\lambda_{e(cw)}^{\mathrm{Hur}_a}/\Lambda$。特别地，如果没有休假策略，则有

$$\lambda_{e(cw-wv)}^{Hur_a}=\frac{1}{2}\left(\mu_1+\mu_2-\frac{c}{R}\right)-\frac{1}{2}\sqrt{\left(\mu_1+\mu_2-\frac{c}{R}\right)^2+\alpha\frac{4c\mu_1}{R}+(1-\alpha)\frac{4c\mu_2}{R}-4\mu_1\mu_2}$$

证明 根据以下的顾客效用表格，可得以下结论。

效用	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	$R-\dfrac{c}{\mu_1-\Lambda}$	$R-\dfrac{c}{\mu_1-\Lambda}-\dfrac{cv^{(2)}}{2v}$	$R-\dfrac{c}{\mu_2-\Lambda}$	$R-\dfrac{c}{\mu_2-\Lambda}-\dfrac{cv^{(2)}}{2v}$
止步	0	0	0	0

其中假定所有的潜在顾客均选择进入系统，则进入顾客的联合效用为

$$U_{\text{join}}=\alpha\left(R-\frac{c}{\mu_2-\Lambda}\right)+(1-\alpha)\left(R-\frac{c}{\mu_1-\Lambda}-\frac{cv^{(2)}}{2v}\right)$$

而显然，离开顾客的联合效用为 $U_{balk}=0$。因此，最大的联合效用为

$$U^{\text{Hur}}=\max\left\{\alpha\left(R-\frac{c}{\mu_2-\Lambda}\right)+(1-\alpha)\left(R-\frac{c}{\mu_1-\Lambda}-\frac{cv^{(2)}}{2v}\right),0\right\}$$

如果 $U_{\text{join}}>0$，所有顾客选择进入，否则，他们将选择混合策略。因此，通过求解 $U_{\text{join}}=0$ 得到顾客的均衡有效到达率为 λ^*。根据 λ^* 与 Λ 间的大小关系，得到 $\lambda_{e(cw)}^{\text{Hur}_a}$。

4. 最小最大后悔值准则

定理 3.2.2 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，以下方程

$$2R-\frac{cv^{(2)}}{2v}-\left(\frac{c}{\mu_1-\lambda}+\frac{c}{\mu_2-\lambda}\right)=0 \tag{3-5}$$

的解为 λ^*，且顾客的决策准则为最小最大后悔值准则，则具有线性等待损耗 $c(w)=cw$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw)}^{\text{reg}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-6}$$

以及相应的均衡进入概率为 $q_{e(cw)}^{\text{reg}_a}=\lambda_{e(cw)}^{\text{reg}_a}/\Lambda$。特别地，如果没有休假策略，则有

$$\lambda_{e(cw-wv)}^{\text{reg}_a}=\frac{1}{2}\left(\mu_1+\mu_2-\frac{c}{R}\right)-\frac{1}{2}\sqrt{\left(\mu_1+\mu_2-\frac{c}{R}\right)^2+2(\mu_1+\mu_2)\frac{c}{R}-4\mu_1\mu_2}$$

证明 我们将分以下三种情况证明该定理。

（1）如果 $\Lambda\le\mu_1-c/(R-cv^{(2)}/2v)$，有如下的后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	0	0	0	0
止步	$R-\dfrac{c}{\mu_1-\Lambda}$	$R-\dfrac{c}{\mu_1-\Lambda}-\dfrac{cv^{(2)}}{2v}$	$R-\dfrac{c}{\mu_2-\Lambda}$	$R-\dfrac{c}{\mu_2-\Lambda}-\dfrac{cv^{(2)}}{2v}$

因此，进入系统的最大后悔值为 $\rho_{\text{join}}=0$，而离开系统的最大后悔值为 $\rho_{\text{balk}}=R-c/(\mu_2-\Lambda)$。这样，最小最大后悔值为 $\rho^*=\min\{0, R-c/(\mu_2-\Lambda)\}=0$，所有顾客选择进入，即 $q_{e(cw)}^{\text{reg}_a}=1$。

（2）如果 $\mu_1-c/(R-cv^{(2)}/2v)<\Lambda\leq\mu_2-c/R$，有如下的后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	0 或 $\dfrac{c}{\mu_1-\Lambda}-R$	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\mu_1-\Lambda}-R$	0	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\mu_2-\Lambda}-R$ 或 0
止步	$R-\dfrac{c}{\mu_1-\Lambda}$或0	0	$R-\dfrac{c}{\mu_2-\Lambda}$	0或$R-\dfrac{c}{\mu_2-\Lambda}-\dfrac{cv^{(2)}}{2v}$

因此，$\rho_{\text{join}}=cv^{(2)}/(2v)+c/(\mu_1-\Lambda)-R$ 且 $\rho_{\text{balk}}=R-c/(\mu_2-\Lambda)$，最小最大后悔值为 $\rho^*=\min\{cv^{(2)}/(2v)+c/(\mu_1-\Lambda)-R,\ R-c/(\mu_2-\Lambda)\}$。令

$$\frac{cv^{(2)}}{2v}+\frac{c}{\mu_1-\Lambda}-R=R-\frac{c}{\mu_2-\Lambda}$$

用其他参数来表示 Λ，并用 λ^*表示该表达式。如果 $\mu_1-c/(R-cv^{(2)}/2v)<\Lambda<\lambda^*$，所有顾客选择进入，即 $q_{e(cw)}^{\text{reg}_a}=1$；如果 $\lambda^*\leq\Lambda\leq\mu_2-c/R$，所有顾客将以概率 $q_{e(cw)}^{\text{reg}_a}=\lambda^*/\Lambda$ 选择进入。

（3）如果 $\Lambda>\mu_2-c/R$，有如下的后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	$\dfrac{c}{\mu_1-\Lambda}-R$	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\mu_1-\Lambda}-R$	$\dfrac{c}{\mu_2-\Lambda}-R$	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\mu_2-\Lambda}-R$
止步	0	0	0	0

因此，$\rho_{\text{join}}=cv^{(2)}/(2v)+c/(\mu_1-\Lambda)-R$ 和 $\rho_{\text{balk}}=0$，最小最大后悔值为 $\rho^*=\min\{cv^{(2)}/(2v)+c/(\mu_1-\Lambda)-R,0\}=0$，说明所有顾客需以概率 $q_{e(cw)}^{\text{reg}_a}=\lambda^*/\Lambda$ 选择混合策略而非选择进入系统这种纯策略。

结合以上讨论，得到定理中的结论。

5. Laplace 准则

定理 3.2.3　如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且顾客的决策准则为 Laplace 准则，则具有线性等待损耗 $c(w)=cw$ 的顾客的均衡有效到达率和具有最小最大后悔值准则的顾客的均衡有效到达率一致，即 $\lambda_{e(cw)}^{\text{Lap}_a}=\lambda_{e(cw)}^{\text{reg}_a}$。

证明　由于存在四种可能的系统状态，则顾客进入系统的平均剩余效用为

$$
\begin{aligned}
U_{\text{join}} &= \frac{1}{4}\left(R-\frac{c}{\mu_1-\lambda}\right)+\frac{1}{4}\left(R-\frac{c}{\mu_1-\lambda}-\frac{cv^{(2)}}{2v}\right) \\
&\quad +\frac{1}{4}\left(R-\frac{c}{\mu_2-\lambda}\right)+\frac{1}{4}\left(R-\frac{c}{\mu_2-\lambda}-\frac{cv^{(2)}}{2v}\right) \\
&= \frac{1}{2}\left(R-\frac{c}{\mu_1-\lambda}\right)+\frac{1}{2}\left(R-\frac{c}{\mu_2-\lambda}\right)-\frac{1}{2}\frac{cv^{(2)}}{2v}
\end{aligned}
$$

而离开系统的平均剩余效用为 0。因此，最大的平均剩余效用为

$$
U^{\text{Lap}} = \max\left\{\frac{1}{2}\left(R-\frac{c}{\mu_1-\lambda}\right)+\frac{1}{2}\left(R-\frac{c}{\mu_2-\lambda}\right)-\frac{1}{2}\frac{cv^{(2)}}{2v},0\right\}
$$

通过求解 $U_{\text{join}}=0$，得到 $\lambda_{e(cw)}^{\text{Lap}_a}$，且观察到 $\lambda_{e(cw)}^{\text{Lap}_a}=\lambda_{e(cw)}^{\text{reg}_a}$。

假定 Λ 足够大，图 3-1[①]比较了各种决策准则下顾客的均衡有效到达率。我们观察到，$\lambda_e^{\text{opt}} \geq \lambda_e^{\text{Hur}} \geq \lambda_e^{\text{reg}} = \lambda_e^{\text{Lap}} \geq \lambda_e^{\text{pes}}$，其直观地反映出具有不同决策准则的顾客对系统风险的处置态度。图 3-2 比较了顾客的均衡进入概率，类似于图 3-1，$q_e^{\text{opt}} \geq q_e^{\text{Hur}} \geq q_e^{\text{reg}} = q_e^{\text{Lap}} \geq q_e^{\text{pes}}$。

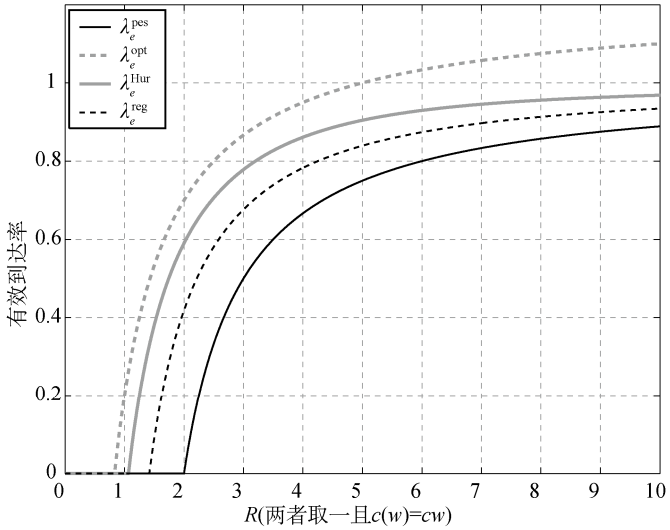

图 3-1　两者取一信息下顾客的均衡有效到达率[$c(w)=cw$]

① 为了使图更加清晰，我们在图例中简化了参数的上下标，并在横轴位置对部分信息类型和等待损耗函数形式进行了说明。

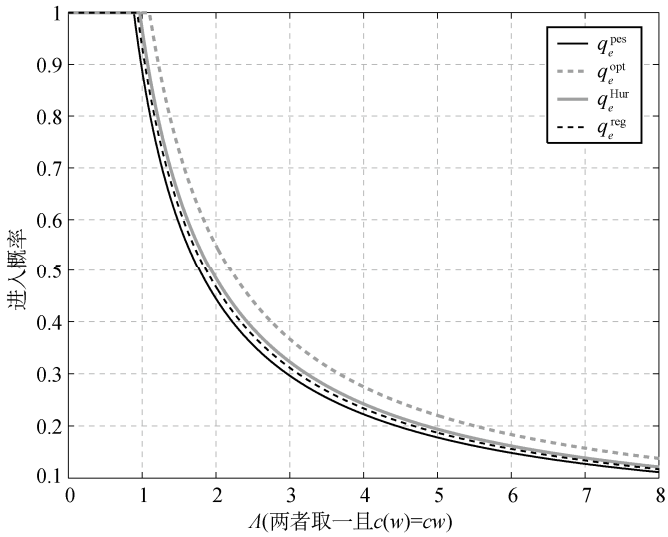

图 3-2　两者取一信息下顾客的均衡进入概率[$c(w) = cw$]

3.2.2.2　取值范围

我们继续考虑关于服务时间参数的第二类部分信息, 即 $\mu \in \left[\overline{\mu},\ \underline{\mu}\right]$, $0 < \underline{\mu} < \overline{\mu}$ ①。对于具有最大最小准则的顾客，他们总是假设系统状态为 $(\underline{\mu}, \mathbf{0})$。因此，得到顾客的均衡有效到达率为

$$\lambda_{e(cw)}^{\mathrm{pes}_r} = \underline{\mu} - \frac{c}{R - \dfrac{cv^{(2)}}{2v}} \tag{3-8}$$

对于具有最大最大准则的顾客，他们总是假设系统状态为$(\overline{\mu},\ \mathbf{1})$。因此，得到顾客的均衡有效到达率为

$$\lambda_{e(cw)}^{\mathrm{opt}_r} = \overline{\mu} - \frac{c}{R} \tag{3-9}$$

对于具有 Hurwicz 准则的顾客，他们假设系统状态以概率 α 为$(\overline{\mu},\ \mathbf{1})$，而以概率 $1-\alpha$ 为$(\underline{\mu}, \mathbf{0})$。因此，通过求解

$$R - (1-\alpha)\frac{cv^{(2)}}{2v} - \left(\alpha\frac{c}{\overline{\mu}-\lambda} + (1-\alpha)\frac{c}{\underline{\mu}-\lambda}\right) = 0 \tag{3-10}$$

可得到顾客的均衡有效到达率为 $\lambda_{e(cw)}^{\mathrm{Hur}_r}$。

在此，我们重点分析在最小最大后悔值准则和 Laplace 准则下顾客的均衡有效到达率。

① 在此省略顾客在各种系统状态下的平均剩余效用。

定理 3.2.4 如果服务时间参数 μ 处于区间 $[\underline{\mu},\overline{\mu}](0<\underline{\mu}<\overline{\mu})$ 内，记以下方程

$$2R-\frac{cv^{(2)}}{2v}-\left(\frac{c}{\underline{\mu}-\lambda}+\frac{c}{\overline{\mu}-\lambda}\right)=0 \tag{3-11}$$

的解为 λ^*，且顾客的决策准则为最小最大后悔值准则，则具有线性等待损耗 $c(w)=cw$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw)}^{\mathrm{reg}_r}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-12}$$

以及相应的均衡进入概率为 $q_{e(cw)}^{\mathrm{reg}_r}=\lambda_{e(cw)}^{\mathrm{reg}_r}/\Lambda$。

证明 （1）如果 $\Lambda\le\underline{\mu}-c/(R-cv^{(2)}/2v)$，有如下的后悔值表格：

后悔值	$(\underline{\mu},1)$	$(\underline{\mu},0)$	$(\forall\mu\in(\underline{\mu},\ \overline{\mu}),1)$
进入	0	0	0
止步	$R-\dfrac{c}{\underline{\mu}-\Lambda}$	$R-\dfrac{c}{\underline{\mu}-\Lambda}-\dfrac{cv^{(2)}}{2v}$	$R-\dfrac{c}{\mu-\Lambda}$

后悔值	$(\forall\mu\in(\underline{\mu},\ \overline{\mu}),0)$	$(\overline{\mu},1)$	$(\overline{\mu},0)$
进入	0	0	0
止步	$R-\dfrac{c}{\mu-\Lambda}-\dfrac{cv^{(2)}}{2v}$	$R-\dfrac{c}{\overline{\mu}-\Lambda}$	$R-\dfrac{c}{\overline{\mu}-\Lambda}-\dfrac{cv^{(2)}}{2v}$

因此，$\rho_{\mathrm{join}}=0$ 且 $\rho_{\mathrm{balk}}=R-c/(\overline{\mu}-\Lambda)$，最小最大后悔值为 $\rho^*=\min\{0,R-c/(\overline{\mu}-\Lambda)\}=0$，所有顾客将选择进入系统，即 $q_{e(cw)}^{\mathrm{reg}_r}=1$。

（2）如果 $\underline{\mu}-c/(R-cv^{(2)}/2v)<\Lambda\le\overline{\mu}-c/R$，有如下的后悔值表格：

后悔值	$(\underline{\mu},1)$	$(\underline{\mu},0)$	$(\forall\mu\in(\underline{\mu},\ \overline{\mu}),1)$
进入	0 或 $\dfrac{c}{\underline{\mu}-\Lambda}-R$	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\underline{\mu}-\Lambda}-R$	0 或 $\dfrac{c}{\mu-\Lambda}-R$
止步	$R-\dfrac{c}{\underline{\mu}-\Lambda}$ 或 0	0	$R-\dfrac{c}{\mu-\Lambda}$ 或 0

后悔值	$(\forall\mu\in(\underline{\mu},\ \overline{\mu}),0)$	$(\overline{\mu},1)$	$(\overline{\mu},0)$
进入	0 或 $\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\mu-\Lambda}-R$	0	0 或 $\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\overline{\mu}-\Lambda}-R$
止步	$R-\dfrac{c}{\mu-\Lambda}-\dfrac{cv^{(2)}}{2v}$ 或 0	$R-\dfrac{c}{\overline{\mu}-\Lambda}$	$R-\dfrac{c}{\overline{\mu}-\Lambda}-\dfrac{cv^{(2)}}{2v}$ 或 0

因此，$\rho_{\text{join}} = cv^{(2)}/(2v)+c/(\underline{\mu}-\Lambda) - R$ 且 $\rho_{\text{balk}} = R - c/(\overline{\mu}-\Lambda)$，最小最大后悔值为 $\rho^* = \min\{cv^{(2)}/(2v) + c/(\underline{\mu}-\Lambda) - R, R - c/(\overline{\mu}-\Lambda)\}$。令

$$\frac{cv^{(2)}}{2v}+\frac{c}{\underline{\mu}-\Lambda}-R = R-\frac{c}{\overline{\mu}-\Lambda}$$

用其他参数来表示 Λ，并用 λ^*表示该表达式。如果 $\underline{\mu}-c/(R-cv^{(2)}/2v) < \Lambda < \lambda^*$，所有顾客选择进入，即 $q_{e(cw)}^{\text{reg}_r}=1$；如果 $\lambda^* \le \Lambda \le \overline{\mu}-c/R$，所有顾客将以概率 $q_{e(cw)}^{\text{reg}_r}=\lambda^*/\Lambda$ 选择进入系统。

（3）如果 $\Lambda > \overline{\mu}-c/R$，有如下的后悔值表格：

后悔值	$(\underline{\mu},1)$	$(\underline{\mu},0)$	$(\forall\mu\in(\underline{\mu},\ \overline{\mu}),1)$
进入	$\dfrac{c}{\underline{\mu}-\Lambda}-R$	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\underline{\mu}-\Lambda}-R$	$\dfrac{c}{\mu-\Lambda}-R$
止步	0	0	0

后悔值	$(\forall\mu\in(\underline{\mu},\ \overline{\mu}),0)$	$(\overline{\mu},1)$	$(\overline{\mu},0)$
进入	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\mu-\Lambda}-R$	$\dfrac{c}{\overline{\mu}-\Lambda}-R$	$\dfrac{cv^{(2)}}{2v}+\dfrac{c}{\overline{\mu}-\Lambda}-R$
止步	0	0	0

因此，$\rho_{\text{join}} = cv^{(2)}/(2v)+c/(\underline{\mu}-\Lambda)-R$ 且 $\rho_{\text{balk}}=0$，最小最大后悔值为 $\rho^* = \min\{cv^{(2)}/(2v) +c/(\underline{\mu}-\Lambda)-R, 0\}=0$，说明所有顾客将以概率 $q_{e(cw)}^{\text{reg}_r}=\lambda^*/\Lambda$ 选择混合策略。

结合以上讨论，得到定理中的结论。

定理 3.2.5　如果服务时间参数 μ 处于区间 $[\underline{\mu},\overline{\mu}]$ $(0<\underline{\mu}<\overline{\mu})$ 内，且顾客的决策准则为 Laplace 准则，则具有线性等待损耗 $c(w) = cw$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw)}^{\text{Lap}_r}=\begin{cases}\Lambda,\ \Lambda \le \dfrac{\underline{\mu}\mathrm{e}^{(\overline{\mu}-\underline{\mu})\left(\frac{R}{c}-\frac{v^{(2)}}{4v}\right)}-\overline{\mu}}{\mathrm{e}^{(\overline{\mu}-\underline{\mu})\left(\frac{R}{c}-\frac{v^{(2)}}{4v}\right)}-1} \\ \dfrac{\underline{\mu}\mathrm{e}^{(\overline{\mu}-\underline{\mu})\left(\frac{R}{c}-\frac{v^{(2)}}{4v}\right)}-\overline{\mu}}{\mathrm{e}^{(\overline{\mu}-\underline{\mu})\left(\frac{R}{c}-\frac{v^{(2)}}{4v}\right)}-1},\ \Lambda > \dfrac{\underline{\mu}\mathrm{e}^{(\overline{\mu}-\underline{\mu})\left(\frac{R}{c}-\frac{v^{(2)}}{4v}\right)}-\overline{\mu}}{\mathrm{e}^{(\overline{\mu}-\underline{\mu})\left(\frac{R}{c}-\frac{v^{(2)}}{4v}\right)}-1}\end{cases} \tag{3-13}$$

以及相应的均衡进入概率为 $q_{e(cw)}^{\mathrm{Lap}_r}=\lambda_{e(cw)}^{\mathrm{Lap}_r}/\Lambda$。

证明 由于服务员以等概率处于休假或忙期状态，且 μ 在区间$[\underline{\mu},\overline{\mu}]$上均匀分布，则顾客进入系统的平均剩余效用为

$$\begin{aligned}U_{\mathrm{join}}&=\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}}\left(R-\frac{c}{\mu-\lambda}\right)\frac{1}{\overline{\mu}-\underline{\mu}}\mathrm{d}\mu+\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}}\left(R-\frac{c}{\mu-\lambda}-\frac{cv^{(2)}}{2v}\right)\frac{1}{\overline{\mu}-\underline{\mu}}\mathrm{d}u\\&=\frac{1}{2\left(\overline{\mu}-\underline{\mu}\right)}\left(R\mu-c\ln(\mu-\lambda)\right)\Big|_{\underline{\mu}}^{\overline{\mu}}+\frac{1}{2(\overline{\mu}-\underline{\mu})}\left(R\mu-\frac{cv^{(2)}}{2v}\mu-c\ln(\mu-\lambda)\right)\Big|_{\underline{\mu}}^{\overline{\mu}}\\&=R-\frac{1}{2}\frac{cv^{(2)}}{2v}-\frac{c}{\overline{\mu}-\underline{\mu}}\ln\left(\frac{\overline{\mu}-\lambda}{\underline{\mu}-\lambda}\right)\end{aligned}$$

而顾客离开的平均效用为 $U_{\mathrm{balk}}=0$。因此，最大的平均效用为

$$U^{\mathrm{Lap}}=\max\left\{R-\frac{1}{2}\frac{cv^{(2)}}{2v}-\frac{c}{\overline{\mu}-\underline{\mu}}\ln\left(\frac{\overline{\mu}-\lambda}{\underline{\mu}-\lambda}\right),0\right\}$$

求解 $U_{\mathrm{join}}=0$，得到 $\lambda_{e(cw)}^{\mathrm{Lap}_r}$。

平行于图 3-1，图 3-3 中比较了此种情况下各种决策准则下顾客的均衡有效到达率。我们观察到，$\lambda_e^{\mathrm{opt}}\geq\lambda_e^{\mathrm{Lap}}\geq\lambda_e^{\mathrm{reg}}\geq\lambda_e^{\mathrm{Hur}}\geq\lambda_e^{\mathrm{pes}}$，说明最小最大后悔值准则是一个最理性和人性化的决策准则，也是顾客最常采用的决策准则。图 3-4 比较了顾客的均衡进入概率，类似于图 3-3，$q_e^{\mathrm{opt}}\geq q_e^{\mathrm{Lap}}\geq q_e^{\mathrm{reg}}\geq q_e^{\mathrm{Hur}}\geq q_e^{\mathrm{pes}}$。

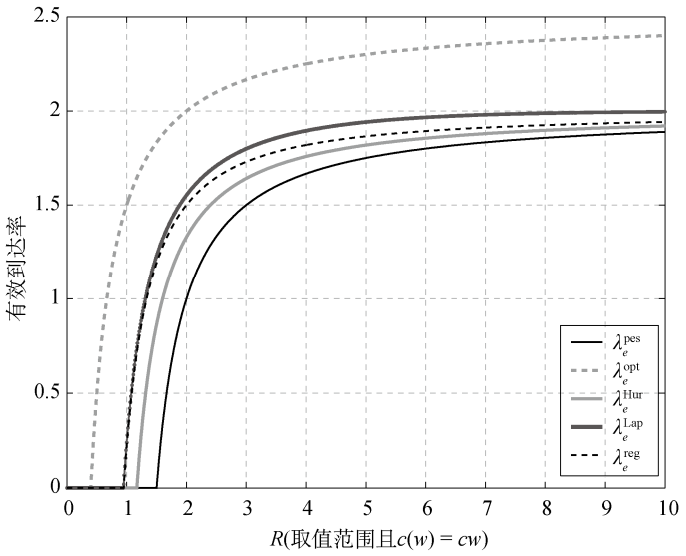

图 3-3 取值范围信息下的均衡有效到达率[$c(w)=cw$]

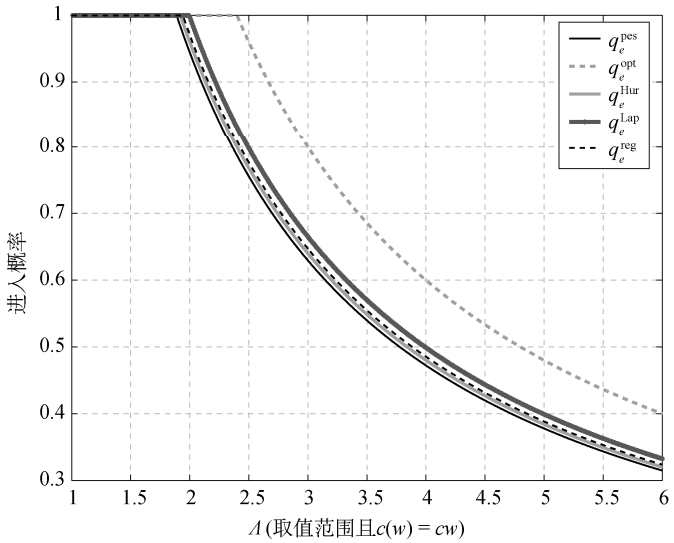

图 3-4　取值范围信息下的均衡进入概率[$c(w)=cw$]

3.2.3　二次等待损耗

不同于风险中立顾客, 具有二次等待损耗函数的风险规避顾客在等待 w 个单位时间后的等待损耗为 $c(w) = cw^2$。

3.2.3.1　两者取一

如果系统状态(μ_1, **1**)最终实现，顾客完成服务之后的平均剩余效用为

$$U_{\text{join}}^{(\mu_1,1)} = R - c\left(\frac{2}{(\mu_1-\lambda)^2}\right)$$

如果系统状态(μ_1, **0**)最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_1,0)} = R - c\left(\frac{2}{(\mu_1-\lambda)^2} + \frac{v^{(2)}}{(\mu_1-\lambda)v} + \frac{v^{(3)}}{3v}\right)$$

如果系统状态(μ_2, **1**)最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_2,1)} = R - c\left(\frac{2}{(\mu_2-\lambda)^2}\right)$$

如果系统状态(μ_2, **0**)最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_2,0)} = R - c\left(\frac{2}{(\mu_2-\lambda)^2} + \frac{v^{(2)}}{(\mu_2-\lambda)v} + \frac{v^{(3)}}{3v}\right)$$

1. 最大最小准则

求解下列方程

$$U_{\text{join}}^{(\mu_1,0)}=R-c\left(\frac{2}{(\mu_1-\lambda)^2}+\frac{v^{(2)}}{(\mu_1-\lambda)v}+\frac{v^{(3)}}{3v}\right)=0$$

得到如下命题中的结果。

命题 3.2.3 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且顾客的决策准则为最大最小准则，令

$$\Lambda^*=\mu_1-\frac{3cv^{(2)}+\sqrt{9c^2v^{(2)2}-24c^2v^{(3)}v+72Rcv^2}}{6Rv-2cv^{(3)}}$$

则具有二次等待损耗 $c(w)=cw^2$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\text{pes}_a}=\begin{cases}\Lambda, & \Lambda\leq\Lambda^*\\ \Lambda^*, & \Lambda>\Lambda^*\end{cases} \tag{3-14}$$

以及相应的均衡进入概率为 $q_{e(cw^2)}^{\text{pes}_a}=\lambda_{e(cw^2)}^{\text{pes}_a}/\Lambda$。

2. 最大最大准则

求解下列方程

$$U_{\text{join}}^{(\mu_2,1)}=R-c\left(\frac{2}{(\mu_2-\lambda)^2}\right)=0$$

得到如下命题中的结果。

命题 3.2.4 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且顾客的决策准则为最大最大准则，则具有二次等待损耗 $c(w)=cw^2$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\text{opt}_a}=\begin{cases}\Lambda, & \Lambda\leq\mu_2-\sqrt{\dfrac{2c}{R}}\\ \mu_2-\sqrt{\dfrac{2c}{R}}, & \Lambda>\mu_2-\sqrt{\dfrac{2c}{R}}\end{cases} \tag{3-15}$$

以及相应的均衡进入概率为 $q_{e(cw^2)}^{\text{opt}_a}=\lambda_{e(cw^2)}^{\text{opt}_a}/\Lambda$。

3. Hurwicz 准则

命题 3.2.5 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，以下方程

$$R-(1-\alpha)c\frac{v^{(3)}}{3v}-c\left(\alpha\frac{2}{(\mu_2-\lambda)^2}+(1-\alpha)\left(\frac{2}{(\mu_1-\lambda)^2}+\frac{v^{(2)}}{(\mu_1-\lambda)v}\right)\right)=0 \tag{3-16}$$

的解为 λ^*，且顾客的决策准则为 Hurwicz 准则，则具有二次等待损耗 $c(w)=cw^2$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\mathrm{Hur}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-17}$$

以及相应的均衡进入概率为 $q_{e(cw^2)}^{\mathrm{Hur}_a}=\lambda_{e(cw^2)}^{\mathrm{Hur}_a}/\Lambda$。

4. 最小最大后悔值准则

定理 3.2.6　如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，记以下方程

$$2R-\frac{2c}{(\mu_2-\lambda)^2}-c\left(\frac{2}{(\mu_1-\lambda)^2}+\frac{v^{(2)}}{(\mu_1-\lambda)v}+\frac{v^{(3)}}{3v}\right)=0 \tag{3-18}$$

的解为 λ^*，且顾客的决策准则为最小最大后悔值准则，则具有二次等待损耗 $c(w)=cw^2$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\mathrm{reg}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-19}$$

以及相应的均衡进入概率为 $q_{e(cw^2)}^{\mathrm{reg}_a}=\lambda_{e(cw^2)}^{\mathrm{reg}_a}/\Lambda$。

证明　（1）如果 $\Lambda\le\Lambda^*$，有如下的后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	0	0	0	0
止步	$R-\dfrac{2c}{(\mu_1-\Lambda)^2}$	$R-c\left(\dfrac{2c}{(\mu_1-\Lambda)^2}+\dfrac{v^{(2)}}{(\mu_1-\Lambda)v}+\dfrac{v^{(3)}}{3v}\right)$	$R-\dfrac{2c}{(\mu_2-\Lambda)^2}$	$R-c\left(\dfrac{2}{(\mu_2-\Lambda)^2}+\dfrac{v^{(2)}}{(\mu_2-\Lambda)v}+\dfrac{v^{(3)}}{3v}\right)$

因此，$\rho_{\mathrm{join}}=0$ 且 $\rho_{\mathrm{balk}}=R-2c/(\mu_2-\Lambda)^2$，最小最大后悔值为 $\rho^*=\min\{0,R-2c/(\mu_2-\Lambda)^2\}=0$，所有顾客将选择进入系统，即 $q_{e(cw^2)}^{\mathrm{reg}_a}=1$。

（2）如果 $\Lambda^*<\Lambda\le\mu_2-\sqrt{2c/R}$，有如下的后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$
进入	0 或 $\dfrac{2}{(\mu_1-\Lambda)^2}-R$	$c\left(\dfrac{2}{(\mu_1-\Lambda)^2}+\dfrac{v^{(2)}}{(\mu_1-\Lambda)v}+\dfrac{v^{(3)}}{3v}\right)-R$
止步	$R-\dfrac{2c}{(\mu_1-\Lambda)^2}$ 或 0	0

后悔值	$(\mu_2,1)$	$(\mu_2,0)$
进入	0	$c\left(\frac{2}{(\mu_2-\Lambda)^2}+\frac{v^{(2)}}{(\mu_2-\Lambda)v}+\frac{v^{(3)}}{3v}\right)-R$ 或 0
止步	$R-\frac{2c}{(\mu_2-\Lambda)^2}$	0 或 $R-c\left(\frac{2}{(\mu_2-\Lambda)^2}+\frac{v^{(2)}}{(\mu_2-\Lambda)v}+\frac{v^{(3)}}{3v}\right)$

因此，$\rho_{\text{join}}=c(v^{(3)}/(3v)+v^{(2)}/((\mu_1-\Lambda))v+2/(\mu_1-\Lambda)^2)-R$ 且 $\rho_{\text{balk}}=R-2c/(\mu_2-\Lambda)^2$，最小最大后悔值为 $\rho^*=\min\{c(v^{(3)}/(3v)+v^{(2)}/((\mu_1-\Lambda))v+2/(\mu_1-\Lambda)^2)-R, R-2c/(\mu_2-\Lambda)^2\}$。令

$$c\left(\frac{v^{(3)}}{3v}+\frac{v^{(2)}}{(\mu_1-\Lambda)v}+\frac{2}{(\mu_1-\Lambda)^2}\right)-R=R-\frac{2c}{(\mu_2-\Lambda)^2}$$

用其他参数表示 Λ，并用 λ^*表示该表达式。如果 $\Lambda^*<\Lambda<\lambda^*$，所有顾客选择进入，即 $q^{\text{reg}_a}_{e(cw^2)}=1$ ；如果 $\lambda^*\le\Lambda\le\mu_2-\sqrt{2c/R}$，所有顾客将以概率 $q^{\text{reg}_a}_{e(cw^2)}=\lambda^{\text{reg}_a}_{e(cw^2)}/\Lambda$ 选择进入。

（3）如果 $\Lambda>\mu_2-\sqrt{2c/R}$，有如下的后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	$\frac{2c}{(\mu_1-\Lambda)^2}-R$	$c\left(\frac{2}{(\mu_1-\Lambda)^2}+\frac{v^{(2)}}{(\mu_1-\Lambda)v}+\frac{v^{(3)}}{3v}\right)-R$	$\frac{2c}{(\mu_2-\Lambda)^2}-R$	$c\left(\frac{2}{(\mu_2-\Lambda)^2}+\frac{v^{(2)}}{(\mu_2-\Lambda)v}+\frac{v^{(3)}}{3v}\right)-R$
止步	0	0	0	0

因此，$\rho_{\text{join}}=c(v^{(3)}/(3v)+v^{(2)}/((\mu_1-\Lambda))v+2/(\mu_1-\Lambda)^2)-R$ 且 $\rho_{\text{balk}}=0$，最小最大后悔值为 $\rho^*=\min\left\{c(v^{(3)}/(3v)+v^{(2)}/((\mu_1-\Lambda))v+2/(\mu_1-\Lambda)^2)-R,0\right\}=0$，说明所有顾客需以概率 $q^{\text{reg}_a}_{e(cw^2)}=\lambda^*/\Lambda$ 选择混合策略。

结合以上讨论，得到定理中的结论。

5. Laplace 准则

定理 3.2.7 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，以下方程

$$4R-4c\left(\frac{1}{(\mu_1-\lambda)^2}+\frac{1}{(\mu_2-\lambda)^2}\right)-c\left(\frac{v^{(2)}}{(\mu_1-\lambda)v}+\frac{v^{(2)}}{(\mu_2-\lambda)v}\right)-\frac{2cv^{(3)}}{3v}=0 \quad (3\text{-}20)$$

的解为 λ^*，且顾客的决策准则为 Laplace 准则，则具有二次等待损耗 $c(w)=cw^2$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\mathrm{Lap}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-21}$$

以及相应的均衡进入概率为 $q_{e(cw^2)}^{\mathrm{Lap}_a}=\lambda_{e(cw^2)}^{\mathrm{Lap}_a}/\Lambda$。

证明　顾客进入系统的平均剩余效用为

$$\begin{aligned}U_{join}&=\frac{1}{4}\left(R-\frac{2c}{(\mu_1-\lambda)^2}\right)+\frac{1}{4}\left(R-c\left(\frac{2}{(\mu_1-\lambda)^2}+\frac{v^{(2)}}{(\mu_1-\lambda)v}+\frac{v^{(3)}}{3v}\right)\right)\\&\quad+\frac{1}{4}\left(R-\frac{2c}{(\mu_2-\lambda)^2}\right)+\frac{1}{4}\left(R-c\left(\frac{2}{(\mu_2-\lambda)^2}+\frac{v^{(2)}}{(\mu_2-\lambda)v}+\frac{v^{(3)}}{3v}\right)\right)\\&=R-c\left(\frac{1}{(\mu_1-\lambda)^2}+\frac{1}{(\mu_2-\lambda)^2}\right)-\frac{1}{4}c\left(\frac{v^{(2)}}{(\mu_1-\lambda)v}+\frac{v^{(2)}}{(\mu_2-\lambda)v}\right)-\frac{cv^{(3)}}{6v}\end{aligned}$$

而离开系统的平均效用为 $U_{\mathrm{balk}}=0$。因此，最大平均效用为

$$U^{\mathrm{Lap}}=\max\left\{R-c\left(\frac{1}{(\mu_1-\lambda)^2}+\frac{1}{(\mu_2-\lambda)^2}\right)-\frac{1}{4}c\left(\frac{v^{(2)}}{(\mu_1-\lambda)v}+\frac{v^{(2)}}{(\mu_2-\lambda)v}\right)-\frac{cv^{(3)}}{6v},0\right\}$$

求解 $U_{\mathrm{join}}=0$，得到 $\lambda_{e(cw^2)}^{\mathrm{Lap}_a}$。

图 3-5 和图 3-6 比较了此种情况下具有各种决策准则的顾客的均衡有效到达率。我们观察到，$\lambda_e^{\mathrm{opt}}\ge\lambda_e^{\mathrm{Hur}}\ge\lambda_e^{\mathrm{Lap}}\ge\lambda_e^{\mathrm{reg}}\ge\lambda_e^{\mathrm{pes}}$ 且 $q_e^{\mathrm{opt}}\ge q_e^{\mathrm{Hur}}\ge q_e^{\mathrm{Lap}}\ge q_e^{\mathrm{reg}}\ge q_e^{\mathrm{pes}}$。显然，它们的值都小于（至少不大于）风险中立情况下的值。

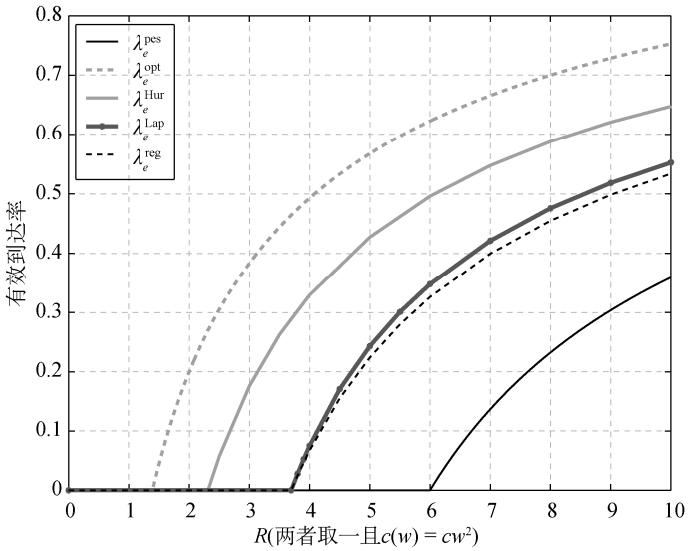

图 3-5　两者取一信息下的均衡有效到达率[$c(w)=cw^2$]

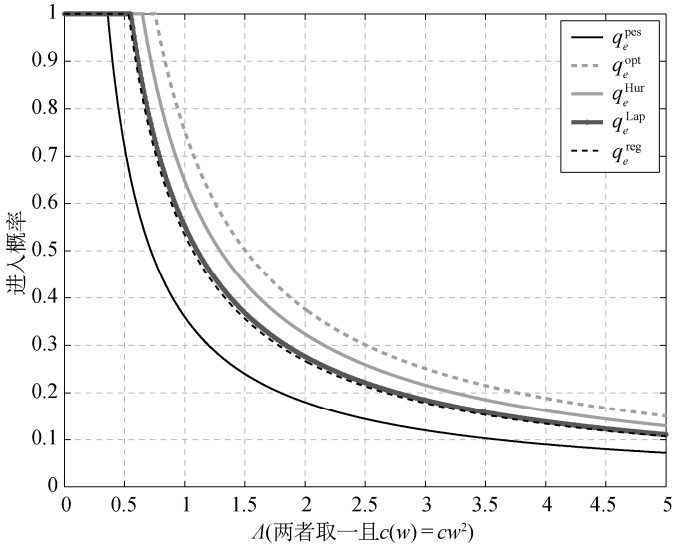

图 3-6　两者取一信息下的均衡进入概率[$c(w)=cw^2$]

3.2.3.2　取值范围

假定关于服务时间参数的部分信息为 $\mu\in\left[\underline{\mu},\overline{\mu}\right]$，$0<\underline{\mu}<\overline{\mu}$。对于具有最大最小决策准则的顾客，他们的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\mathrm{pes}_r}=\underline{\mu}-\frac{3cv^{(2)}+\sqrt{9c^2v^{(2)2}-24c^2v^{(3)}v+72Rcv^2}}{6Rv-2cv^{(3)}} \tag{3-22}$$

具有最大最大决策准则的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\mathrm{opt}_r}=\overline{\mu}-\sqrt{\frac{2c}{R}} \tag{3-23}$$

求解以下方程

$$R-(1-\alpha)c\frac{v^{(3)}}{3v}-c\left(\alpha\frac{2}{(\overline{\mu}-\lambda)^2}+(1-\alpha)\left(\frac{2}{(\underline{\mu}-\lambda)^2}+\frac{v^{(2)}}{(\underline{\mu}-\lambda)v}\right)\right)=0 \tag{3-24}$$

可得到具有 Hurwicz 准则的顾客的均衡有效到达率 $\lambda_{e(cw^2)}^{\mathrm{Hur}_r}$。

定理 3.2.8　如果服务时间参数 μ 处于区间$\left[\underline{\mu},\overline{\mu}\right]$, $0<\underline{\mu}<\overline{\mu}$ 内，以下方程

$$2R-\frac{2c}{(\overline{\mu}-\lambda)^2}-c\left(\frac{2}{(\underline{\mu}-\lambda)^2}+\frac{v^{(2)}}{(\underline{\mu}-\lambda)v}+\frac{v^{(3)}}{3v}\right)=0 \tag{3-25}$$

的解为 λ^*，且顾客的决策准则为最小最大后悔值准则，则具有二次等待损耗 $c(w)=cw^2$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\mathrm{reg}_r}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-26}$$

以及相应的均衡进入概率为 $q_{e(cw^2)}^{\mathrm{reg}_r}=\lambda_{e(cw^2)}^{\mathrm{reg}_r}/\Lambda$。

定理 3.2.9　如果服务时间参数 μ 处于区间 $[\underline{\mu},\overline{\mu}]$ 内，$0<\underline{\mu}<\overline{\mu}$，以下方程

$$R-\frac{2c}{\overline{\mu}-\underline{\mu}}\left(\frac{1}{\underline{\mu}-\lambda}-\frac{1}{\overline{\mu}-\lambda}\right)-\frac{c}{\overline{\mu}-\underline{\mu}}\frac{v^{(2)}}{2v}\ln\left(\frac{\overline{\mu}-\lambda}{\underline{\mu}-\lambda}\right)-\frac{cv^{(3)}}{6v}=0 \tag{3-27}$$

的解为 λ^*，且顾客的决策准则为 Laplace 准则，则具有二次等待损耗 $c(w)=cw^2$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^2)}^{\mathrm{Lap}_r}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-28}$$

以及相应的均衡进入概率为 $q_{e(cw^2)}^{\mathrm{Lap}_r}=\lambda_{e(cw^2)}^{\mathrm{Lap}_r}/\Lambda$。

证明　基于 Laplace 准则的特点，顾客进入系统的平均剩余效用为

$$\begin{aligned}U_{\mathrm{join}}&=\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}}\left(R-\frac{2c}{(\mu-\lambda)^2}\right)\frac{1}{\overline{\mu}-\underline{\mu}}\mathrm{d}\mu+\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}}\left(R-c\left(\frac{2}{(\mu-\lambda)^2}+\frac{v^{(2)}}{(\mu-\lambda)v}+\frac{v^{(3)}}{3v}\right)\right)\frac{1}{\overline{\mu}-\underline{\mu}}\mathrm{d}\mu\\&=\frac{1}{2(\overline{\mu}-\underline{\mu})}\left(R\mu+\frac{2c}{\mu-\lambda}\right)\Big|_{\underline{\mu}}^{\overline{\mu}}+\frac{1}{2(\overline{\mu}-\underline{\mu})}\left(R\mu\Big|_{\underline{\mu}}^{\overline{\mu}}-c\int_{\underline{\mu}}^{\overline{\mu}}\left(\frac{2}{(\mu-\lambda)^2}+\frac{v^{(2)}}{(\mu-\lambda)v}+\frac{v^{(3)}}{3v}\right)\mathrm{d}\mu\right)\\&=R-\frac{c}{(\overline{\mu}-\underline{\mu})}\left(\frac{1}{\underline{\mu}-\lambda}-\frac{1}{\overline{\mu}-\lambda}\right)-\frac{c}{2(\overline{\mu}-\underline{\mu})}\left(\frac{-2}{\mu-\lambda}+\frac{v^{(2)}}{v}\ln(\mu-\lambda)+\frac{v^{(3)}}{3v}\mu\right)\Big|_{\underline{\mu}}^{\overline{\mu}}\\&=R-\frac{2c}{\overline{\mu}-\underline{\mu}}\left(\frac{1}{\underline{\mu}-\lambda}-\frac{1}{\overline{\mu}-\lambda}\right)-\frac{c}{(\overline{\mu}-\underline{\mu})}\frac{v^{(2)}}{2v}\ln\left(\frac{\overline{\mu}-\lambda}{\underline{\mu}-\lambda}\right)-\frac{cv^{(3)}}{6v}\end{aligned}$$

而顾客离开的平均效用为 $U_{\mathrm{balk}}=0$。因此，最大平均效用为

$$U^{Lap}=\max\left\{R-\frac{2c}{\overline{\mu}-\underline{\mu}}\left(\frac{1}{\underline{\mu}-\lambda}-\frac{1}{\overline{\mu}-\lambda}\right)-\frac{c}{\overline{\mu}-\underline{\mu}}\frac{v^{(2)}}{2v}\ln\left(\frac{\overline{\mu}-\lambda}{\underline{\mu}-\lambda}\right)-\frac{cv^{(3)}}{6v},0\right\}$$

求解 $U_{\mathrm{join}}=0$，得到 $\lambda_{e(cw^2)}^{\mathrm{Lap}_r}$。

图 3-7 和图 3-8 比较了此种情况下具有各种决策准则的顾客的均衡有效到达率。显然，$\lambda_e^{\mathrm{opt}}\ge\lambda_e^{\mathrm{Lap}}\ge\lambda_e^{\mathrm{reg}}\ge\lambda_e^{\mathrm{Hur}}\ge\lambda_e^{\mathrm{pes}}$ 且 $q_e^{\mathrm{opt}}\ge q_e^{\mathrm{Lap}}\ge q_e^{\mathrm{reg}}\ge q_e^{\mathrm{Hur}}\ge q_e^{\mathrm{pes}}$。与图 3-5 和图 3-6 类似，它们的值也都不大于风险中立情况下的值。

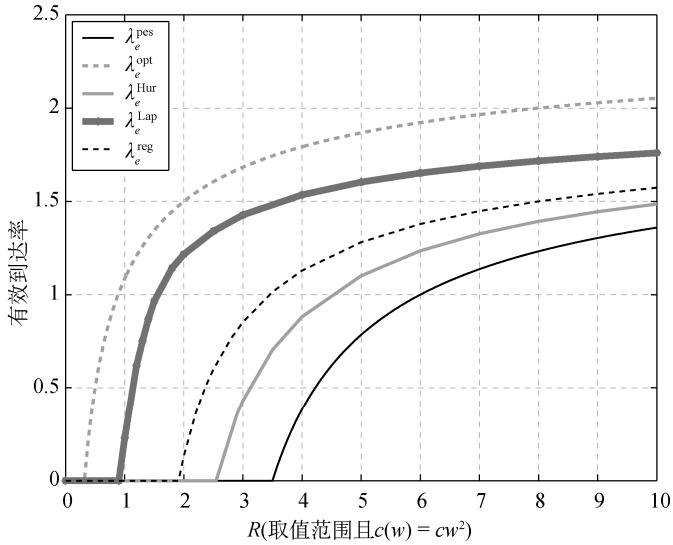

图 3-7　取值范围信息下的均衡有效到达率[$c(w) = cw^2$]

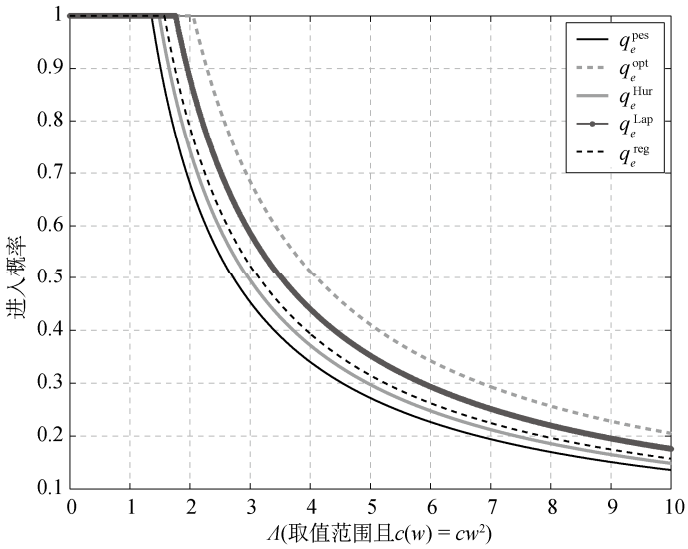

图 3-8　取值范围信息下的均衡进入概率[$c(w)= cw^2$]

3.2.4　指数等待损耗

除了具有二次等待损耗的风险规避顾客，我们继续考虑另一类具有指数等待损耗的风险规避顾客，即 $c(w) = ce^{aw}$ $(a > 0)$。特别地，假设休假时间服从参数为 $1/\nu(> a)$的指数分布。

3.2.4.1　两者取一

如果系统状态$(\mu_1, \mathbf{1})$最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_1,1)} = R - \frac{c(\mu_1 - \lambda)}{(\mu_1 - \lambda) - a}$$

如果系统状态$(\mu_1, \mathbf{0})$最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_1,0)} = R - \frac{c(\mu_1 - \lambda)}{((\mu_1 - \lambda) - a)(1 - va)}$$

如果系统状态$(\mu_2, \mathbf{1})$最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_2,1)} = R - \frac{c(\mu_2 - \lambda)}{(\mu_2 - \lambda) - a}$$

如果系统状态$(\mu_2, \mathbf{0})$最终实现，顾客的平均剩余效用为

$$U_{\text{join}}^{(\mu_2,0)} = R - \frac{c(\mu_2 - \lambda)}{((\mu_2 - \lambda) - a)(1 - va)}$$

1. 最大最小准则

求解以下方程

$$U_{\text{join}}^{(\mu_1,0)} = R - \frac{c(\mu_1 - \lambda)}{((\mu_1 - \lambda) - a)(1 - va)} = 0$$

得到如下命题中的结果。

命题 3.2.6　如果服务时间参数μ为μ_1和μ_2两者中的一个，$0 < \mu_1 < \mu_2$，且满足

$$\mu_1 - \frac{Ra(1 - va)}{R(1 - va) - c} \in \left(\mu_1 - \frac{1}{v}, \mu_1 - a \right) \tag{3-29}$$

则具有最大最小准则和指数等待损耗$c(w) = ce^{aw}\ (a > 0)$的顾客的均衡有效到达率为

$$\lambda_{e(ce^{aw})}^{pes_a} = \begin{cases} \Lambda, & \Lambda \leq \mu_1 - \dfrac{Ra(1 - va)}{R(1 - va) - c} \\ \mu_1 - \dfrac{Ra(1 - va)}{R(1 - va) - c}, & \Lambda > \mu_1 - \dfrac{Ra(1 - va)}{R(1 - va) - c} \end{cases} \tag{3-30}$$

以及相应的均衡进入概率为$q_{e(ce^{aw})}^{\text{pes}_a} = \lambda_{e(ce^{aw})}^{\text{pes}_a} / \Lambda$。

2. 最大最大准则

求解以下方程

$$U_{\text{join}}^{(\mu_2,1)} = R - \frac{c(\mu_2 - \lambda)}{(\mu_2 - \lambda) - a} = 0$$

得到如下命题中的结果。

命题 3.2.7　如果服务时间参数μ为μ_1和μ_2两者中的一个，$0 < \mu_1 < \mu_2$，且满足

$$\mu_2-\frac{Ra}{R-c}\in\left(\mu_2-\frac{1}{v},\ \mu_2-a\right) \tag{3-31}$$

则具有最大最大准则和指数等待损耗 $c(w)=ce^{aw}\ (a>0)$的顾客的均衡有效到达率为

$$\lambda_{e(ce^{aw})}^{\mathrm{opt}_a}=\begin{cases}\Lambda, & \Lambda\le\mu_2-\dfrac{Ra}{R-c}\\ \mu_2-\dfrac{Ra}{R-c}, & \Lambda>\mu_2-\dfrac{Ra}{R-c}\end{cases} \tag{3-32}$$

以及相应的均衡进入概率为 $q_{e(ce^{aw})}^{\mathrm{opt}_a}=\lambda_{e(ce^{aw})}^{\mathrm{opt}_a}/\Lambda$。

3. Hurwicz 准则

命题 3.2.8 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且如果以下方程

$$R-\alpha\frac{c(\mu_2-\lambda)}{(\mu_2-\lambda)-a}-(1-\alpha)\frac{c(\mu_1-\lambda)}{((\mu_1-\lambda)-a)(1-va)}=0 \tag{3-33}$$

的解 λ^*处于区间 $(\alpha\mu_2+(1-\alpha)\mu_1-1/v,\ \alpha\mu_2+(1-\alpha)\mu_1-a)$ 之内，则具有 Hurwicz 准则和指数等待损耗 $c(w)=ce^{aw}\ (a>0)$的顾客的均衡有效到达率为

$$\lambda_{e(cw^{aw})}^{\mathrm{Hur}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-34}$$

以及相应的均衡进入概率为 $q_{e(ce^{aw})}^{\mathrm{Hur}_a}=\lambda_{e(ce^{aw})}^{\mathrm{Hur}_a}/\Lambda$。

4. 最小最大后悔值准则

定理 3.2.10 如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且如果以下方程

$$2R-\frac{c(\mu_2-\lambda)}{(\mu_2-\lambda)-a}-\frac{c(\mu_1-\lambda)}{((\mu_1-\lambda)-a)(1-va)}=0 \tag{3-35}$$

的解 λ^*处于区间 $(\mu_1-1/v,\mu_1-a)$ 之内，则具有最小最大后悔值准则和指数等待损耗 $c(w)=ce^{aw}\ (a>0)$ 的顾客的均衡有效到达率为

$$\lambda_{e(cw^{aw})}^{\mathrm{reg}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-36}$$

以及相应的均衡进入概率为 $q_{e(ce^{aw})}^{\mathrm{reg}_a}=\lambda_{e(ce^{aw})}^{\mathrm{reg}_a}/\Lambda$。

证明 （1）如果 $\Lambda\le\mu_1-Ra(1-va)/(R-(1-va)-c)$，有如下后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	0	0	0	0
止步	$R-\dfrac{c(\mu_1-\Lambda)}{(\mu_1-\Lambda)-a}$	$R-\dfrac{c(\mu_1-\Lambda)}{((\mu_1-\Lambda)-a)(1-va)}$	$R-\dfrac{c(\mu_2-\Lambda)}{(\mu_2-\Lambda)-a}$	$R-\dfrac{c(\mu_2-\Lambda)}{((\mu_2-\Lambda)-a)(1-va)}$

因此，$\rho_{\text{join}}=0$ 且 $\rho_{\text{balk}}=R-c(\mu_2-\Lambda)/((\mu_2-\Lambda)-a)$，最小最大后悔值为 $\rho^*=\min\{0, R-c(\mu_2-\Lambda)/((\mu_2-\Lambda)-a)\}=0$，所有顾客将选择进入系统，即 $q_{e(ce^{aw})}^{\text{reg}_a}=1$。

（2）如果 $\mu_1-Ra(1-va)/(R(1-va)-c)<\Lambda\leq\mu_2-Ra/(R-c)$，有如下后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	0 或 $\dfrac{c(\mu_1-\Lambda)}{(\mu_1-\Lambda)-a}-R$	$\dfrac{c(\mu_1-\Lambda)}{((\mu_1-\Lambda)-a)(1-va)}-R$	0	$\dfrac{c(\mu_2-\Lambda)}{((\mu_2-\Lambda)-a)(1-va)}-R$ 或 0
止步	$R-\dfrac{c(\mu_1-\Lambda)}{(\mu_1-\Lambda)-a}$ 或 0	0	$R-\dfrac{c(\mu_2-\Lambda)}{(\mu_2-\Lambda)-a}$	0 或 $R-\dfrac{c(\mu_2-\Lambda)}{((\mu_2-\Lambda)-a)(1-va)}$

因此，$\rho_{\text{join}}=c(\mu_1-\Lambda)/(((\mu_1-\Lambda)-a)(1-va))-R$ 且 $\rho_{\text{balk}}=R-c(\mu_2-\Lambda)/((\mu_2-\Lambda)-a)$，最小最大后悔值为 $\rho^*=\min\{c(\mu_1-\Lambda)/(((\mu_1-\Lambda)-a)(1-va))-R, R-c(\mu_2-\Lambda)/((\mu_2-\Lambda)-a)\}$。令

$$\frac{c(\mu_1-\Lambda)}{((\mu_1-\Lambda)-a)(1-va)}-R=R-\frac{c(\mu_2-\Lambda)}{(\mu_2-\Lambda)-a}$$

用其他参数表示 Λ，且用 λ^*表示该表达式。如果 $\mu_1-Ra(1-va)/(R(1-va)-c)<\Lambda<\lambda^*$，所有顾客选择进入；如果 $\lambda^*\leq\Lambda\leq\mu_2-Ra/(R-c)$，顾客以概率 $q_{e(ce^{aw})}^{\text{reg}_a}=\lambda^*/\Lambda$ 选择进入。

（3）如果 $\Lambda>\mu_2-Ra/(R-c)$，有如下后悔值表格：

后悔值	$(\mu_1,1)$	$(\mu_1,0)$	$(\mu_2,1)$	$(\mu_2,0)$
进入	$\dfrac{c(\mu_1-\Lambda)}{(\mu_1-\Lambda)-a}-R$	$\dfrac{c(\mu_1-\Lambda)}{((\mu_1-\Lambda)-a)(1-va)}-R$	$\dfrac{c(\mu_2-\Lambda)}{(\mu_2-\Lambda)-a}-R$	$\dfrac{c(\mu_2-\Lambda)}{((\mu_2-\Lambda)-a)(1-va)}-R$
止步	0	0	0	0

因此，$\rho_{\text{join}}=c(\mu_1-\Lambda)/(((\mu_1-\Lambda)-a)(1-va))-R$ 且 $\rho_{\text{balk}}=0$，最小最大后悔值为 $\rho^*=\min\{c(\mu_1-\Lambda)/(((\mu_1-\Lambda)-a)(1-va))-R, 0\}=0$，说明所有顾客需以概率 $q_{e(ce^{aw})}^{\text{reg}_a}=\lambda^*/\Lambda$ 选择混合策略。

结合以上讨论，得到定理中的结论。

5. Laplace 准则

定理 3.2.11　如果服务时间参数 μ 为 μ_1 和 μ_2 两者中的一个，$0<\mu_1<\mu_2$，且如果以下方程

$$4R-\frac{c(\mu_1-\lambda)(2-va)}{((\mu_1-\lambda)-a)(1-va)}-\frac{c(\mu_2-\lambda)(2-va)}{((\mu_2-\lambda)-a)(1-va)}=0 \tag{3-37}$$

的解λ^*处于区间$((\mu_1+\mu_2)/2-1/v,(\mu_1+\mu_2)/2-a)$之内，则具有Laplace准则和指数等待损耗$c(w)=ce^{aw}(a>0)$的顾客的均衡有效到达率为

$$\lambda_{e(ce^{aw})}^{\mathrm{Lap}_a}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-38}$$

以及相应的均衡进入概率为$q_{e(ce^{aw})}^{\mathrm{Lap}_a}=\lambda_{e(ce^{aw})}^{\mathrm{Lap}_a}/\Lambda$。

证明 顾客进入系统的平均剩余效用为

$$\begin{aligned}U_{join}&=\frac{1}{4}\left(R-\frac{c(\mu_1-\lambda)}{(\mu_1-\lambda)-a}\right)+\frac{1}{4}\left(R-\frac{c(\mu_1-\lambda)}{((\mu_1-\lambda)-a)(1-va)}\right)\\&\quad+\frac{1}{4}\left(R-\frac{c(\mu_2-\lambda)}{(\mu_2-\lambda)-a}\right)+\frac{1}{4}\left(R-\frac{c(\mu_1-\lambda)}{((\mu_1-\lambda)-a)(1-va)}\right)\\&=R-\frac{c(\mu_1-\lambda)(2-va)}{4((\mu_1-\lambda)-a)(1-va)}-\frac{c(\mu_2-\lambda)(2-va)}{4((\mu_2-\lambda)-a)(1-va)},\end{aligned}$$

而顾客离开的平均效用为$U_{\mathrm{balk}}=0$。因此，最大平均效用为

$$U^{\mathrm{Lap}}=\max\left\{R-\frac{c(\mu_1-\lambda)(2-va)}{4((\mu_1-\lambda)-a)(1-va)}-\frac{c(\mu_2-\lambda)(2-va)}{4((\mu_2-\lambda)-a)(1-va)},0\right\}$$

求解$U_{\mathrm{join}}=0$，得到$\lambda_{e(ce^{aw})}^{\mathrm{Lap}_a}$。

图 3-9 和图 3-10 比较了此种情况下具有各种决策准则的顾客的均衡有效到达率和均衡进入概率。令$a=0.1$，显然，$\lambda_e^{\mathrm{opt}}\ge\lambda_e^{\mathrm{Hur}}\ge\lambda_e^{\mathrm{Lap}}\ge\lambda_e^{\mathrm{reg}}\ge\lambda_e^{\mathrm{pes}}$且$q_e^{\mathrm{opt}}\ge q_e^{\mathrm{Hur}}\ge q_e^{\mathrm{Lap}}\ge q_e^{\mathrm{reg}}\ge q_e^{\mathrm{pes}}$。而且，$\lambda_e^{\mathrm{Lap}}\approx\lambda_e^{\mathrm{reg}}$且$q_e^{\mathrm{Lap}}\approx q_e^{\mathrm{reg}}$。重置$a=1.0$，图 3-11 和图 3-12 表明顾客的均衡有效到达率和均衡进入概率相比图 3-9 和图 3-10 发生了锐减。这说明具有参数$a=1.0$的顾客比具有参数$a=0.1$的顾客对部分信息可能导致的附加延迟更加敏感。

3.2.4.2 取值范围

对于具有最大最小决策准则的顾客，如果$\underline{\mu}-Ra(1-va)/(R(1-va)-c)\in(\underline{\mu}-1/v,\ \underline{\mu}-a)$，则他们的均衡有效到达率为

$$\lambda_{e(ce^{aw})}^{\mathrm{pes}_r}=\underline{\mu}-\frac{Ra(1-va)}{R(1-va)-c} \tag{3-39}$$

对于具有最大最大决策准则的顾客，如果$\overline{\mu}-Ra/(R-c)\in(\overline{\mu}-1/v,\ \overline{\mu}-a)$，则他们的均衡有效到达率为

$$\lambda_{e(ce^{aw})}^{\mathrm{opt}_r}=\overline{\mu}-\frac{Ra}{R-c} \tag{3-40}$$

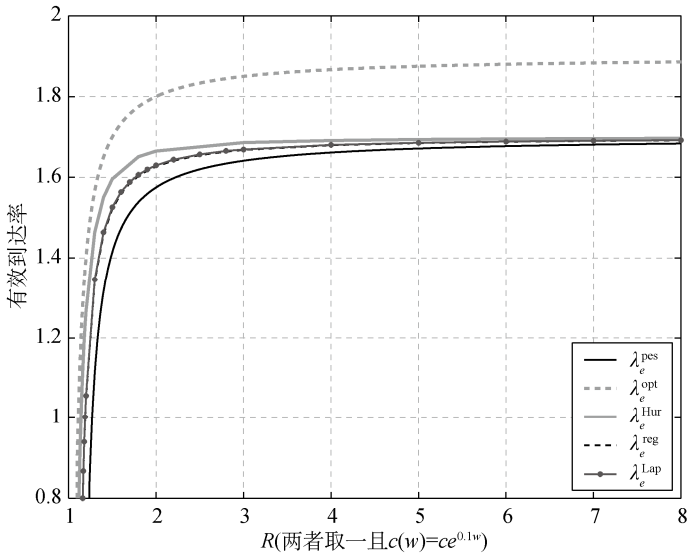

图 3-9　两者取一信息下的均衡有效到达率[$c(w) = ce^{aw}$ 且 $a = 0.1$]

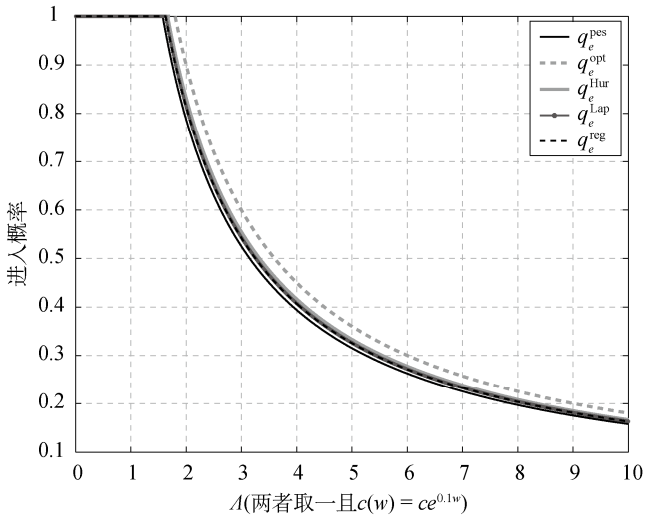

图 3-10　两者取一信息下的均衡进入概率[$c(w) = ce^{aw}$ 且 $a = 0.1$]

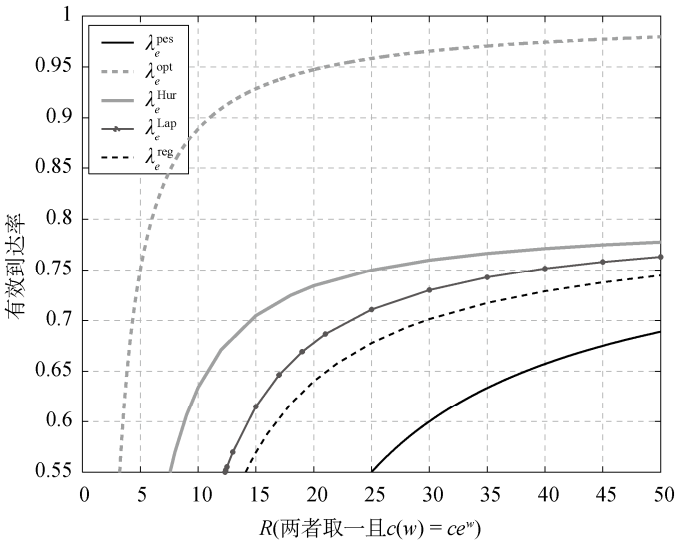

图 3-11　两者取一信息下的均衡有效到达率[$c(w) = ce^{aw}$ 且 $a = 1$]

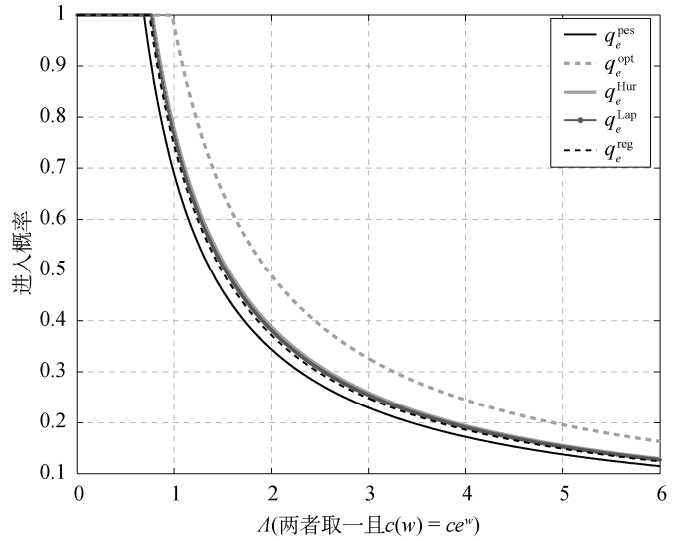

图 3-12　两者取一信息下的均衡进入概率[$c(w) = ce^{aw}$ 且 $a = 1$]

对于具有 Hurwicz 决策准则的顾客，如果 $\lambda \in (\alpha\overline{\mu}+(1-\alpha)\underline{\mu}-1/v, \alpha\overline{\mu}+(1-\alpha)\underline{\mu}-a)$，可以通过求解方程

$$R-\alpha\frac{c(\overline{\mu}-\lambda)}{(\overline{\mu}-\lambda)-a}-(1-\alpha)\frac{c(\underline{\mu}-\lambda)}{((\underline{\mu}-\lambda)-a)(1-va)}=0 \tag{3-41}$$

得到顾客的均衡有效到达率 $\lambda^{\mathrm{Hur}_r}_{e(ce^{aw})}$。

定理 3.2.12　如果服务时间参数 μ 处于区间 $[\underline{\mu},\overline{\mu}]$ 内，$0<\underline{\mu}<\overline{\mu}$，且如果以下方程

$$2R-\frac{c(\overline{\mu}-\lambda)}{(\overline{\mu}-\lambda)-a}-\frac{c(\underline{\mu}-\lambda)}{((\underline{\mu}-\lambda)-a)(1-va)}=0 \tag{3-42}$$

的解 λ^* 处于区间 $(\underline{\mu}-1/v, \underline{\mu}-a)$ 之内，则具有最小最大后悔值准则和指数等待损耗 $c(w)=ce^{aw}(a>0)$ 的顾客的均衡有效到达率为

$$\lambda^{\mathrm{reg}_r}_{e(ce^{aw})}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-43}$$

以及相应的均衡进入概率为 $q^{\mathrm{reg}_r}_{e(ce^{aw})}=\lambda^{\mathrm{reg}_r}_{e(ce^{aw})}/\Lambda$。

定理 3.2.13　如果服务时间参数 μ 处于区间 $[\underline{\mu},\overline{\mu}],0<\underline{\mu}<\overline{\mu}$ 内，且如果以下方程

$$2R-\frac{c(2-va)}{1-va}-\frac{ac(2-va)}{(\overline{\mu}-\underline{\mu})(1-va)}\ln\left(\frac{(\overline{\mu}-\lambda)-a}{(\underline{\mu}-\lambda)-a}\right)=0 \tag{3-44}$$

的解 λ^* 处于区间 $((\underline{\mu}+\overline{\mu})/2-1/v, (\underline{\mu}+\overline{\mu})/2-a)$ 之内，则具有 Laplace 准则和指数等待损耗 $c(w)=ce^{aw}(a>0)$ 的顾客均衡有效到达率为

$$\lambda^{\mathrm{Lap}_r}_{e(ce^{aw})}=\begin{cases}\Lambda, & \Lambda\le\lambda^*\\ \lambda^*, & \Lambda>\lambda^*\end{cases} \tag{3-45}$$

以及相应的均衡进入概率为 $q^{\mathrm{Lap}_r}_{e(ce^{aw})}=\lambda^{\mathrm{Lap}_r}_{e(ce^{aw})}/\Lambda$。

证明　基于 Laplace 准则的特点，顾客进入系统的平均剩余效用为

$$U_{\mathrm{join}}=\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}}\left(R-\frac{c(\mu-\lambda)}{(\mu-\lambda)-a}\right)\frac{1}{\overline{\mu}-\underline{\mu}}\mathrm{d}\mu$$
$$+\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}}\left(R-\frac{c(\mu-\lambda)}{1-v(\mu-\lambda)}\left(\frac{1}{(\mu-\lambda)-a}-\frac{v}{1-va}\right)\right)\frac{1}{\overline{\mu}-\underline{\mu}}\mathrm{d}\mu$$

其中

$$\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}}\frac{c(\mu-\lambda)}{(\mu-\lambda)-a}\frac{1}{\overline{\mu}-\underline{\mu}}\mathrm{d}\mu=\frac{1}{2(\overline{\mu}-\underline{\mu})}\int_{\underline{\mu}}^{\overline{\mu}}\left(c+\frac{ac}{(\mu-\lambda)-a}\right)\mathrm{d}\mu$$

$$= \frac{c}{2} + \frac{ac}{2(\overline{\mu}-\underline{\mu})} \ln\left(\frac{(\overline{\mu}-\lambda)-a}{(\underline{\mu}-\lambda)-a} \right)$$

$$\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}} \frac{c(\mu-\lambda)}{1-v(\mu-\lambda)} \frac{v}{1-va} \frac{1}{\overline{\mu}-\underline{\mu}} \mathrm{d}\mu = \frac{c}{2(\overline{\mu}-\underline{\mu})(1-va)} \int_{\underline{\mu}}^{\overline{\mu}} \left(\frac{1}{1-v(\mu-\lambda)} - 1 \right) \mathrm{d}\mu$$

$$= \frac{c}{2v(\overline{\mu}-\underline{\mu})(1-va)} \ln\left(\frac{1-v(\overline{\mu}-\lambda)}{1-v(\underline{\mu}-\lambda)} \right) - \frac{c}{2(1-va)}$$

$$\frac{1}{2}\int_{\underline{\mu}}^{\overline{\mu}} \frac{c(\mu-\lambda)}{((\mu-\lambda)-a)(1-v(\mu-\lambda))} \frac{1}{\overline{\mu}-\underline{\mu}} \mathrm{d}\mu$$

$$= \frac{1}{2(\overline{\mu}-\underline{\mu})} \left(\int_{\underline{\mu}}^{\overline{\mu}} \frac{ac}{1-va} \frac{1}{(\mu-\lambda)-a} \mathrm{d}\mu + \int_{\underline{\mu}}^{\overline{\mu}} \frac{c}{1-va} \frac{1}{1-v(\mu-\lambda)} \mathrm{d}\mu \right)$$

$$= \frac{1}{2(\overline{\mu}-\underline{\mu})} \left(\frac{ac}{1-va} \ln\left(\frac{(\overline{\mu}-\lambda)-a}{(\underline{\mu}-\lambda)-a} \right) - \frac{c}{v(1-va)} \ln\left(\frac{1-v(\overline{\mu}-\lambda)}{1-v(\underline{\mu}-\lambda)} \right) \right)$$

因此，

$$U_{\text{join}} = R - \frac{c(2-va)}{2(1-va)} - \frac{ac(2-va)}{2(\overline{\mu}-\underline{\mu})(1-va)} \ln\left(\frac{(\overline{\mu}-\lambda)-a}{(\underline{\mu}-\lambda)-a} \right)$$

而顾客离开的平均效用为$U_{\text{balk}} = 0$。因此，最大平均效用为

$$U^{\text{Lap}} = \max\left\{ R - \frac{c(2-va)}{2(1-va)} - \frac{ac(2-va)}{2(\overline{\mu}-\underline{\mu})(1-va)} \ln\left(\frac{(\overline{\mu}-\lambda)-a}{(\underline{\mu}-\lambda)-a} \right), 0 \right\}$$

求解 $U_{\text{join}} = 0$，得到 $\lambda_{e(ce^{aw})}^{\text{Lap}_r}$。

图 3-13 和图 3-14 比较了此种情况下具有各种决策准则的顾客的均衡有效到达率和均衡进入概率。令 $a = 0.1$，$\lambda_e^{\text{opt}} \geq \lambda_e^{\text{Lap}} \geq \lambda_e^{\text{reg}} \geq \lambda_e^{\text{Hur}} \geq \lambda_e^{\text{pes}}$ 且 $q_e^{\text{opt}} \geq q_e^{\text{Lap}} \geq$

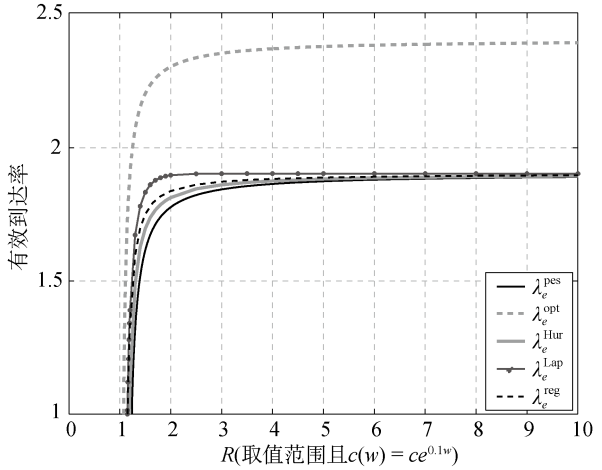

图 3-13　取值范围信息下的均衡有效到达率[$c(w) = ce^{aw}$ 且 $a = 0.1$]

$q_e^{reg} \geq q_e^{Hur} \geq q_e^{pes}$。类似于图 3-9，图 3-13 说明 λ_e^{Lap}, λ_e^{reg}, λ_e^{Hur} 和 λ_e^{pes} 的极限值几乎相等。重置 a=1.1，图 3-15 和图 3-16 也表明顾客的均衡有效到达率和均衡进入概率相比图 3-13 和图 3-14 发生了锐减。

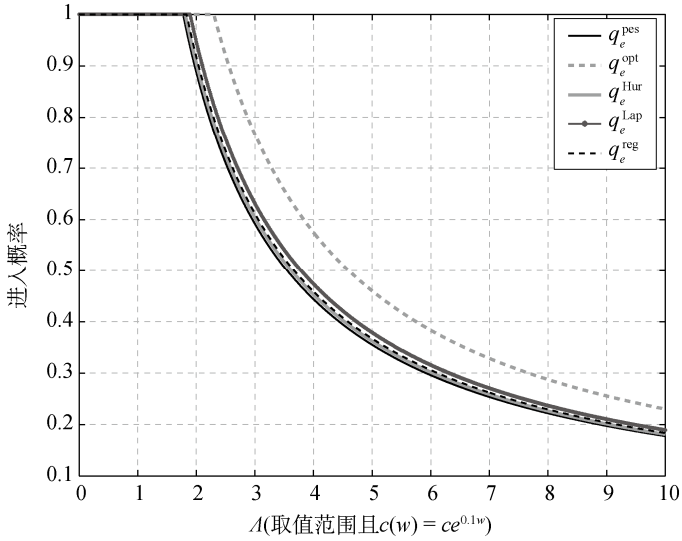

图 3-14　取值范围信息下的均衡进入概率[$c(w) = ce^{aw}$ 且 $a = 0.1$]

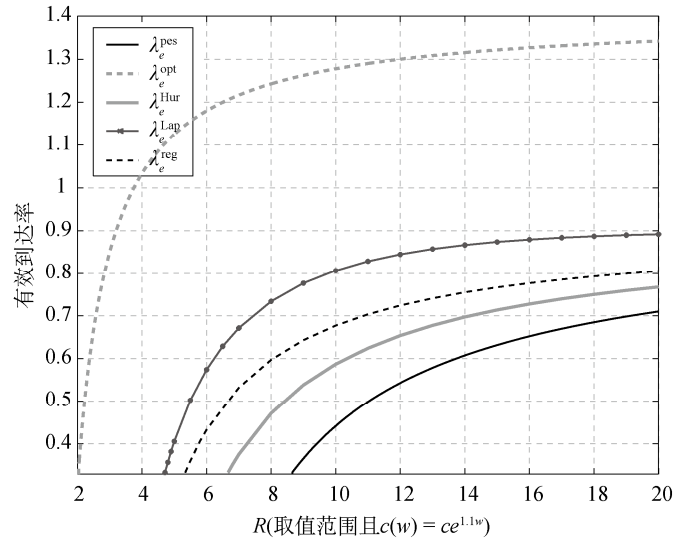

图 3-15　取值范围信息下的均衡有效到达率[$c(w) = ce^{aw}$ 且 $a = 1.1$]

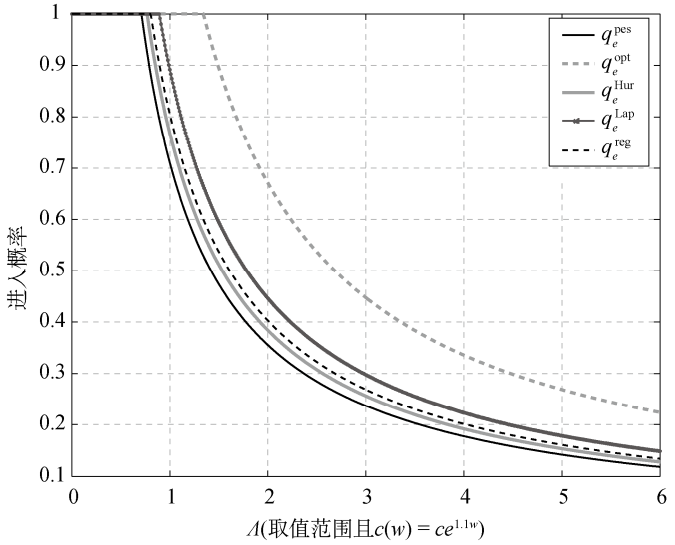

图 3-16　取值范围信息下的均衡进入概率[$c(w) = ce^{aw}$ 且 $a = 1.1$]

3.3　具有阈值多重休假策略的排队系统

本节将研究具有多重休假与 N 策略的 Markov 单服务员排队系统中顾客的均衡和最优止步策略，且分别讨论完全可视与完全不可视两种信息水平。

本节内容为作者研究成果[57]的进一步完善和补充。

3.3.1　模型描述

本节考虑一些具有多重休假和 N 策略的 Markov 单服务员排队系统。假定顾客的潜在到达率为 Λ，服务员在忙期的服务率为 μ。一旦系统空竭，服务员开始进行多重休假，单次休假 V 服从参数为 θ 的指数分布。当某次休假结束后，系统中的顾客数不少于 N 个，服务员进入忙期：否则，继续休假。

令 $(L(t), I(t))$表示时刻 t 时的系统状态，其中 $L(t)$表示时刻 t 时的系统队长，$I(t)$表示时刻 t 时的服务员状态，且

$$I(t)=\begin{cases}0, \text{服务员处于休假期}\\ 1, \text{服务员处于忙期}\end{cases}$$

因此，完全可视情形意味着到达顾客可同时观测到系统信息 $L(t)$和 $I(t)$，完全不可视情形意味着系统信息 $L(t)$和 $I(t)$均对顾客封闭。

在此，标记一个到达顾客，服务完成可获得服务收益 R，但也需承担单位等待损耗 C。我们采用线性等待损耗函数，因此平均剩余效用为 $U = R - CE[W]$，其中 $E[W]$表示他的平均逗留时间。显然，如果他离开，则 $U = 0$。假设服务顺序为先到先服务，不允许出现插队和中途退出队列行为。

3.3.2　完全可视排队系统

在完全可视排队系统中，定义标记顾客在观测到系统状态(n, i)后的逗留时间为 $W(n, i)$，均值为 $E[W(n, i)]$，服务完成后的平均剩余效用为 $U(n, i)$。因此，在服务员状态 i 下顾客的均衡止步阈值记为 $n_e(i)$，整体阈值策略记为$(n_e(0), n_e(1))$。另一方面，在服务员状态 i 下顾客的社会最优止步阈值记为 $n^*(i)$，整体阈值策略记为$(n^*(0), n^*(1))$。

3.3.2.1　纳什均衡

首先考虑完全可视情形下顾客的均衡止步行为，即时刻 t 到达的顾客可同时观测到 $L(t)$和 $I(t)$。对于顾客在观测到系统状态$(n, 1)$后的逗留时间 $W(n, 1)$，显然等于 $n+1$ 个顾客（包括他自己）的服务时间，即

$$E\left[W(n,1)\right]=\frac{n+1}{\mu} \tag{3-46}$$

均衡状态下，顾客服务完成后的平均剩余效用 $U_e(n, 1)$ 等于零，即 $U_e(n, 1) = R-CE[W(n, 1)] = 0$。求解 $U_e(n, 1) = 0$，得到

$$n_e(0)=\left\lfloor \frac{R\mu}{C} \right\rfloor -1 \tag{3-47}$$

可见 $n_e(1)$与 N 和 θ 并不相关。对于顾客在观测到系统状态$(n, 0)$后的逗留时间 $W(n, 0)$，我们需要讨论两种情况：$n < N-1$ 和 $n \geq N-1$。如果 $n \geq N-1$，$W(n, 0)$等于一个剩余休假时间和 $n+1$ 个服务时间之和，因此

$$E\left[W(n,0)\right]=\frac{1}{\theta}+\frac{n+1}{\mu} \tag{3-48}$$

均衡状态下，顾客服务完成后的平均剩余效用 $U_e(n, 0)$等于零，即 $U_e(n, 0) = R - CE[W(n, 0)] = 0$。求解 $U_e(n, 0) = 0$，得到

$$n_e(0)=\left\lfloor \left(\frac{R}{C}-\frac{1}{\theta}\right)\mu \right\rfloor -1 \tag{3-49}$$

为了保证服务员始终能被激活，显然需要保证条件 $n_e(0) \geq N-1$ 成立，而没有必要再考虑另一种情况 $n < N-1$。因此，$n_e(1) > n_e(0) \geq N-1$ 且 $n_e(0)$随 θ 递增。

为了得到均衡状态下所有顾客的单位社会福利，首先需要导出队长的平稳分布。定义整体止步阈值策略为$(n(0), n(1))$，状态转移图如图 3-17 所示，其中$N-1\leq n(0)\leq n(1)$。因此，该 Markov 过程$\{L(t), I(t)\}$的状态空间为

$$\Omega_{ob}^{e}=\{(n,0):0\leq n\leq n(0)+1\}\cup\{(n,1):1\leq n\leq n(1)+1\}$$

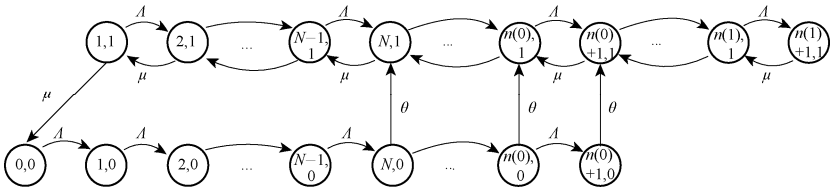

图 3-17　完全可视排队系统的状态转移图[$N-1\leq n(0)\leq n(1)$]

定义稳态队长分布为

$$\pi_{n,j}=P\{L=n,I=j\}=\lim_{t\to\infty}P\{L(t)=n,I(t)=j\},(n,j)\in\Omega_{ob}^{e},\ j=0,1$$

得到稳态状态转移概率方程为

$$\pi_{0,0}\Lambda=\pi_{1,1}\mu \tag{3-50}$$

$$\pi_{n,0}=\pi_{n-1,0},\ 1\leq n\leq N-1 \tag{3-51}$$

$$\pi_{n,0}(\Lambda+\theta)=\pi_{n-1,0}\Lambda,\ N\leq n\leq n(0) \tag{3-52}$$

$$\pi_{n(0)+1,0}\theta=\pi_{n(0),0}\Lambda \tag{3-53}$$

$$\pi_{1,1}(\Lambda+\mu)=\pi_{2,1}\mu \tag{3-54}$$

$$\pi_{n,1}(\Lambda+\mu)=\pi_{n-1,1}\Lambda+\pi_{n+1,1}\mu,\ 2\leq n\leq N-1\ \&\ n(0)+2\leq n\leq n(1) \tag{3-55}$$

$$\pi_{n,1}(\Lambda+\mu)=\pi_{n-1,1}\Lambda+\pi_{n+1,1}\mu+\pi_{n,0}\theta,\ N\leq n\leq n(0)+1 \tag{3-56}$$

$$\pi_{n(1)+1,1}\mu=\pi_{n(1),1}\Lambda \tag{3-57}$$

首先考虑概率$\{\pi_{n,0}\,|\,0\leq n\leq n(0)+1\}$。由方程（3-50）和方程（3-51），可以得到

$$\pi_{n,0}=\frac{\mu}{\Lambda}\pi_{1,1},\ 0\leq n\leq N-1 \tag{3-58}$$

再由方程（3-52）和方程（3-58），得到

$$\begin{aligned}\pi_{n,0}&=\frac{\Lambda}{\Lambda+\theta}\pi_{n-1,0}=\left(\frac{\Lambda}{\Lambda+\theta}\right)^{n-N+1}\pi_{N-1,0}\\&=\left(\frac{\Lambda}{\Lambda+\theta}\right)^{n-N+1}\frac{\mu}{\Lambda}\pi_{1,1},\ N\leq n\leq n(0)\end{aligned} \tag{3-59}$$

然后根据方程（3-53）和方程（3-59），导出

$$\pi_{n(0)+1,0}=\frac{\Lambda}{\theta}\left(\frac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\frac{\mu}{\Lambda}\pi_{1,1}=\frac{\mu}{\theta}\left(\frac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1} \tag{3-60}$$

接下来考虑概率$\{\pi_{n,1}\mid 1\leq n\leq N-1\}$。根据方程（3-55），可知上述概率是下列齐次线性差分方程

$$\mu x_{n+1}-(\Lambda+\mu)x_n+\Lambda x_{n-1}=0,\ 2\leq n\leq N-1 \tag{3-61}$$

的解。因此，它的特征方程为

$$\mu x^2-(\Lambda+\mu)x+\Lambda=0$$

且有两个解：1 和$\rho=\Lambda/\mu\,(\neq 1)$，所以方程（3-60）的通解为$x_n^{\text{hom}}=A_1+B_1\rho^n$，其中$A_1$、$B_1$是未知系数。由方程（3-53），得到

$$\begin{cases}A_1+B_1\rho=\pi_{1,1}\\(\Lambda+\mu)(A_1+B_1\rho)=\mu(A_1+B_1\rho^2)\end{cases}$$

从而解得

$$\begin{cases}A_1=\dfrac{\rho-1}{2\rho-1}\pi_{1,1}\\B_1=\dfrac{1}{2\rho-1}\pi_{1,1}\end{cases} \tag{3-62}$$

因此，

$$\pi_{n,1}=A_1+B_1\rho^n,\ 1\leq n\leq N-1 \tag{3-63}$$

然后再考虑概率$\{\pi_{n,1}\mid N\leq n\leq n(0)+1\}$。由方程（3-56）可知，以上概率是以下非齐次线性差分方程

$$\mu x_{n+1}-(\Lambda+\mu)x_n+\Lambda x_{n-1}=-\left(\frac{\Lambda}{\Lambda+\theta}\right)^{n-N+1}\frac{\mu\theta}{\Lambda}\pi_{1,1},\ N\leq n\leq n(0) \tag{3-64}$$

的解。由于方程（3-64）对应的齐次差分方程的通解为$x_n^{\text{hom}}=A_2+B_2\rho^n$，所以方程（3-64）的通解为$x_n^{\text{gen}}=x_n^{\text{hom}}+x_n^{\text{spec}}$，其中$x_n^{\text{spec}}$是一个特解，考虑形式为$x_n^{\text{spec}}=C_1(\Lambda/(\Lambda+\theta))^n$的一个特解，将其代入方程（3-64）中，得到

$$C_1=\frac{\mu\theta(\Lambda+\theta)^N}{\Lambda^N((\Lambda+\theta)^2-(\Lambda+\mu)(\Lambda+\theta)+\mu\Lambda)}\pi_{1,1} \tag{3-65}$$

因此，

$$x_n^{\text{gen}}=A_2+B_2\rho^n+C_1\left(\frac{\Lambda}{\Lambda+\theta}\right)^n,\ N\leq n\leq n(0)+1 \tag{3-66}$$

其中，A_2和B_2为未知系数。由方程（3-63），得到

$$\begin{cases} A_2 = A_1 - C_1\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{N-1}\dfrac{\Lambda(1-\rho)-\rho\theta}{(\Lambda+\theta)(1-\rho)} \\ B_2 = B_1 - C_1\left(\dfrac{\Lambda}{\rho(\Lambda+\theta)}\right)^{N-1}\dfrac{\theta}{(\Lambda+\theta)(1-\rho)} \end{cases} \tag{3-67}$$

所以由方程（3-66），得到

$$\pi_{n,1} = A_2 + B_2\rho^n + C_1\left(\frac{\Lambda}{\Lambda+\theta}\right)^n，\ N \le n \le n(0)+1 \tag{3-68}$$

特别地，基于方程（3-60）和方程（3-68），可求解出

$$\pi_{n(0)+2,1} = A_2 + B_2\rho^{n(0)+2} + C_1\left(\frac{\Lambda}{\Lambda+\theta}\right)^{n(0)+1}\left(1-\frac{\theta}{\mu}\right) - \left(\frac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1} \tag{3-69}$$

最后，考虑概率$\{\pi_{n,1} \mid n(0)+3 \le n \le n(1)+1\}$。在此情形下，方程（3-61）的通解为$x_n^{\text{hom}} = A_3 + B_3\rho^n$，其中$A_3$、$B_3$是未知系数。由方程（3-57）和方程（3-69），得到$A_3 = 0$和

$$B_3 = \frac{\pi_{n(0)+2,1}}{\rho^{n(0)+2}} \tag{3-70}$$

因此，

$$\pi_{n,1} = B_3\rho^n，\ n(0)+3 \le n \le n(1)+1 \tag{3-71}$$

总之，所有导出的稳态队长概率$\{\pi_{n,j} \mid (n,j) \in \Omega_{ob}^e\}$都与$\pi_{1,1}$相关。利用平衡方程，可得到如下定理中的结果。

定理 3.3.1 对于具有多重休假与N策略的完全可视 Markov 排队系统，状态空间为$\Omega_{ob}^e = \{(n,0): 0 \le n \le n(0)+1\} \cup \{(n,1): 1 \le n \le n(1)+1\}$，如果$N-1 \le n(0) \le n(1)$，则稳态队长分布$\{\pi_{n,j} \mid (n,j) \in \Omega_{ob}^e\}$为

$$\begin{cases} \pi_{n,0} = \dfrac{\mu}{\Lambda}\pi_{1,1}，\ 0 \le n \le N-1 \\ \pi_{n,0} = \left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n-N+1}\dfrac{\mu}{\Lambda}\pi_{1,1}，\ N \le n \le n(0) \\ \pi_{n(0)+1,0} = \dfrac{\mu}{\theta}\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1} \end{cases} \tag{3-72}$$

且

$$\begin{cases}\pi_{n,1}=A_1+B_1\rho^n, 1\leq n\leq N-1\\ \pi_{n,1}=A_2+B_2\rho^n+C_1\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^n,\ N\leq n\leq n(0)+1\\ \pi_{n(0)+2,1}=A_2+B_2\rho^{n(0)+2}+C_1\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(0)+1}\left(1-\dfrac{\theta}{\mu}\right)-\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1}\\ \pi_{n,1}=B_3\rho^n,\ n(0)+3\leq n\leq n(1)+1\end{cases} \tag{3-73}$$

其中 $\rho=\Lambda/\mu\neq 1$；$A_i(i=1,2)$，$B_j(j=1,2,3)$ 和 C_1 由方程（3-62）、方程（3-65）、方程（3-67）、方程（3-70）分别给出，且 $\pi_{1,1}$ 可由下列平衡方程

$$\sum_{(n,j)\in\Omega_{ob}^e}\pi_{n,j}=1 \tag{3-74}$$

解出。

基于图 3-17 和定理 3.3.1，顾客的止步状态为(n(0)+1, 0)和(n(1)+1, 1)。因此，如果 $N-1\leq n(0)\leq n(1)$，单位社会福利为

$$U_e(n(0),n(1))=\Lambda R(1-\pi_{n(0)+1,0}-\pi_{n(1)+1,1})-C\left(\sum_{n=1}^{n(0)+1}n\pi_{n,0}+\sum_{n=1}^{n(1)+1}n\pi_{n,1}\right) \tag{3-75}$$

显然，均衡社会福利为 $U_e(n_e(0),n_e(1))$。

3.3.2.2　社会最优

接下来，我们考虑完全可视排队系统中的顾客的社会最优止步行为。在社会最优条件下，除了以上讨论的情形 $N-1\leq n(0)\leq n(1)$ 之外，其他两种情形 $N-1\leq n(1)\leq n(0)$ 和 $n(1)<N-1\leq n(0)$ 也应该考虑。状态转移图分别如图 3-18 和图 3-19 所示。因此，这两种情形的状态空间相同，都为

$$\Omega_{ob}^{s_1}=\Omega_{ob}^{s_2}=\{(n,0):0\leq n\leq n(0)+1\}\cup\{(n,1):1\leq n\leq n(0)+1\}$$

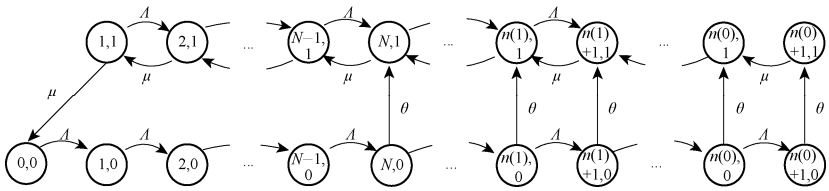

图 3-18　完全可视排队系统的状态转移图[$N-1\leq n(1)\leq n(0)$]

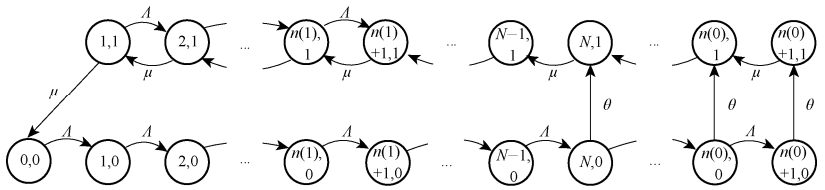

图 3-19　完全可视排队系统的状态转移图[$n(1)<N-1\leq n(0)$]

对于第二种情形 $N-1\leq n(1)\leq n(0)$，稳态转移概率方程为

$$\pi_{0,0}\Lambda=\pi_{1,1}\mu \tag{3-76}$$

$$\pi_{n,0}=\pi_{n-1,0}\ ,\ 1\leq n\leq N-1 \tag{3-77}$$

$$\pi_{n,0}(\Lambda+\theta)=\pi_{n-1,0}\Lambda\ ,\ N\leq n\leq n(0) \tag{3-78}$$

$$\pi_{n(0)+1,0}\theta=\pi_{n(0),0}\Lambda \tag{3-79}$$

$$\pi_{1,1}(\Lambda+\mu)=\pi_{2,1}\mu \tag{3-80}$$

$$\pi_{n,1}(\Lambda+\mu)=\pi_{n-1,1}\Lambda+\pi_{n+1,1}\mu,\ 2\leq n\leq N-1 \tag{3-81}$$

$$\pi_{n,1}(\Lambda+\mu)=\pi_{n-1,1}\Lambda+\pi_{n+1,1}\mu+\pi_{n,0}\theta,\ N\leq n\leq n(1) \tag{3-82}$$

$$\pi_{n(1)+1,1}\mu=\pi_{n(1),1}\Lambda+\pi_{n(1)+1,0}\theta+\pi_{n(1)+2,1}\mu \tag{3-83}$$

$$\pi_{n,1}\mu=\pi_{n,0}\theta+\pi_{n+1,1}\mu,\ n(1)+2\leq n\leq n(0) \tag{3-84}$$

$$\pi_{n(0)+1,1}\mu=\pi_{n(0)+1,0}\theta \tag{3-85}$$

应用与第一种情形类似的分析方法，可以得到第二种情形下的稳态队长分布 $\{\pi_{n,j}\mid(n,j)\in\Omega_{ob}^{s_1}\}$。

定理 3.3.2 对于具有多重休假与 N 策略的完全可视 Markov 排队系统，状态空间为 $\Omega_{ob}^{s_1}\{(n,0):0\leq n\leq n(0)+1\}\cup\{(n,1):1\leq n\leq n(0)+1\}$，如果 $N-1\leq n(1)\leq n(0)$，则稳态队长分布 $\{\pi_{n,j}\mid(n,j)\in\Omega_{ob}^{s_1}\}$ 为

$$\begin{cases}\pi_{n,0}=\dfrac{\mu}{\Lambda}\pi_{1,1}\ ,0\leq n\leq N-1\\ \pi_{n,0}=\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n-N+1}\dfrac{\mu}{\Lambda}\pi_{1,1}\ ,\ N\leq n\leq n(0)\\ \pi_{n(0)+1,0}=\dfrac{\mu}{\theta}\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1}\end{cases} \tag{3-86}$$

且

$$\begin{cases}\pi_{n,1}=A_1+B_1\rho^n,1\leq n\leq N-1\\ \pi_{n,1}=A_2+B_2\rho^n+C_1\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^n\ ,\ N\leq n\leq n(1)+1\\ \pi_{n,1}=A_2(1-\rho)+C_1\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(1)}\left(\dfrac{\Lambda}{\Lambda+\theta}-\rho\right)\\ \qquad+\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(1)-N}\left(\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n-n(1)}-\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^2-\dfrac{\theta}{\Lambda}\right)\pi_{1,1}\ ,n(1)+2\leq n\leq n(0)\\ \pi_{n(0)+1,1}=\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1}\end{cases} \tag{3-87}$$

其中 $\rho=\Lambda/\mu\neq 1$，A_i，B_i (i=1,2)和 C_1 由方程（3-62）、方程（3-65）、方程（3-67）分别给出，且 $\pi_{1,1}$ 可由下列平衡方程

$$\sum_{(n,j)\in\Omega_{ob}^{s_1}}\pi_{n,j}=1 \tag{3-88}$$

解出。

基于图 3-18 和定理 3.3.2，顾客的止步状态为$(n(0)+1,0)$和$\{(n,1)\mid n(1)+1\leq n\leq n(0)+1\}$。因此，如果 $N-1\leq n(1)\leq n(0)$，单位社会福利为

$$U_{s_1}(n(0),n(1))=\Lambda R\left(1-\pi_{n(0)+1,0}-\sum_{n=n(1)+1}^{n(0)+1}\pi_{n,1}\right)-C\left(\sum_{n=1}^{n(0)+1}0\pi_{n,0}+\sum_{n=1}^{n(1)+1}n\pi_{n,1}\right) \tag{3-89}$$

类似地，对于情形 $n(1)<N-1\leq n(0)$来说，稳态转移概率方程为

$$\pi_{0,0}\Lambda=\pi_{1,1}\mu \tag{3-90}$$

$$\pi_{n,0}=\pi_{n-1,0},1\leq n\leq N-1 \tag{3-91}$$

$$\pi_{n,0}(\Lambda+\theta)=\pi_{n-1,0}\Lambda,N\leq n\leq n(0) \tag{3-92}$$

$$\pi_{n(0)+1,0}\theta=\pi_{n(0),0}\Lambda \tag{3-93}$$

$$\pi_{1,1}(\Lambda+\mu)=\pi_{2,1}\mu \tag{3-94}$$

$$\pi_{n,1}(\Lambda+\mu)=\pi_{n-1,1}\Lambda+\pi_{n+1,1}\mu,2\leq n\leq n(1) \tag{3-95}$$

$$\pi_{n(1)+1,1}\mu=\pi_{n(1),1}\Lambda+\pi_{n(1)+2,1}\mu \tag{3-96}$$

$$\pi_{n,1}=\pi_{n+1,1},n(1)+2\leq n\leq N-1 \tag{3-97}$$

$$\pi_{n,1}\mu=\pi_{n,0}\theta+\pi_{n+1,1}\mu,\ N\leq n\leq n(0) \tag{3-98}$$

$$\pi_{n(0)+1,1}\mu=\pi_{n(0)+1,0}\theta \tag{3-99}$$

定理 3.3.3　对于具有多重休假与 N 策略的完全可视 Markov 排队系统，状态空间为 $\Omega_{ob}^{s_2}=\{(n,0):0\leq n\leq n(0)+1\}\bigcup\{(n,1):1\leq n\leq n(0)+1\}$，如果 $n(1)<N-1\leq n(0)$，则稳态队长分布 $\{\pi_{n,j}\mid(n,j)\in\Omega_{ob}^{s_2}\}$ 为

$$\begin{cases}\pi_{n,0}=\dfrac{\mu}{\Lambda}\pi_{1,1},0\leq n\leq N-1\\ \pi_{n,0}=\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n-N+1}\dfrac{\mu}{\Lambda}\pi_{1,1},\ N\leq n\leq n(0)\\ \pi_{n(0)+1,0}=\dfrac{\mu}{\theta}\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1}\end{cases} \tag{3-100}$$

且

$$\begin{cases}\pi_{n,1}=A_1+B_1\rho^n, 1\le n\le n(1)+1\\ \pi_{n,1}=A_1(1-\rho), n(1)+2\le n\le N\\ \pi_{n,1}=A_1(1-\rho)+\left(\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n-N}-1\right)\pi_{1,1},\ N+1\le n\le n(0)\\ \pi_{n(0)+1,1}=\left(\dfrac{\Lambda}{\Lambda+\theta}\right)^{n(0)-N+1}\pi_{1,1}\end{cases} \tag{3-101}$$

其中 $\rho=\Lambda/\mu\neq 1$，A_1、B_1 由方程（3-62）给出，且 $\pi_{1,1}$ 可由下列平衡方程

$$\sum_{(n,j)\in\Omega_{ob}^{S2}}\pi_{n,j}=1 \tag{3-102}$$

解出。

基于图 3-19 和定理 3.3.3，如果 $n(1)<N-1\le n(0)$，单位社会福利与式（3-89）形式相同。总之，定义单位社会福利为 $U_s(n(0), n(1))$，且

$$U_s(n(0),n(1))=\begin{cases}U_e(n(0),n(1)),\text{如果}N-1\le n(0)\le n(1)\\ U_{s_1}(n(0),n(1)),\text{如果}N-1\le n(1)\le n(0)\\ U_{s_2}(n(0),n(1)),\text{如果}n(1)<N-1\le n(0)\end{cases} \tag{3-103}$$

因此，最优社会福利为 $U_s(n^*(0), n^*(1))=\max\{U_s(n(0), n(1))\}$。

图 3-20 表明 $n^*(0)$和 $n^*(1)$均随 N 递增，这说明即使 N 增大，社会计划者也希望顾客能够积极踊跃地进入系统。显然，当 $N\le 10$ 时，$n^*(0)\le n^*(1)$，而当 $N>10$ 时，$n^*(0)>n^*(1)$。原因在于当 N 较小时，$U_s(n^*(0), n^*(1))=U_e(n^*(0), n^*(1)$，而随着 N 的增大，逐渐转变为 $U_s(n^*(0), n^*(1))=U_{s2}(n^*(0), n^*(1))$。而且，当 $N>5$ 时，$n^*(0)=N-1$，即当 $N>5$ 时，休假状态下的最优止步阈值降到最低，这个阈值只能保证服务员能够被再次激活。相反，图 3-21 表明 $n^*(0)$和 $n^*(1)$均随 θ 递减，说明顾客的个人行为偏离最优行为。而且，当 $\theta>4$ 时，它们的值几乎保持不变，尤其是 $n^*(0)=N-1$。因此，有趣的是，社会计划者不希望在休假期积压太多的顾客，他希望通过加速服务员状态转换的方法来减少顾客的等待损耗。

当 $N=3$ 和 $N=15$ 时，图 3-22 和图 3-23 分别比较了 $n_e(i)$和 $n^*(i)$ (i=0, 1)的大小。图中说明 $n_e(i)$和 $n^*(i)$均随 μ 递增，且显然 $n_e(i)$的增速要比 $n^*(i)$快。而且，$n_e(i)\ge n^*(i)$始终成立，因此顾客的均衡行为相比最优行为会使系统更拥塞。再者，图 3-24 和图 3-25 说明 $U_s(n^*(0), n^*(1))$随 N 递减但随 θ 递增。也就是说，社会计划者希望系统退化为一个经典的无休假排队系统，即 $N=1$ 和 $\theta=\infty$，尤其是当 Λ

较小时更是如此。由于顾客也希望系统无休假以最大化他的剩余效用，因此在这点上顾客和社会计划者的目标是一致的。

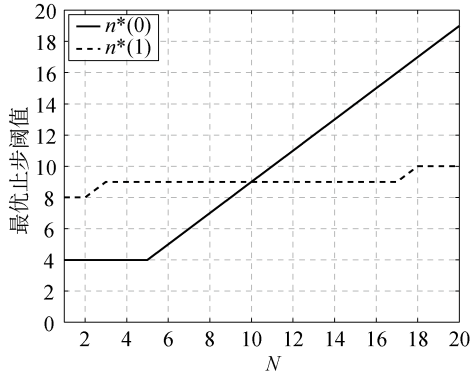

图 3-20　最优止步阈值关于 N 的敏感度（$R=10, C=1, \mu=3, \Lambda=5, \theta=0.1$）

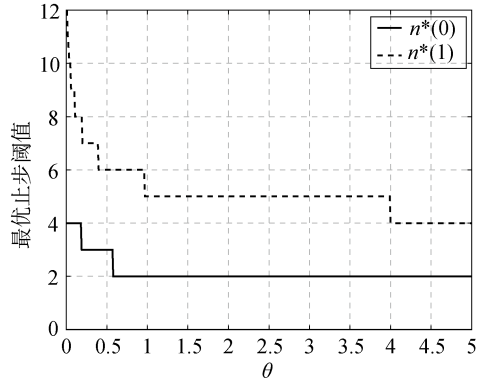

图 3-21　最优止步阈值关于 θ 的敏感度（$R=10, C=1, \mu=3, \Lambda=5, N=3$）

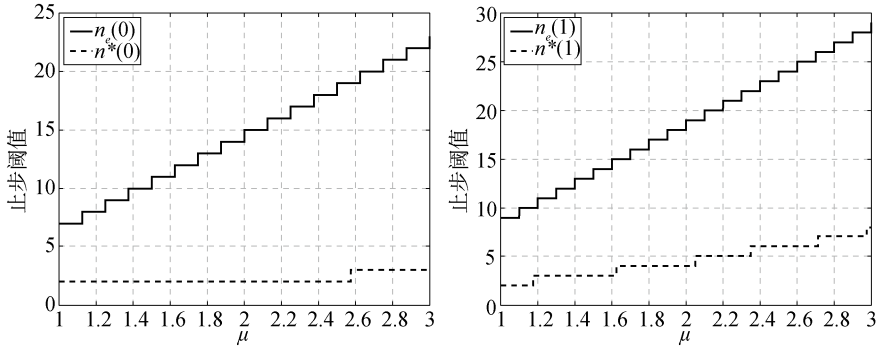

图 3-22　均衡和最优止步阈值（$R=10, C=1, \Lambda=4, \mu=0.5, N=3$）

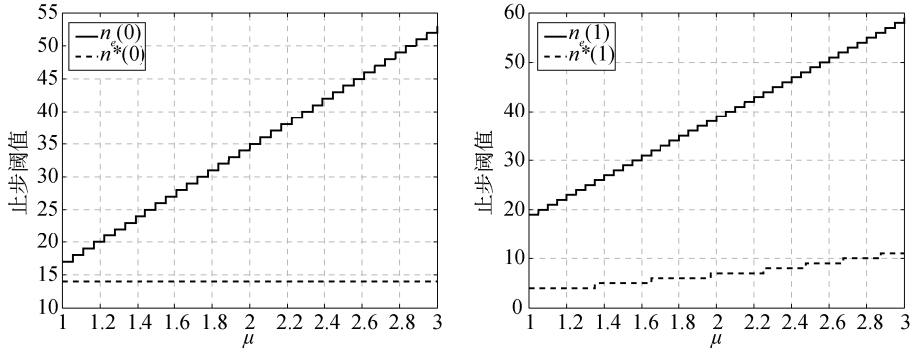

图 3-23　均衡和最优止步阈值（$R = 20$, $C = 1$, $\Lambda = 4$, $\mu = 0.5$, $N = 15$）

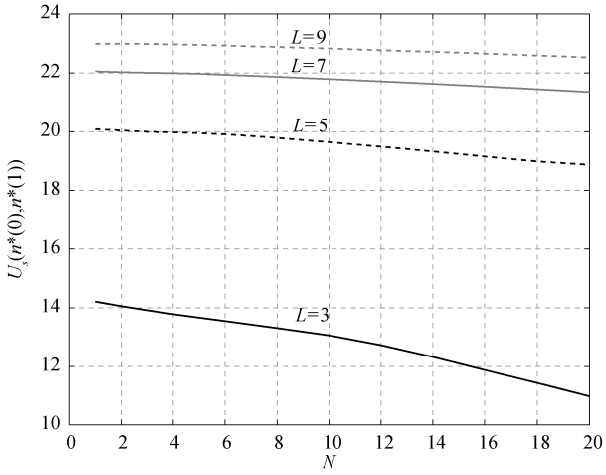

图 3-24　最优社会福利关于 N 的敏感度（$R = 10$, $C = 1$, $\mu = 3$, $\theta = 0.1$）

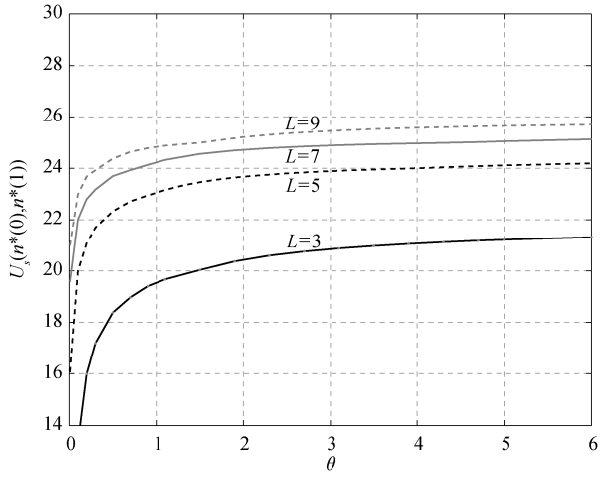

图 3-25　最优社会福利关于 θ 的敏感度（R=10, C=1, μ=3, N=3）

3.3.3　完全不可视排队系统

在完全不可视排队系统中，顾客的决策问题即为选择一个进入概率$q(0\leq q\leq 1)$，使得有效到达率为$\lambda=\Lambda q$。因此，顾客的均衡混合止步策略即为均衡进入概率q_e或均衡到达率$\lambda=\Lambda q_e$。另一方面，顾客的最优混合止步策略即为最优进入概率q^*或最优到达率$\lambda^*=\Lambda q^*$。状态转移图如图3-26所示。

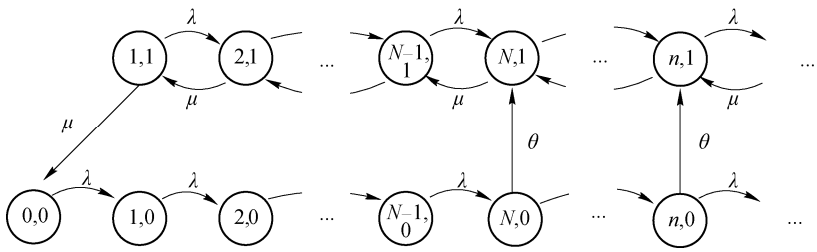

图 3-26　完全不可视排队系统的状态转移图

3.3.3.1　纳什均衡

首先讨论完全不可视情形下顾客的均衡止步行为，即时刻 t 到达的顾客观测不到系统状态 $I(t)$和 $L(t)$。为了得到 λ_e（或 q_e），需导出系统稳态队长分布。根据图 3-26，观察到过程$\{L(t), I(t)\}$为一个拟生灭过程，状态空间为

$$\Omega_{un}=\{(0,0)\}\cup\{(n,j):n\geq 1, j=0,1\}$$

如果 $\rho=\lambda/\mu<1$，令（L, I）为过程$\{L(t), I(t)\}$的稳态极限。记稳态队长分布为

$$\boldsymbol{\pi}=(\pi_{0,0},\ \boldsymbol{\pi}_1,\ \boldsymbol{\pi}_2\ldots),\ \boldsymbol{\pi}_n=(\pi_{n,0},\ \pi_{n,1}),\ n\geq 1$$

$$\pi_{n,j}=P\{L=n,\ I=j\}=\lim_{t\to\infty}P\{L(t)=n,\ I(t)=j\},\ (n,j)\in\Omega_{un},\ j=0,1$$

该过程的最小生成元为 $\boldsymbol{Q}$，得到稳态状态转移概率方程 $\boldsymbol{\pi Q}=\boldsymbol{0}$ 为

$$\pi_{0,0}\lambda=\pi_{1,1}\mu \tag{3-104}$$

$$\pi_{n,0}=\pi_{n-1,0},\ 1\leq n\leq N-1 \tag{3-105}$$

$$\pi_{n,0}(\lambda+\theta)=\pi_{n-1,0}\lambda,\ n\geq N \tag{3-106}$$

$$\pi_{1,1}(\lambda+\mu)=\pi_{2,1}\mu \tag{3-107}$$

$$\pi_{n,1}(\lambda+\mu)=\pi_{n-1,1}\lambda+\pi_{n+1,1}\mu,\ 2\leq n\leq N-1 \tag{3-108}$$

$$\pi_{n,1}(\lambda+\mu)=\pi_{n-1,1}\lambda+\pi_{n+1,1}\mu+\pi_{n,0}\theta,\ n\geq N \tag{3-109}$$

应用字典序方法，$\boldsymbol{Q}$ 可写成分块矩阵的形式

$$\boldsymbol{Q}=\begin{pmatrix} A_0 & \boldsymbol{C}_0 & & & & & & \\ \boldsymbol{B}_1 & \boldsymbol{A}_1 & \boldsymbol{C} & & & & & \\ & \boldsymbol{B} & \boldsymbol{A}_1 & \boldsymbol{C} & & & & \\ & & \boldsymbol{B} & \boldsymbol{A}_1 & \boldsymbol{C} & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & \boldsymbol{B} & \boldsymbol{A}_1 & \boldsymbol{C} & \\ & & & & & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} \\ & & & & & & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} \\ & & & & & & & \ddots & \ddots & \ddots \end{pmatrix} \tag{3-110}$$

其中

$$A_0=-\lambda,\quad \boldsymbol{B}_1=(0\ \ \mu)^{\mathrm{T}},\quad \boldsymbol{C}_0=(\lambda\ \ 0),\quad \boldsymbol{A}_1=\begin{pmatrix}-\lambda & 0\\ 0 & -(\lambda+\mu)\end{pmatrix}$$

$$\boldsymbol{B}=\begin{pmatrix}0 & 0\\ 0 & \mu\end{pmatrix},\qquad \boldsymbol{C}=\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix},\qquad \boldsymbol{A}=\begin{pmatrix}-(\lambda+\theta) & \theta\\ 0 & -(\lambda+\mu)\end{pmatrix}$$

求解以下二次方程

$$\boldsymbol{R}^2\boldsymbol{B}+\boldsymbol{R}\boldsymbol{A}+\boldsymbol{C}=0 \tag{3-111}$$

得到 R 的最小正根为

$$\boldsymbol{R}=\begin{pmatrix}\dfrac{\lambda}{\lambda+\theta} & \rho\\ 0 & \rho\end{pmatrix}$$

应用矩阵解析方法[58]，得到

$$\boldsymbol{\pi}_n=\boldsymbol{\pi}_N\boldsymbol{R}^{n-N},\quad n\geq N \tag{3-112}$$

向量（$\pi_{0,0}, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \cdots, \boldsymbol{\pi}_{N-1}, \boldsymbol{\pi}_N$ ）满足

$$(\pi_{0,0},\boldsymbol{\pi}_1,\boldsymbol{\pi}_2,\cdots,\boldsymbol{\pi}_{N-1},\boldsymbol{\pi}_N)\boldsymbol{B}[\boldsymbol{R}]=0 \tag{3-113}$$

其中

$$\boldsymbol{B}[\boldsymbol{R}]=\begin{pmatrix} A_0 & \boldsymbol{C}_0 & & & & & \\ \boldsymbol{B}_1 & \boldsymbol{A}_1 & \boldsymbol{C} & & & & \\ & \boldsymbol{B} & \boldsymbol{A}_1 & \boldsymbol{C} & & & \\ & & \boldsymbol{B} & \boldsymbol{A}_1 & \boldsymbol{C} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \boldsymbol{B} & \boldsymbol{A}_1 & \boldsymbol{C} \\ & & & & & \boldsymbol{B} & \boldsymbol{R}\boldsymbol{B}+\boldsymbol{A} \end{pmatrix} \tag{3-114}$$

将式（3-114）代入式（3-113），得到

$$\begin{cases}\pi_{0,0}\lambda=\pi_{1,1}\mu\\ \pi_{n,0}=\pi_{n-1,0},1\leq n\leq N-1\\ \pi_{N,0}(\lambda+\theta)=\pi_{N-1,0}\lambda\\ \pi_{1,1}(\lambda+\mu)=\pi_{2,1}\mu\\ \pi_{n,1}(\lambda+\mu)=\pi_{n-1,1}\lambda+\pi_{n+1,1}\mu,2\leq n\leq N-1\\ \pi_{N,1}\mu=\pi_{N-1,1}\lambda+\pi_{N,0}(\lambda+\theta)\end{cases}\tag{3-115}$$

求解方程（3-112）和方程（3-115），得到

$$\begin{cases}\pi_{n,0}=\dfrac{\mu}{\lambda+\theta}\left(\dfrac{\lambda}{\lambda+\theta}\right)^{n-N}\pi_{1,1},\quad n\geq N\\ \pi_{n,1}=\left(\dfrac{\lambda\mu}{\lambda+\theta-\mu}\rho^{n-N-1}\left(1-\left(\dfrac{\mu}{\lambda+\theta}\right)^{n-N}\right)+\dfrac{\mu}{\mu-\lambda}(1-\rho^{N})\rho^{n-N}\right)\pi_{1,1},\ n\geq N\end{cases}\tag{3-116}$$

其中 $\pi_{1,1}$ 可由平衡方程

$$\pi_{0,0}+\sum_{n=1}^{N-1}\boldsymbol{\pi}_n\boldsymbol{e}+\boldsymbol{\pi}_N(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e}=1\tag{3-117}$$

求得，其中 $\boldsymbol{I}$ 为单位矩阵，$\boldsymbol{e}$ 为单位向量。因此，

$$\pi_{1,1}=\frac{\lambda\theta(\mu-\lambda)}{\mu^2(\lambda+N\theta)}\tag{3-118}$$

根据式（3-116）和式（3-118），系统无条件平均队长为

$$E\left[L(\lambda)\right]=\sum_{n=1}^{\infty}n(\pi_{n,0}+\pi_{n,1})=\frac{\lambda}{\theta}+\frac{\lambda}{\mu-\lambda}+\frac{N(N-1)\theta}{2(\lambda+N\theta)}\tag{3-119}$$

顾客无条件平均逗留时间为

$$E\left[W(\lambda)\right]=\frac{E\left[L(\lambda)\right]}{\lambda}=\frac{1}{\theta}+\frac{1}{\mu-\lambda}+\frac{N(N-1)\theta}{2\lambda(\lambda+N\theta)}\tag{3-120}$$

由于 $E[W(\lambda)]$ 关于 λ 的二阶导数为

$$E\left[W''(\lambda)\right]=\frac{2}{(\mu-\lambda)^3}+\frac{\theta N(N-1)}{\lambda^3(\lambda+N\theta)}+\frac{\theta N(N-1)}{\lambda^2(\lambda+N\theta)^2}+\frac{\theta N(N-1)}{\lambda(\lambda+N\theta)^3}\tag{3-121}$$

如果 $\rho=\lambda/\mu<1$，该二阶导数为正，显然，$E[W(\lambda)]$ 关于 λ 是严格凸的。因此，顾客的平均剩余效用为

$$U_e(\lambda)=R-CE\left[W(\lambda)\right]=R-C\left(\frac{1}{\theta}+\frac{1}{\mu-\lambda}+\frac{N(N-1)\theta}{2\lambda(\lambda+N\theta)}\right)\tag{3-122}$$

以下命题对均衡结果进行了总结。

命题 3.3.1 （1）如果 $R - CE[W(\tilde{\lambda})] < 0$，则不存在正均衡到达率；

（2）如果 $R-CE[W(\tilde{\lambda})] = 0$，则存在唯一一个均衡到达率 $\lambda_e = \tilde{\lambda}$ 当且仅当 $\tilde{\lambda} \leq \Lambda$；

（3）如果 $R-CE[W(\tilde{\lambda})] > 0$，且

- 如果 $\lambda_2 \leq \Lambda$，则存在两个正均衡到达率 $\lambda_e=\{\lambda_1, \lambda_2\}$；
- 如果 $\lambda_1 < \Lambda < \lambda_2$，则存在两个正均衡到达率 $\lambda_e=\{\lambda_1, \Lambda\}$；
- 如果 $\Lambda = \lambda_1$，则存在一个正均衡到达率 $\lambda_e = \Lambda$；
- 如果 $\lambda_1 > \Lambda$，则不存在正均衡到达率。

其中 $\tilde{\lambda}$ 是方程

$$\frac{1}{(\mu-\lambda)^2} - \frac{\theta N(N-1)}{2\lambda(\lambda+\theta N)}\left(\frac{1}{\lambda}+\frac{1}{\lambda+\theta N}\right) = 0 \tag{3-123}$$

的唯一解，λ_1 和 λ_2 $(0 \leq \lambda_1 \leq \lambda_2)$是方程

$$R - C\left(\frac{1}{\theta}+\frac{1}{\mu-\lambda}+\frac{N(N-1)\theta}{2\lambda(\lambda+N\theta)}\right) = 0 \tag{3-124}$$

的解。

3.3.3.2 社会最优

下面讨论完全不可视情形下顾客的最优止步行为。根据方程（3-119），可以得到单位社会福利为

$$U_s(\lambda) = \lambda R - CE\left[L(\lambda)\right] = \lambda R - C\left(\frac{\lambda}{\theta}+\frac{\lambda}{\mu-\lambda}+\frac{N(N-1)\theta}{2(\lambda+N\theta)}\right) \tag{3-125}$$

由于 $U_s(\lambda)$关于 λ 的二阶导数为

$$U_s''(\lambda) = -\frac{6}{(\mu-\lambda)^2} - \frac{6\lambda}{(\mu-\lambda)^3} - \frac{3\theta N(N-1)}{(\lambda+\theta N)^3} \tag{3-126}$$

如果 $\rho = \lambda/\mu < 1$，该二阶导数显然是负的，所以 $U_s(\lambda)$关于 λ 严格凹。因此，当 $\Lambda \geq \lambda^*$时，顾客的社会最优混合止步策略（$\lambda^* > 0$ 或 $q^* > 0$）是唯一的，且可通过求解方程

$$U_s'(\lambda) = R - C\left(\frac{1}{\theta}+\frac{\mu}{(\mu-\lambda)^2}-\frac{\theta N(N-1)}{2(\lambda+\theta N)^2}\right) = 0 \tag{3-127}$$

得到。否则，当 $\Lambda < \lambda^*$时，$\lambda^* = \Lambda$。

鉴于均衡到达率 λ_1 并不稳定，我们只考虑稳定的均衡到达率 $\tilde{\lambda}$、λ_2 或 Λ。图 3-27 表明 $\lambda_e \geq \lambda^*$，且观察到当 $\Lambda > \lambda_1$ 时，N 的取值对 λ_e 和 λ^*并没有显著影响，而 λ_1 随 N 递增。因此，除非 Λ 很大，否则较高的 N 值将致使顾客选择止步离开。图 3-28 说明 N 越大，顾客越不愿意选择进入系统，而社会计划者为了能激活服

务员，希望顾客能更踊跃地加入队列。然而，图 3-29 说明顾客和社会计划者都希望缩短平均休假时间以减少顾客的等待损耗。

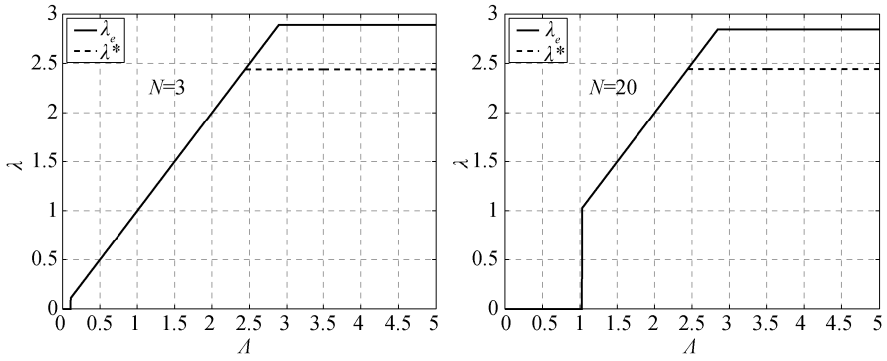

图 3-27　完全不可视情形下的均衡和最优到达率（$R = 10, C = 1, \theta = 2, \mu = 3$）

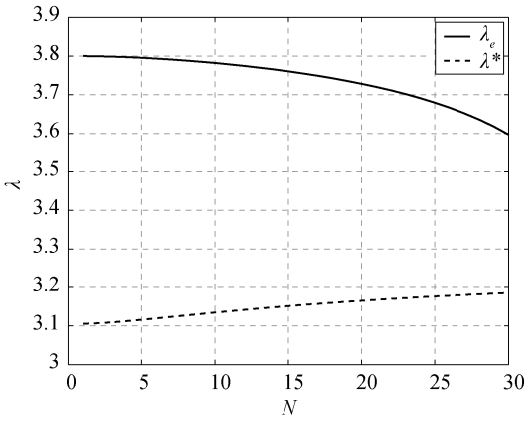

图 3-28　均衡和最优到达率关于 N 的敏感度
（$R = 10, C = 1, \theta = 0.2, \Lambda = 5, \mu = 4$）

当所有顾客均选择均衡混合止步策略 λ_e 时，他们的均衡社会福利可记为 $U_s(\lambda_e)$。另一方面，他们的最优社会福利可记为 $U_s(\lambda^*)$。类似于图 3-24 和图 3-25，图 3-30 和图 3-31 也说明社会计划者希望系统退化为无休假系统。当然，无论是完全可视系统还是完全不可视系统，如果考虑系统运行费用或状态转换费用的话，社会计划者的偏好也许会改变。

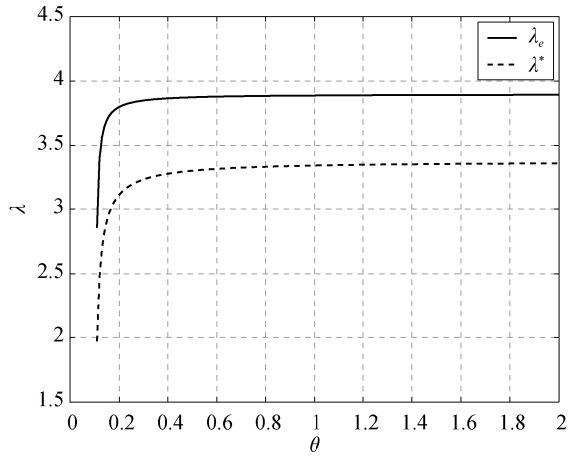

图 3-29　均衡和最优到达率关于 θ 的敏感度（$R=10, C=1, N=3, \Lambda=5, \mu=4$）

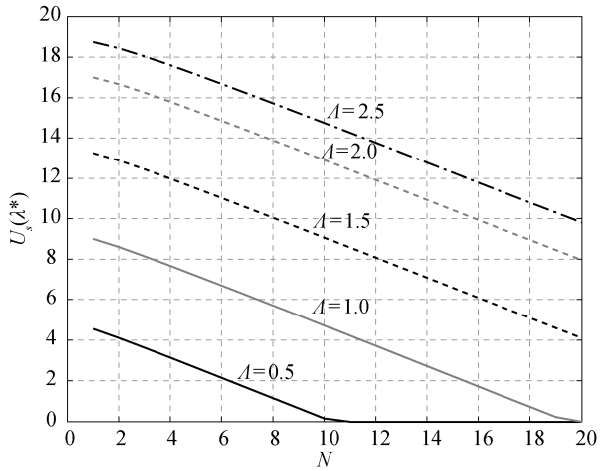

图 3-30　最优社会福利关于 N 的敏感度（$R=10, C=1, \theta=2, \mu=3$）

图 3-32 比较了完全可视和完全不可视情形下顾客的均衡社会福利。对于完全可视情形，表明当 Λ 很小时，顾客的均衡策略是离开，且一旦系统空竭，服务员将不能再被激活。然后，随着 Λ 的增大，$U_s(n_e(0), n_e(1))$迅速增大达到峰值。这表明先到顾客的进入行为对后到顾客来说是有利的，尤其是在休假期且 N 较小时。然而，到达峰值后，$U_s(n_e(0), n_e(1))$开始减小并最终归于零。这是由于随着 Λ 的增大，顾客将面临更高的等待损耗。类似地，$U_s(\lambda_e)$也具有先增后减的变化趋势。通过比较，图 3-32 说明当 $\Lambda < T_{\Lambda e}$ 时，$U_s(\lambda_e) \geq U_s(n_e(0), n_e(1))$，而当 $\Lambda \geq T_{\Lambda e}$ 时，$U_s(\lambda_e) \leq U_s(n_e(0), n_e(1))$。这说明当 Λ 低于某个阈值时，系统管理者应封闭系

统信息，否则，应公开系统信息。

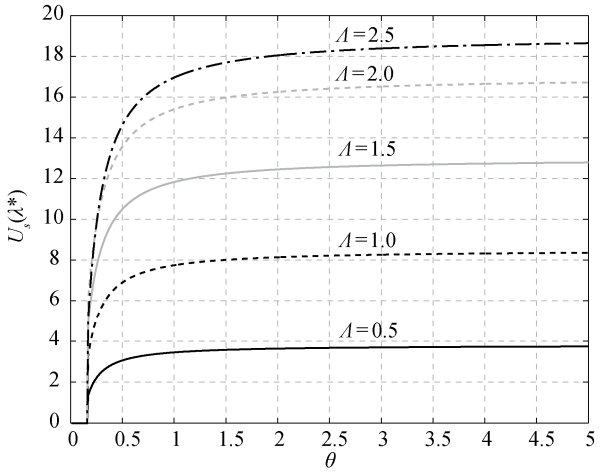

图 3-31　最优社会福利关于 θ 的敏感度（$R=10, C=1, N=3, \mu=3$）

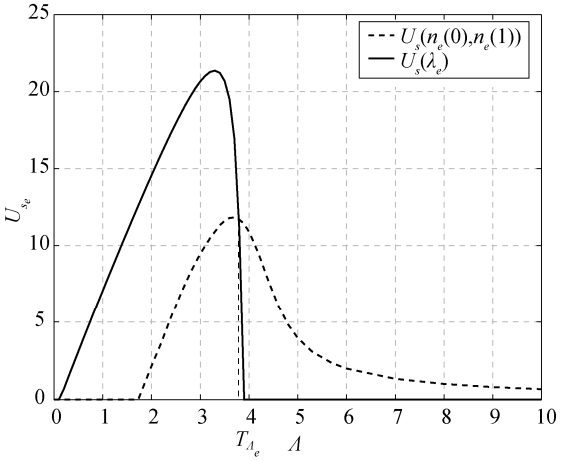

图 3-32　两类信息水平下的均衡社会福利（$R=10, C=1, \theta=0.5, N=3, \mu=4$）

相应地，图 3-33 比较了完全可视和完全不可视情形下顾客的最优社会福利。对于完全可视情形，图中表明 $U_s(n^*(0), n^*(1))$始终随 Λ 递增，但递增率趋缓。原因在于 $n^*(0)$和 $n^*(1)$随 Λ 递减，且当 Λ 足够大时，$n^*(0)=n^*(1)=N-1(=2)$，这使得所有顾客均选择止步离开的概率趋近于 1。因此，$U_s(n^*(0), n^*(1))$的递增率受到限制。对于完全不可视情形，显然，当 $\Lambda<\lambda^*$时，$U_s(\lambda^*)$ 先单增，然后当 $\Lambda\geq\lambda^*$时，再保持恒定。类似于图 3-32, 图 3-33 也表明当 $\Lambda < T_{\Lambda^*}$ 时, $U_s(\lambda^*)\geq U_s(n^*(0), n^*(1))$,

而当 $\Lambda \geq T_{\Lambda^*}$ 时，$U_s(\lambda^*) \leq U_s(n^*(0), n^*(1))$。这说明社会计划者应该和系统管理者采取同样的策略。当然，两个阈值 T_{Λ_e} 和 T_{Λ^*} 一般情况下并不一致。

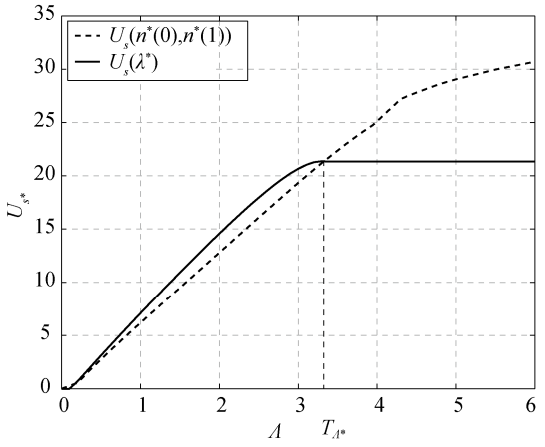

图 3-33　两类信息水平下的最优社会福利（$R = 10$, $C = 1$, $\theta = 0.5$, $N = 3$, $\mu = 4$）

3.4　相关文献评述

对于具有经典休假策略的排队系统中顾客的均衡行为, Burnetas 和 Economou[35] 首先研究了具有启动期的 Markov 单服务员排队系统中顾客的均衡止步策略。接下来，Economou 和 Kanta[59]讨论了具有不可靠服务员的队长可视排队系统中顾客的均衡止步策略。Adan[60]详细分析了具有多重休假策略的排队系统中顾客的同步退出行为。针对完全可视与几乎可视两种信息水平，刘维奇和马琰等[61]考虑了具有单重休假策略的排队系统中顾客的均衡止步行为，而王金亭和张峰[62]则研究了具有不可靠服务员与延迟修理的排队系统中顾客的均衡止步行为。随后，孙微等[56]研究了具有多重休假策略的队长不可视 Markov 排队系统，导出并比较具有五类决策标准与三类等待费用函数的同类顾客的均衡止步策略。对于离散时间排队系统，马琰和刘维奇等[63]研究了具有多重休假策略的 Geo/Geo/1 排队系统中顾客的均衡止步行为。

除了研究顾客的均衡止步行为之外, 许多文献也对顾客的最优止步行为进行了对比研究。比如，孙微等[43,44]讨论了具有三类启动关闭策略的排队系统，并研究了顾客的均衡和最优止步策略。Economou 等[64]则进一步讨论了具有一般服务时间和休假时间分布的完全不可视与几乎不可视排队系统中顾客的均衡和最优止步策略。

除了以上文献中涉及的经典休假策略之外，N 策略也是一类常用的休假策略[7,8]。然而，从经济学的角度研究具有 N 策略的排队系统的工作较少。针对空竭服务系统，Guo 和 Hassin[38,39]分别研究了具有同类顾客与异类顾客的完全可视与完全不可视 N 策略排队系统，得到了顾客的均衡和最优止步策略。而后，Guo 和 Li[65]又分析了与 Guo 和 Hassin[38]的研究中类似的系统，考虑了另两类信息水平，即几乎可视与几乎不可视情形。最近，孙微等[57]研究具有多重休假与 N 策略的完全可视与完全不可视排队系统中顾客的均衡和最优止步策略。另一方面，针对非空竭服务系统，Dimitrakopoulos 和 Burnetas[66]考虑了具有动态 N 策略的队长不可视 M/M/1 排队系统中顾客的均衡止步行为。

第 4 章　具有工作休假策略的排队系统

4.1　工作休假策略

随着电子商务和现代技术的迅猛发展，管理机构和通信网络的运行越来越复杂，对服务质量和系统性能要求也越来越高。随机服务系统理论作为解决系统理论分析和应用研究的有效工具，需要不断引入新的休假策略，建立新的排队模型。在此背景下，提出了工作休假策略，即休假期间，系统以较低速率继续为顾客提供服务，而并非完全停止服务。

与经典休假排队系统相比，工作休假排队系统具有更加复杂的结构，其建模和顾客的行为分析也更加困难。本章针对几类具有工作休假策略的排队系统，从经济学的角度对顾客的均衡和最优止步行为进行了研究。

本章内容为作者研究成果[67–70]的进一步完善和补充。

4.2　具有多重工作休假策略的排队系统

本节将分别对具有多重工作休假策略的完全可视、几乎不可视与完全不可视 Markov 单服务员排队系统进行研究，得到顾客的均衡和最优止步策略并进行比较，并对均衡和最优社会福利进行灵敏度分析。

4.2.1　模型描述

考虑具有多重工作休假策略的 Markov 单服务员排队系统。假设顾客潜在到达率为 λ，服务员正常服务率为 μ_b。一旦系统空竭，服务员开始一个工作休假期 V，且 V 服从参数为 θ 的指数分布。在休假期间，到达顾客以服务率 μ_v 接受服务且 $\mu_v<\mu_b$。休假结束后，如果系统中无顾客等待，则继续进行下一次休假，否则，服

务员将服务率由 μ_v 调整为 μ_b，开始一个正常忙期。记该排队系统为 M/M/1/MWV 排队系统。

令$(N_v(t), I(t))$表示时刻 t 时的系统状态，其中 $N_v(t)$表示时刻 t 的系统队长，$I(t)$表示时刻 t 的服务员状态，且

$$I(t)=\begin{cases}0, & \text{服务员处于工作休假期}\\ 1, & \text{服务员处于忙期}\end{cases}$$

因此，完全可视情形意味着到达顾客可同时观测到系统信息 $N_v(t)$和 $I(t)$，几乎不可视情形意味着到达顾客可观测到系统信息 $I(t)$而观测不到 $N_v(t)$，完全不可视情形意味着系统信息 $N_v(t)$和 $I(t)$均对顾客封闭。

在此，标记一个到达顾客，服务完成可获得服务收益 R，但也需承担单位等待损耗 C。我们采用线性等待损耗函数，即平均剩余效用 $U = R - CE[W]$，其中 $E[W]$表示顾客的平均逗留时间。假设到达过程、服务过程及工作休假时间相互独立，服务顺序为先到先服务。

4.2.2　完全可视排队系统

首先，考虑完全可视情形下顾客的均衡止步行为。由于时刻 t 到达的顾客可观测到系统状态 $N_v(t)$和 $I(t)$，所以当其观测到系统状态为$(n, i)(n\geq 0)$时，需要选择止步阈值 $n(i)(i = 1, 2)$，即整体止步策略$(n(0), n(1))$。记在服务员状态 i 进入的顾客的逗留时间为 $W(i)$，均值为 $E[W(i)]$，该顾客的平均剩余效用为 $U(i)$。因此，顾客在服务员状态 i 下的均衡止步阈值可记为 $n_e(i)$，整体止步策略记为$(n_e(0), n_e(1))$。

为了导出一标记顾客在系统状态$(n, 0)$下的平均逗留时间 $E[W(0)]$，我们做一些准备工作。记工作休假期的一个服务时间为 S_v，显然，其服从参数为 μ_v 的指数分布，记 $S_v^{(n)}$ 为 $n(n\geq 0, S_v^{(0)}\equiv 0)$ 个独立同分布的 S_v 之和。因此，它服从参数为 (n, μ_v)的 Erlang 分布，且密度函数为

$$f_n(x)=\frac{\mu_v(\mu_v x)^{n-1}}{(n-1)!}e^{-\mu_v x},\quad x\geq 0, n\geq 1$$

显然，$S_v^{(n)}$ 和 V 相互独立。接下来给出如下引理。

引理 4.2.1　如果 $S_v^{(n)} < V$，$S_v^{(n)}$ 的条件分布服从参数为$(n, \theta + \mu_v)$的 Erlang 分布。

证明　首先，导出事件 $\{S_v^{(n)} < V\}$ 发生的概率为

$$
\begin{aligned}
P\{S_v^{(n)} < V\} &= \int_0^{\infty} P\{u < V\} \frac{\mu_v(\mu_v u)^{n-1}}{(n-1)!} e^{-\mu_v u} \mathrm{d}u \\
&= \int_0^{\infty} \frac{\mu_v(\mu_v u)^{n-1}}{(n-1)!} e^{-(\theta+\mu_v)u} \mathrm{d}u \\
&= \left(\frac{\mu_v}{\theta+\mu_v}\right)^n \int_0^{\infty} \frac{(\theta+\mu_v)((\theta+\mu_v)u)^{n-1}}{(n-1)!} e^{-(\theta+\mu_v)u} \mathrm{d}u \\
&= \left(\frac{\mu_v}{\theta+\mu_v}\right)^n
\end{aligned}
\tag{4-1}
$$

然后可以导出 $S_v^{(n)}$ 的条件分布函数为

$$
\begin{aligned}
P\{S_v^{(n)} < x \mid S_v^{(n)} < V\} &= \frac{P\{S_v^{(n)} < x, S_v^{(n)} < V\}}{p\{S_v^{(n)} < V\}} \\
&= \left(\frac{\theta+\mu_v}{\mu_v}\right)^n \int_0^x P\{u < V\} \frac{\mu_v(\mu_v u)^{n-1}}{(n-1)!} e^{-\mu_v u} \mathrm{d}u \\
&= \int_0^x \frac{(\theta+\mu_v)((\theta+\mu_v)u)^{n-1}}{(n-1)!} e^{-(\theta+\mu_v)u} \mathrm{d}u
\end{aligned}
\tag{4-2}
$$

显然，被积函数是参数为 $(n, \theta+\mu_v)$ 的 Erlang 分布的密度函数。

引理 4.2.2 如果 $S_v^{(n)} < V < S_v^{(n+1)}$，$V$ 的条件分布服从参数为 $(n+1, \theta+\mu_v)$ 的 Erlang 分布。

证明 根据式（4-1），可导出事件 $\{S_v^{(n)} < V < S_v^{(n+1)}\}$ 发生的概率为

$$
P\{S_v^{(n)} < V < S_v^{(n+1)}\} = \left(\frac{\mu_v}{\theta+\mu_v}\right)^n - \left(\frac{\mu_v}{\theta+\mu_v}\right)^{n+1} = \left(\frac{\mu_v}{\theta+\mu_v}\right)^n \frac{\theta}{\theta+\mu_v} \tag{4-3}
$$

因此，V 的条件分布函数为

$$
\begin{aligned}
&P\{V < x \mid S_v^{(n)} < V < S_v^{(n+1)}\} \\
&= \frac{P\{V < x, S_v^{(n)} < V < S_v^{(n+1)}\}}{P\{S_v^{(n)} < V < S_v^{(n+1)}\}} \\
&= \left(\frac{\theta+\mu_v}{\mu_v}\right)^n \frac{\theta+\mu_v}{\theta} \int_0^x P\{S_v^{(n)} < u < S_v^{(n)} + S_v\} \theta e^{-\theta u} \mathrm{d}u \\
&= \frac{(\theta+\mu_v)^{n+1}}{\mu_v^n} \int_0^x \int_0^u \frac{\mu_v(\mu_v w)^{n-1}}{(n-1)!} e^{-\mu_v w} P\{S_v > u - w\} \mathrm{d}w e^{-\theta u} \mathrm{d}u \\
&= (\theta+\mu_v)^{n+1} \int_0^x e^{-(\theta+\mu_v)u} \mathrm{d}u \int_0^u \frac{w^{n-1}}{(n-1)!} \mathrm{d}w \\
&= \int_0^x \frac{(\theta+\mu_v)((\theta+\mu_v)u)^n}{n!} e^{-(\theta+\mu_v)u} \mathrm{d}u.
\end{aligned}
\tag{4-4}
$$

显然，被积函数是参数为 $(n+1,\theta+\mu_v)$ 的 Erlang 分布的密度函数。

基于以上引理中的结论，需要在状态$(n, 0)$时讨论两种情况：如果剩余休假时间，记为 V_R，足够服务完 n+1（包括标记顾客自己）个顾客，顾客的逗留时间 $W(0)$等于 n+1 个参数为 μ_v 的服务时间，否则，有 $j(0\leq j\leq n)$个顾客在 V_R 内完成服务，$W(0)$等于 V_R 与 $n+1-j$ 个参数为 μ_v 的服务时间之和。因此，得到 $W(0)$的 Laplace 变换为

$$\begin{aligned}\widetilde{W}_0^*(s)&=P\left\{S_v^{(n+1)}<V_R\right\}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{n+1}\\&\quad+\sum_{j=0}^{n}P\{S_v^{(j)}<V_R<S_v^{(j+1)}\}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{j+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1-j}\\&=\left(\frac{\mu_v}{\mu_v+\theta}\right)^{n+1}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{n+1}\\&\quad+\sum_{j=0}^{n}\left(\frac{\theta}{\mu_v+\theta}\right)\left(\frac{\mu_v}{\mu_v+\theta}\right)^{j}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{j+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1-j}\\&=\left(\frac{\mu_v}{\mu_v+\theta+s}\right)^{n+1}+\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1}\frac{\theta}{\mu_v+\theta+s}\sum_{j=0}^{n}\left(\frac{\mu_v(\mu_b+s)}{\mu_b(\mu_v+\theta+s)}\right)^{j}\\&=\left(\frac{\mu_v}{\mu_v+\theta+s}\right)^{n+1}+\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1}\frac{\mu_b\theta}{\mu_b\theta+(\mu_b-\mu_v)s}\\&\quad\times\left(1-\left(\frac{\mu_v(\mu_b+s)}{\mu_b(\mu_v+\theta+s)}\right)^{n+1}\right)\end{aligned}\tag{4-5}$$

从而得到 $E[W(0)]=-\widetilde{W}_0^{*\prime}(0)$，即

$$E[W(0)]=\frac{(n+1)\theta}{\theta\mu_b}+\frac{\mu_b-\mu_v}{\theta\mu_b}\left(1-\left(\frac{\mu_v}{\mu_v+\theta}\right)^{n+1}\right)\tag{4-6}$$

另一方面，在状态$(n, 1)$下标记顾客的平均逗留时间 $E[W(1)]$显然等于 n+1 个参数为 μ_b 的逗留时间之和，即

$$E[W(1)]=\frac{n+1}{\mu_b}\tag{4-7}$$

根据式（4-6）和式（4-7），容易得到 $U(0)$和 $U(1)$分别为

$$U(0)=R-CE[W(0)]=R-C\left(\frac{(n+1)\theta}{\theta\mu_b}+\frac{\mu_b-\mu_v}{\theta\mu_b}\left(1-\left(\frac{\mu_v}{\mu_v+\theta}\right)^{n+1}\right)\right)\tag{4-8}$$

$$U(1) = R - CE[W(1)] = R - \frac{C(n+1)}{\mu_b} \tag{4-9}$$

均衡状态下，顾客服务完成后的平均剩余效用为零，即 $U_e(i) = 0$。求解 $U_e(0) = 0$ 和 $U_e(1) = 0$，可得到唯一可行解 $n_e^*(0)$ 和 $n_e^*(1)$，且顾客的均衡阈值止步策略为 $(n_e(0), n_e(1)) = (\lfloor n_e^*(0) \rfloor, \lfloor n_e^*(1) \rfloor)$。图 4-1 表明 $n_e(0)$和 $n_e(1)$均随 μ_b 递增，然而 $n_e(1)$的增速比 $n_e(0)$快得多。

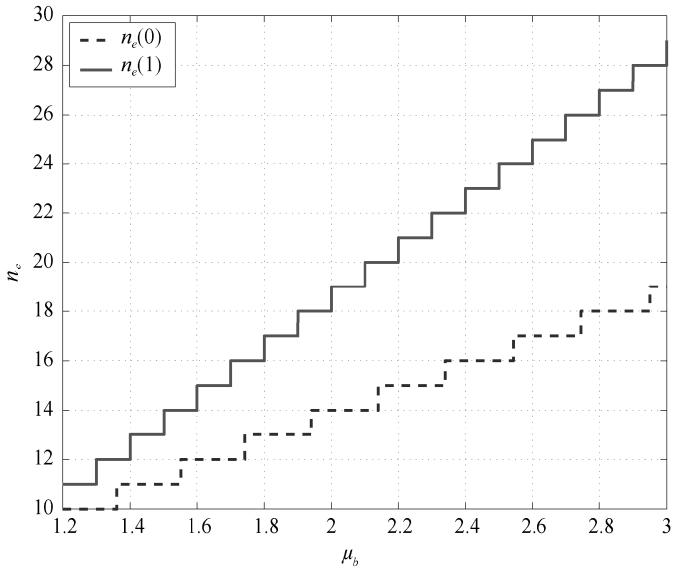

图 4-1　完全可视情形下的均衡止步阈值（$R = 10$, $C = 1$, $\mu_v = 1$, $\theta = 0.2$）

接下来，导出完全可视情形下系统的稳态队长分布，状态转移图如图 4-2 所示。因此，过程$\{N_v(t), I(t)\}$的状态空间为

$$\Omega_{ob} = \{(n,0) \mid 0 \le n \le n(0)+1\} \cup \{(n,1) \mid 1 \le n \le n(1)+1\}$$

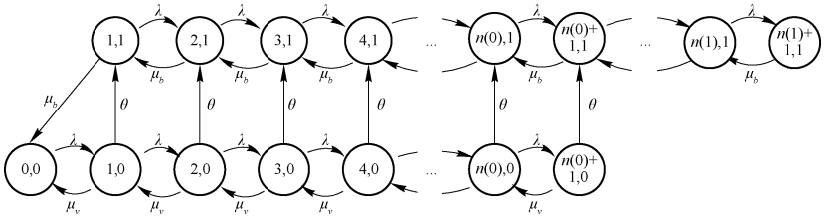

图 4-2　完全可视排队系统的状态转移图

定义稳态队长分布为

$$\pi_{kj} = P\{N_v = k, I = j\} = \lim_{t\to\infty} P\{N_v(t) = k, I(t) = j\}, \quad (k,j) \in \Omega_{ob}$$

状态转移概率方程为

$$\pi_{00}\lambda = \pi_{11}\mu_b + \pi_{10}\mu_v \tag{4-10}$$

$$\pi_{n0}(\lambda + \mu_v + \theta) = \pi_{n-10}\lambda + \pi_{n+10}\mu_v, \quad 1 \le n \le n(0) \tag{4-11}$$

$$\pi_{n(0)+10}(\theta + \mu_v) = \pi_{n(0)0}\lambda \tag{4-12}$$

$$\pi_{11}(\lambda + \mu_b) = \pi_{10}\theta + \pi_{21}\mu_b \tag{4-13}$$

$$\pi_{n1}(\lambda + \mu_b) = \pi_{n-11}\lambda + \pi_{n+11}\mu_b + \pi_{n0}\theta, \quad 2 \le n \le n(0)+1 \tag{4-14}$$

$$\pi_{n1}(\lambda + \mu_b) = \pi_{n-11}\lambda + \pi_{n+11}\mu_b, \quad n(0)+2 \le n \le n(1) \tag{4-15}$$

$$\pi_{n(1)+11}\mu_b = \pi_{n(1)1}\lambda \tag{4-16}$$

根据式(4-10)至式(4-12)，首先考虑概率 $\{\pi_{n0} \mid 0 \le n \le n(0)+1\}$。从式(4-11)中，观察到上述概率为以下齐次线性差分方程

$$\mu_v x_{n+1} - (\lambda + \mu_v + \theta)x_n + \lambda x_{n-1} = 0, \quad 1 \le n \le n(0) \tag{4-17}$$

的解。考虑相应的特征方程

$$\mu_v x^2 - (\lambda + \mu_v + \theta)x + \lambda = 0$$

显然，它有两个解

$$x_{1,2}^* = \frac{(\lambda + \mu_v + \theta) \pm \sqrt{(\lambda + \mu_v + \theta)^2 - 4\lambda\mu_v}}{2\mu_v} \tag{4-18}$$

差分方程（4-17）的通解为 $x_n^{hom} = A_1 x_1^{*n} + B_1 x_2^{*n}$（显然 $x_1^* \ne x_2^*$），其中 A_1、B_1 为未知系数。根据式（4-10）和式（4-12），得到关于 A_1 和 B_1 的方程

$$\begin{cases} A_1(\lambda - \mu_v x_1^*) + B_1(\lambda - \mu_v x_2^*) = \pi_{11}\mu_b \\ A_1(x_1^{*n(0)+1}(\theta + \mu_v) - x_1^{*n(0)}\lambda) + B_1(x_2^{*n(0)+1}(\theta + \mu_v) - x_2^{*n(0)}\lambda) = 0 \end{cases} \tag{4-19}$$

求解式（4-19），得到

$$\begin{cases} A_1 = \dfrac{(x_2^{*n(0)+1}(\theta + \mu_v) - x_2^{*n(0)}\lambda)\mu_b}{(\lambda - \mu_v x_1^*)(x_2^{*n(0)+1}(\theta + \mu_v) - x_2^{*n(0)}\lambda) - (\lambda - \mu_v x_2^*)(x_1^{*n(0)+1}(\theta + \mu_v) - x_1^{*n(0)}\lambda)}\pi_{11} \\ B_1 = \dfrac{(x_1^{*n(0)+1}(\theta + \mu_v) - x_1^{*n(0)}\lambda)\mu_b}{(\lambda - \mu_v x_2^*)(x_1^{*n(0)+1}(\theta + \mu_v) - x_1^{*n(0)}\lambda) - (\lambda - \mu_v x_1^*)(x_2^{*n(0)+1}(\theta + \mu_v) - x_2^{*n(0)}\lambda)}\pi_{11} \end{cases} \tag{4-20}$$

因此，

$$\pi_{n0} = A_1 x_1^{*n} + B_1 x_2^{*n}, \quad 0 \le n \le n(0)+1 \tag{4-21}$$

类似地，再考虑概率 $\{\pi_{n1} \mid 1 \le n \le n(0)+2\}$。由式（4-14）可知，它们是以下非齐次线性差分方程

$$\mu_b x_{n+1} - (\lambda + \mu_b)x_n + \lambda x_{n-1} = -\pi_{n0}\theta = -\theta(A_1 x_1^{*n} + B_1 x_2^{*n}), \quad 2 \le n \le n(0)+1 \tag{4-22}$$

的解。差分方程（4-22）的齐次部分的通解为 $x_n^{hom} = A_2 1^n + B_2 \rho^n$（假设 $\rho \ne 1$），其中 $\rho = \lambda / \mu_b$。因此，差分方程（4-22）的通解为 $x_n^{gen} = x_n^{hom} + x_n^{spec}$，其中 x_n^{spec} 为

式（4-22）的特解。由于式（4-22）的非齐次部分为参数 x_1^* 和 x_2^* 的几何形式，所以考虑一个形为 $x_n^{spec}=C_1x_1^{*n}+D_1x_2^{*n}$ 的特解，并将其代入式（4-22），得到

$$\begin{cases} C_1=\dfrac{\theta x_1^*A_1}{(x_1^*-1)(\lambda-\mu_b x_1^*)} \\ D_1=\dfrac{\theta x_2^*B_1}{(1-x_2^*)(\mu_b x_2^*-\lambda)} \end{cases} \tag{4-23}$$

因此，差分方程（4-22）的通解为

$$x_n^{gen}=A_2 1^n+B_2\rho^n+C_1x_1^{*n}+D_1x_2^{*n},\quad 1\le n\le n(0)+2 \tag{4-24}$$

其中 A_2, B_2 为未知系数。根据式（4-13），得到

$$\begin{cases} A_2+\rho B_2=\pi_{11}-(C_1x_1^*+D_1x_2^*) \\ (A_2+\rho^2B_2)\mu_b=\pi_{11}(\lambda+\mu_b)-(A_1x_1^*+B_1x_2^*)\theta-(C_1x_1^{*2}+D_1x_2^{*2})\mu_b \end{cases} \tag{4-25}$$

求解式（4-25），得到

$$\begin{cases} A_2=\dfrac{\mu_b\pi_{11}+\lambda(C_1x_1^*+D_1x_2^*)-\mu_b(C_1x_1^{*2}+D_1x_2^{*2})-(A_1x_1^*+B_1x_2^*)\theta}{\mu_b(1-\rho)} \\ B_2=\dfrac{\mu_b(C_1x_1^{*2}+D_1x_2^{*2}-C_1x_1^*-D_1x_2^*)-\pi_{11}\lambda+(A_1x_1^*+B_1x_2^*)\theta}{\lambda(1-\rho)} \end{cases} \tag{4-26}$$

因此，根据式（4-24），得到

$$\pi_{n1}=A_2+B_2\rho^n+C_1x_1^{*n}+D_1x_2^{*n},\quad 1\le n\le n(0)+2 \tag{4-27}$$

最后，导出概率 $\{\pi_{n1}\mid n(0)+1\le n\le n(1)+1\}$。由式(4-15)可知，它们是式(4-22)的齐次部分的解，即 $x_n^{hom}=A_3 1^n+B_3\rho^n$，其中 A_3、B_3 为未知系数。考虑式（4-27）和式（4-16），得到

$$\begin{cases} A_3+\rho^{n(0)+2}B_3=A_2+\rho^{n(0)+2}B_2+C_1x_1^{*n(0)+2}+D_1x_2^{*n(0)+2} \\ A_3+\rho^{n(1)+1}B_3-\rho(A_3+\rho^{n(1)}B_3)=0 \end{cases} \tag{4-28}$$

求解式（4-28），得到

$$\begin{cases} A_3=0 \\ B_3=\dfrac{A_2+\rho^{n(0)+2}B_2+C_1x_1^{*n(0)+2}+D_1x_2^{*n(0)+2}}{\rho^{n(0)+2}} \end{cases} \tag{4-29}$$

因此，

$$\pi_{n1}=B_3\rho^n,\quad n(0)+1\le n\le n(1)+1 \tag{4-30}$$

总之，所有稳态概率 $\{\pi_{kj}\mid(k,j)\in\Omega_{ob}\}$ 均与 π_{11} 有关。应用平衡方程，得到如下定理中的结果。

定理 4.2.1 对于具有多重工作休假策略的完全可视 Markov 排除系统，状

态空间为 $\Omega_{ob}=\{(n,0)\mid 0\le n\le n(0)+1\}\cup\{(n,1)\mid 1\le n\le n(1)+1\}$，其中$(n(0),\ n(1))$为顾客的阈值策略，则稳态队长分布 $\{\pi_{kj}\mid(k,j)\in\Omega_{ob}\}$ 为

$$\pi_{n0}=A_1x_1^{*n}+B_1x_2^{*n},\quad 0\le n\le n(0)+1 \tag{4-31}$$

$$\pi_{n1}=A_2+B_2\rho^n+C_1x_1^{*n}+D_1x_2^{*n},\quad 2\le n\le n(0)+1 \tag{4-32}$$

$$\pi_{n1}=B_3\rho^n,\quad n(0)+2\le n\le n(1)+1 \tag{4-33}$$

其中 $\rho=\lambda/\mu_b\ne 1$ 和 x_1^*、x_2^*、A_1、B_2、C_1、D_1、A_2、B_2、B_3 分别由式（4-18）、式（4-20）、式（4-23）、式（4-26）、式（4-29）给出，且 π_{11} 可由平衡方程

$$\sum_{(k,j)\in\Omega_{ob}}\pi_{kj}=1 \tag{4-34}$$

解出。

基于状态转移图 4-2 和定理 4.2.1 中的结论，到达顾客的止步状态为$(n(0)+1,0)$或$(n\ (1)+1,\ 1)$。因此，单位社会福利为

$$U_s(n(0),n(1))=\lambda R(1-\pi_{n(0)+10}-\pi_{n(1)+11})-C\left(\sum_{n=1}^{n(0)+1}n\pi_{n0}+\sum_{n=1}^{n(1)+1}n\pi_{n1}\right) \tag{4-35}$$

一方面，当所有顾客采用均衡阈值止步策略$(n_e(0),\ n_e(1))$时，单位社会福利可表示为 $U_s(n_e(0),\ n_e(1))$。从图 4-3 中我们观察到，$U_s(n_e(0),\ n_e(1))$随 λ 先增后减且在 λ=1.6 处达到峰值。原因在于，当 λ 较小时，系统很少出现拥挤的现象，所以顾客不需要等待太长时间就能完成服务，这使得社会福利不断增加。然而，当社会福利达到峰值后，若 λ 继续增大，顾客将面临越来越大的等待损耗，这将给社会福利带来负效应，使其不断减少。

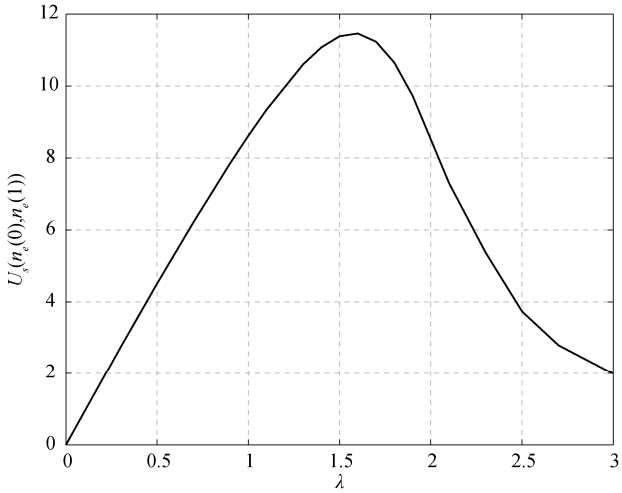

图 4-3　完全可视情形下的均衡社会福利（$R=10,\ C=1,\ \mu_b=2,\ \mu_v=1,\ \theta=1$）

另一方面, 对于社会计划者来说, 将设法使顾客采用社会最优阈值止步策略, 记为$(n^*(0), n^*(1))$, 以使社会福利最大化, 其中$(n^*(0), n^*(1))$可通过求解无约束整数规划 max $U_s(n_e(0), n_e(1))$得到。图 4-4 比较了$(n_e(0), n_e(1))$和$(n^*(0), n^*(1))$, 并且可观察到 $n_e(0)>n^*(0)$且 $n_e(1)>n^*(1)$, 说明顾客个人最优偏离社会最优。而且, 一般情况下, 均衡止步阈值随 μ_b 的增速比社会最优止步阈值的增速要快。

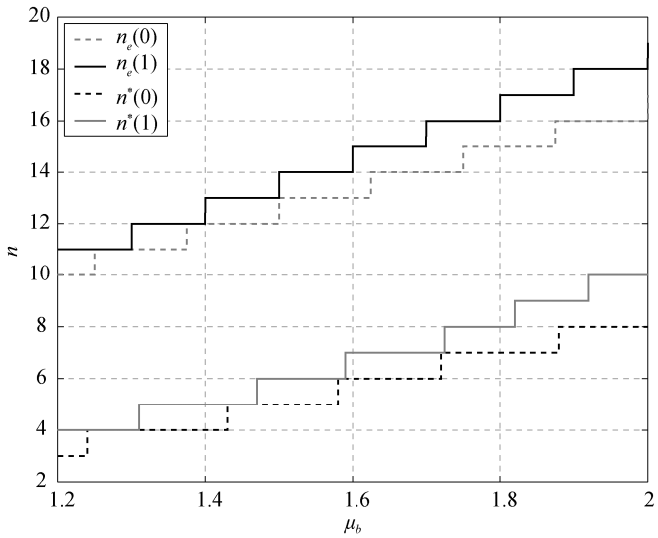

图 4-4　完全可视情形下的均衡和最优止步阈值（$R = 10, C = 1, \lambda = 1, \mu_v = 1, \theta = 0.5$）

4.2.3　几乎不可视排队系统

接下来, 考虑几乎不可视排队系统, 即时刻 t 到达的顾客只可观测到 $I(t)$而观测不到 $N_v(t)$。因此, 顾客的决策行为可表示为一对进入概率 $(q_0, q_1)(0 \le q_i \le 1, i = 0, 1)$, 且他们的均衡混合止步策略可记为$(q_e(0), q_e(1))$。

为了导出顾客的均衡混合止步策略$(q_e(0), q_e(1))$, 首先导出稳态队长分布, 状态转移图如图 4-5 所示。根据状态转移图, 观察到过程$\{N_v(t), I(t)\}$显然为一个拟生灭过程, 且状态空间为

$$\Omega_{po} = \{(0,0)\} \cup \{(k, j) : k \ge 1, j = 0, 1\}$$

如果 $\rho_b = \lambda q_1 / \mu_b < 1$, 令$(N_v, I)$为过程$\{N_v(t), I(t)\}$的稳态极限。记稳态队长分布为

$$\boldsymbol{\pi} = (\pi_0, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \cdots), \quad \pi_0 = \pi_{00}, \quad \boldsymbol{\pi}_k = (\pi_{k0}, \pi_{k1}), \quad k \ge 1$$

$$\pi_{kj} = P\{N_v = k, I = j\} = \lim_{t \to \infty} P\{N_v(t) = k, I(t) = j\}, \quad (k, j) \in \Omega_{po}$$

最小生成元为 $\boldsymbol{Q_v}$, 则状态转移概率方程 $\boldsymbol{\pi Q_v} = \boldsymbol{0}$ 为

$$\pi_{00}\lambda q_0 = \pi_{11}\mu_b + \pi_{10}\mu_v \tag{4-36}$$

$$\pi_{n0}(\lambda q_0 + \mu_v + \theta) = \pi_{n-10}\lambda q_0 + \pi_{n+10}\mu_v, \quad n \geq 1 \tag{4-37}$$

$$\pi_{11}(\lambda q_1 + \mu_b) = \pi_{10}\theta + \pi_{21}\mu_b \tag{4-38}$$

$$\pi_{n1}(\lambda q_1 + \mu_b) = \pi_{n-11}\lambda q_1 + \pi_{n+11}\mu_b + \pi_{n0}\theta, \quad n \geq 2 \tag{4-39}$$

应用字典序，$\boldsymbol{Q_v}$可写成分块矩阵的形式：

$$\boldsymbol{Q_v} = \begin{pmatrix} A_0 & \boldsymbol{C_0} & & & \\ \boldsymbol{B_1} & \boldsymbol{A} & \boldsymbol{C} & & \\ & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} & \\ & & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} \\ & & & \ddots & \ddots & \ddots \end{pmatrix} \tag{4-40}$$

其中

$$A_0 = -\lambda q_0, \quad \boldsymbol{C}_0 = (\lambda q_0 \ \ 0), \quad \boldsymbol{B}_1 = \begin{pmatrix} \mu_v \\ \mu_b \end{pmatrix}$$

$$\boldsymbol{B} = \begin{pmatrix} \mu_v & 0 \\ 0 & \mu_b \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} \lambda q_0 & 0 \\ 0 & \lambda q_1 \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} -(\lambda q_0 + \theta + \mu_v) & \theta \\ 0 & -(\lambda q_1 + \mu_b) \end{pmatrix}$$

现在需要求解以下矩阵二次方程

$$\boldsymbol{R}^2\boldsymbol{B} + \boldsymbol{R}\boldsymbol{A} + \boldsymbol{C} = \boldsymbol{0} \tag{4-41}$$

的最小非负解，即率阵 $\boldsymbol{R}$。

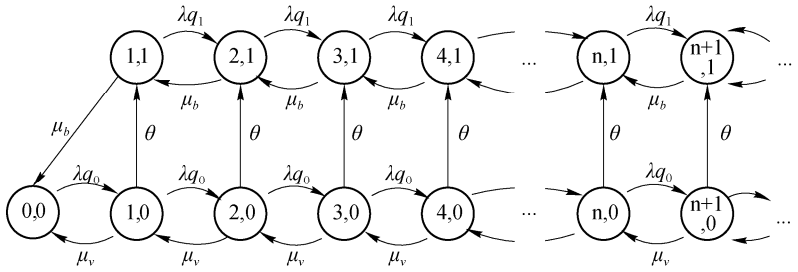

图 4-5　几乎不可视排队系统的状态转移图

引理 4.2.3　以下一元二次方程

$$\mu_v z^2 - (\lambda q_0 + \theta + \mu_v)z + \lambda q_0 = 0 \tag{4-42}$$

具有两个不同的根 $r_v(0 < r_v < 1)$ 和 $r_v^*(r_v^* > 1)$，且 r_v 满足

$$\mu_v + \frac{\theta}{1 - r_v} = \frac{\lambda q_0}{r_v} \text{①} \tag{4-43}$$

① 容易验证 $r_v = (\lambda q_0 + \theta + \mu_v - \sqrt{(\lambda q_0 + \theta + \mu_v)^2 - 4\mu_v\lambda q_0}) / 2\mu_v \in (0,1)$，将其代入式(4-42)得到式(4-43)。

引理 4.2.4 如果 $\rho_b = \lambda q_1 / \mu_b < 1$，则方程（4-41）具有最小非负解

$$\boldsymbol{R} = \begin{pmatrix} r_v & \dfrac{\theta r_v}{\mu_b(1-r_v)} \\ 0 & \rho_b \end{pmatrix} \tag{4-44}$$

证明 因为 $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$ 均为上三角矩阵，所以 $\boldsymbol{R}$ 一定也为上三角阵。假设

$$\boldsymbol{R} = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$$

将其代入式（4-41），得到方程组

$$\begin{cases} \mu_v r_{11}^2 - (\lambda q_0 + \theta + \mu_v) r_{11} + \lambda q_0 = 0 \\ \mu_b r_{22}^2 - (\lambda q_1 + \mu_b) r_{22} + \lambda q_1 = 0 \\ \mu_b r_{12}(r_{11} + r_{22}) + \theta r_{11} - (\lambda q_1 + \mu_b) r_{12} = 0 \end{cases}$$

根据引理 4.2.3 的结果，从第一个方程得到 $r_{11} = r_v$，从第二个方程得到 $r_{22} = \rho_b$。将 r_v 和 ρ_b 代入到第三个方程中，得到 r_{12}。

定理 4.2.2 对于具有多重工作休假策略的几乎不可视 Markov 排队系统，状态空间为 $\Omega_{po} = \{(0,0)\} \cup \{(k,j) : k \geq 1, j = 0,1\}$，如果顾客均服从策略$(q_0,\ q_1)$且 $\rho_b = \lambda q_1 / \mu_b < 1$，则稳态队长分布 $\{\pi_{kj} \mid (k,j) \in \Omega_{po}\}$ 为

$$\begin{cases} \pi_{k0} = K r_v^k, & k \geq 0 \\ \pi_{k1} = K \dfrac{\theta r_v}{\mu_b(1-r_v)} \displaystyle\sum_{j=0}^{k-1} r_v^j \rho_b^{k-1-j}, & k \geq 1 \end{cases} \tag{4-45}$$

其中

$$K = (1-r_v)(1-\rho_b)\left(1 - \rho_b + \frac{\theta r_v}{\mu_b(1-r_v)}\right)^{-1}$$

且 r_v 满足方程

$$\mu_v + \frac{\theta}{1-r_v} = \frac{\lambda q_0}{r_v}$$

证明 应用矩阵分析方法，得到

$$\boldsymbol{\pi}_k = (\pi_{k0}, \pi_{k1}) = (\pi_{10}, \pi_{11}) \boldsymbol{R}^{k-1}, \quad k \geq 1 \tag{4-46}$$

且$(\pi_{00}, \pi_{10}, \pi_{11})$满足方程

$$(\pi_{00}, \pi_{10}, \pi_{11}) B[\boldsymbol{R}] = \boldsymbol{0} \tag{4-47}$$

平衡条件为

$$\pi_{00} + (\pi_{10}, \pi_{11})(\boldsymbol{I} - \boldsymbol{R})^{-1} \boldsymbol{e} = 1 \tag{4-48}$$

其中 $\boldsymbol{I}$ 为单位阵，$\boldsymbol{e}$ 为单位向量，且

$$B[\boldsymbol{R}]=\begin{pmatrix}-\lambda q_0 & \lambda q_0 & 0\\ \mu_v & & \\ \mu_b & \boldsymbol{RB}+\boldsymbol{A} & \end{pmatrix}=\begin{pmatrix}-\lambda q_0 & \lambda q_0 & 0\\ \mu_v & -\lambda q_0-\theta-\mu_v(1-r_v) & \dfrac{\theta}{1-r_v}\\ \mu_b & 0 & -\mu_b\end{pmatrix} \tag{4-49}$$

将 $B[\boldsymbol{R}]$代入式（4-47），得到

$$\begin{cases}-\lambda q_0\pi_{00}+\mu_v\pi_{10}+\mu_b\pi_{11}=0\\ \lambda q_0\pi_{00}-(\lambda q_0+\theta+\mu_v(1-r_v))\pi_{10}=0\\ \dfrac{\theta}{1-r_v}\pi_{10}-\mu_b\pi_{11}=0\end{cases}$$

令 $\pi_{00}=K$，得到

$$(\pi_{00},\pi_{10},\pi_{11})=K\left(1,r_v,\frac{\theta r_v}{\mu_b(1-r_v)}\right)$$

而且，

$$R^{k-1}=\begin{pmatrix}r_v^{k-1} & \dfrac{\theta r_v}{\mu_b(1-r_v)}\displaystyle\sum_{j=0}^{k-2}r_v^j\rho_b^{k-2-j}\\ 0 & \rho_b^{k-1}\end{pmatrix},\quad k\geq 2$$

将 (π_{10},π_{11}) 和 $\boldsymbol{R}^{k-1}$ 代入式（4-46），得到式（4-45），其中 K 可由式（4-48）求得。

根据式（4-45），系统处于休假状态和忙期的概率 p_0 和 p_1 分别为

$$p_0=P\{I=0\}=\sum_{k=0}^{\infty}\pi_{k0}=(1-\rho_b)\left(1-\rho_b+\frac{\theta r_v}{\mu_b(1-r_v)}\right)^{-1}=\frac{K}{1-r_v} \tag{4-50}$$

$$p_1=P\{I=1\}=\sum_{k=1}^{\infty}\pi_{k1}=\frac{\theta r_v}{\mu_b(1-r_v)}\left(1-\rho_b+\frac{\theta r_v}{\mu_b(1-r_v)}\right)^{-1}=\frac{K\theta r_v}{\mu_b(1-\rho_b)(1-r_v)^2} \tag{4-51}$$

因此，可得到休假和忙期状态下的条件平均队长分别为

$$E[L_0]=\frac{\sum_{k=1}^{\infty}k\pi_{k0}}{\sum_{k=0}^{\infty}\pi_{k0}}=\frac{Kr_v}{(1-r_v)^2}\left(\frac{K}{1-r_v}\right)^{-1}=\frac{r_v}{1-r_v} \tag{4-52}$$

$$\begin{aligned}E[L_1]&=\frac{\sum_{k=1}^{\infty}k\pi_{k1}}{\sum_{k=1}^{\infty}\pi_{k1}}=\frac{K\theta r_v}{\mu_b(1-r_v)}\frac{1-r_v\rho_b}{(1-\rho_b)^2(1-r_v)^2}\left(\frac{K\theta r_v}{\mu_b(1-\rho_b)(1-r_v)^2}\right)^{-1}\\ &=\frac{1-r_v\rho_b}{(1-r_v)(1-\rho_b)}\end{aligned} \tag{4-53}$$

用 $E[L_1]$替换式（4-7）中的 n，可得到顾客在忙期状态下的条件平均逗留时间为

$$E[W_1]=\frac{2-\rho_b-r_v}{(1-r_v)(1-\rho_b)\mu_b} \tag{4-54}$$

对于顾客在休假状态下的条件平均逗留时间，记为 $E[W_0]$，我们可类似于完全可视情形，讨论在状态$(k, 0)(k\geq 0)$下的两种同样的情况。因此，可得到顾客在休假状态下逗留时间的 Laplace 变换

$$\begin{aligned}
W_0^*(s) &= \sum_{k=0}^{\infty}\pi_{k0}\left(P\{S_v^{(k+1)}<V\}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{k+1}\right.\\
&\quad\left.+\sum_{j=0}^{k}P\{S_v^{(j)}<V<S_v^{(j+1)}\}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{j+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{k+1-j}\right)\\
&= K\sum_{k=0}^{\infty}r_v^k\left(\frac{\mu_v}{\mu_v+\theta}\right)^{k+1}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{k+1}\\
&\quad+\frac{K\theta}{\mu_v+\theta}\sum_{k=0}^{\infty}r_v\sum_{j=0}^{k}\left(\frac{\mu_v}{\mu_v+\theta}\right)^{j}\left(\frac{\mu_v+\theta}{\mu_v+\theta+s}\right)^{j+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{k+1-j}\\
&= \frac{K\mu_v}{\mu_v+\theta+s}\sum_{k=0}^{\infty}\left(\frac{r_v\mu_v}{\mu_v+\theta+s}\right)^{k}\\
&\quad+\frac{K\theta\mu_b}{(\mu_b+s)(\mu_v+\theta+s)}\sum_{k=0}^{\infty}\left(\frac{r_v\mu_b}{\mu_b+s}\right)^{k}\sum_{j=0}^{k}\left(\frac{\mu_v(\mu_b+s)}{\mu_b(\mu_v+\theta+s)}\right)^{j}\\
&= \frac{K}{\mu_v(1-r_v)+\theta+s}\left(\mu_v+\frac{\theta\mu_b}{\mu_b(1-r_v)+s}\right)
\end{aligned}\tag{4-55}$$

因此，休假状态下条件逗留时间的 Laplace 变换为

$$\begin{aligned}
\tilde{W}_0^*(s) &= \frac{1}{p_0}W_0^*(s) = \frac{1-r_v}{K}\frac{K}{\mu_v(1-r_v)+\theta+s}\left(\mu_v+\frac{\theta\mu_b}{\mu_b(1-r_v)+s}\right)\\
&= \frac{1-r_v}{\mu_v(1-r_v)+\theta+s}\left(\mu_v+\frac{\theta\mu_b}{\mu_b(1-r_v)+s}\right)
\end{aligned}\tag{4-56}$$

从而得到 $E[W_0]=-\tilde{W}_0^{*\prime}(0)$，即

$$E[W_0]=\frac{2\mu_b\sqrt{(\lambda q_0+(\theta-\mu_v))^2+4\theta\mu_v}+2\mu_b(\mu_v-\lambda q_0-\theta)+4\theta\mu_v}{\mu_b\left(\left(\mu_v-\lambda q_0+\sqrt{(\lambda q_0+(\theta-\mu_v))^2+4\theta\mu_v}\right)^2-\theta^2\right)}\tag{4-57}$$

因此，根据式（4-57），得到休假状态下顾客的平均剩余效用为

$$U(0;q_0)=R-C\left(\frac{2\mu_b\sqrt{(\lambda q_0+(\theta-\mu_v))^2+4\theta\mu_v}+2\mu_b(\mu_v-\lambda q_0-\theta)+4\theta\mu_v}{\mu_b\left(\left(\mu_v-\lambda q_0+\sqrt{(\lambda q_0+(\theta-\mu_v))^2+4\theta\mu_v}\right)^2-\theta^2\right)}\right)\tag{4-58}$$

可见 $U(0;\ q_0)$并不依赖于 q_1。另一方面，根据式（4-54），得到忙期状态下顾客

的平均剩余效用为

$$U(0;q_0,q_1)=R-\frac{C(2-\rho_b-r_v)}{(1-r_v)(1-\rho_b)\mu_b} \tag{4-59}$$

求解 $U_e(0;q_0)=0$ 和 $U_e(1;q_0;q_1)=0$，得到唯一可行解 $(q_e^*(0),q_e^*(1))$。因此，顾客的均衡混合止步策略为 $(q_e(0),q_e(1))=(\min\{q_e^*(0),1\},\min\{q_e^*(1),1\})$。

虽然 $\mu_v<\mu_b$，图 4-6 表明 $q_e(0)$不一定小于 $q_e(1)$。这是因为休假状态下，当等待损耗较低且休假时间较短时，μ_v 能够满足顾客需求。然而，当等待损耗很高或休假时间很长时，顾客就很难接受休假期的低服务率了。

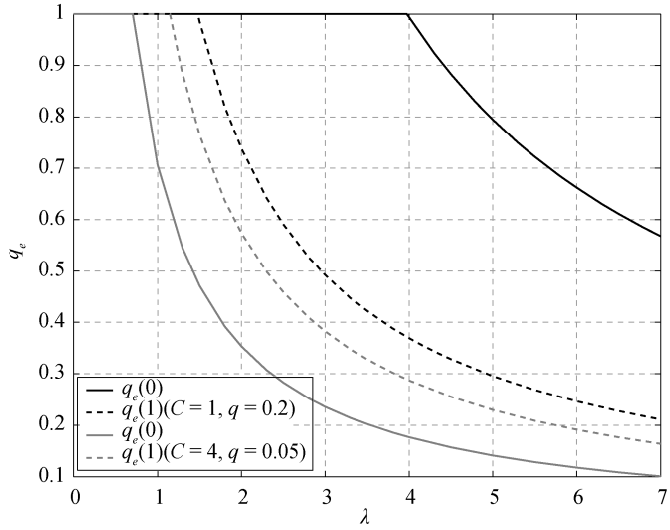

图 4-6　几乎不可视情形下的均衡混合止步策略（$R=10,\ \mu_b=2,\ \mu_v=1$）

然后，我们考虑均衡和最优社会福利。根据定理 4.2.2，条件平均队长为

$$E[L]=\sum_{k=0}^{\infty}k\pi_{k0}+\sum_{k=1}^{\infty}k\pi_{k1}=\frac{Kr_v}{1-r_v}\left(\frac{1}{1-r_v}+\frac{\theta(1-r_v\rho_b)}{\mu_b(1-\rho_b)^2(1-r_v)^2}\right) \tag{4-60}$$

因此，策略$(q_0\ ,q_1)$下的单位社会福利为

$$U_s(q_0,q_1)=\bar{\lambda}R-CE[L]=\lambda(p_0q_0+p_1q_1)R-\frac{CKr_v}{1-r_v}\left(\frac{1}{1-r_v}+\frac{\theta(1-r_v\rho_b)}{\mu_b(1-\rho_b)^2(1-r_v)^2}\right) \tag{4-61}$$

其中 $\bar{\lambda}=\lambda(p_0q_0+p_1q_1)$。

当所有顾客采用均衡混合止步策略$(q_e(0),\ q_e(1))$时，单位社会福利可记为 $U_s(q_e(0),\ q_e(1))$。图 4-7 表明，类似于图 4-3 中的 $U_s(n_e(0),\ n_e(1))$，$U_s(q_e(0),\ q_e(1))$也具有先增后减的变化趋势。

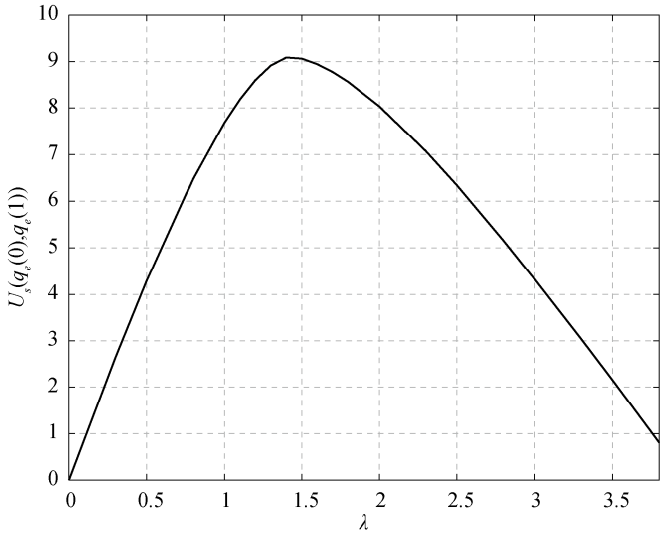

图 4-7　几乎不可视情形下的均衡社会福利（$R=10, C=1, \mu_b=2, \mu_v=1, \theta=1$）

然而，社会计划者希望顾客采用最优混合止步策略 $(q^*(0),q^*(1))$，且 $(q^*(0),q^*(1))$ 可通过 $U_s(q_0,q_1)$ 的一阶最优条件得到。图 4-8 比较了 $(q_e(0),q_e(1))$ 和 $(q^*(0),q^*(1))$，观察到 $q_e(0)>q^*(0)$ 且 $q_e(1)>q^*(1)$，说明在几乎不可视情形下，顾客个人最优同样也偏离社会最优。

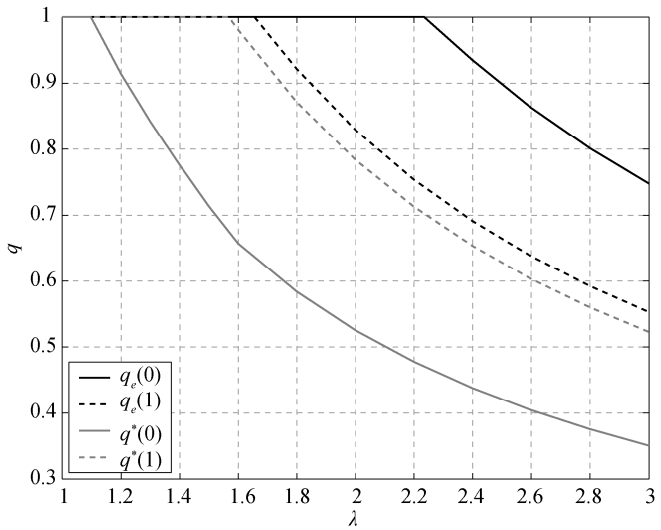

图 4-8　几乎不可视情形下的均衡和最优混合止步策略
（$R=10, C=1, \mu_b=2, \mu_v=1, \theta=0.1$）

4.2.4　完全不可视排队系统

最后，讨论完全不可视排队系统中顾客的均衡和最优止步行为，即时刻 t 到达的顾客观测不到系统状态 $I(t)$和 $N_v(t)$。因此，顾客的决策问题即为选择一个进入概率 $q(0\leq q\leq 1)$，且他们的均衡混合止步策略记为 q_e。状态转移图如图 4-9 所示。

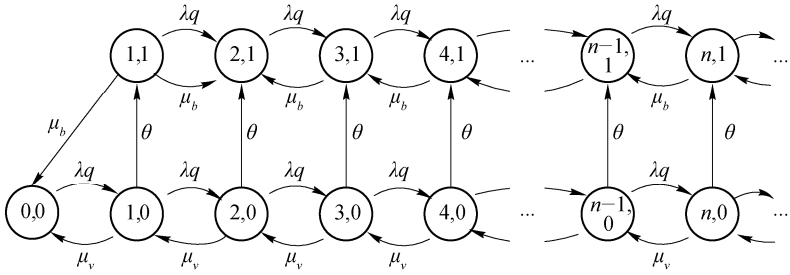

图 4-9　完全不可视排队系统的状态转移图

根据参考文献[71]中的结果，得到顾客的平均逗留时间为

$$E[W]=\frac{1}{\mu_b(1-\rho)}+\left(1-\frac{\mu_v}{\mu_b}\right)\left(1-r\frac{\mu_v}{\mu_b}\right)^{-1}\frac{r}{\lambda q(1-r)} \tag{4-62}$$

其中 $\rho=\lambda q/\mu_b<1$且 r 满足

$$\mu_v+\frac{\theta}{1-r}=\frac{\lambda q}{r} \tag{4-63}$$

因此，顾客的平均剩余效用为

$$U(q)=R-CE[W]=R-C\left(\frac{1}{\mu_b(1-\rho)}+\left(1-\frac{\mu_v}{\mu_b}\right)\left(1-r\frac{\mu_v}{\mu_b}\right)^{-1}\frac{r}{\lambda q(1-r)}\right) \tag{4-64}$$

求解 $U_e(q)=0$ 且考虑式（4-63），得到唯一可行解 q_e^*。因此，顾客的均衡混合止步策略为 $q_e=\min\{q_e^*,1\}$。

根据式（4-62），可得到策略 q 下的单位社会福利为

$$\begin{aligned}U_s(q)&=\lambda q\left(R-CE[W]\right)\\&=\lambda q\left(R-C\left(\frac{1}{\mu_b(1-\rho)}+\left(1-\frac{\mu_v}{\mu_b}\right)\left(1-r\frac{\mu_v}{\mu_b}\right)^{-1}\frac{r}{\lambda q(1-r)}\right)\right)\end{aligned} \tag{4-65}$$

当所有顾客采用均衡混合策略q_e时，单位社会福利可记为$U_s(q_e)$。类似于$U_s(n_e(0),n_e(1))$和$U_s(q_e(0),q_e(1))$，图 4-10 说明$U_s(q_e)$随λ也具有先增后减的变化趋势。

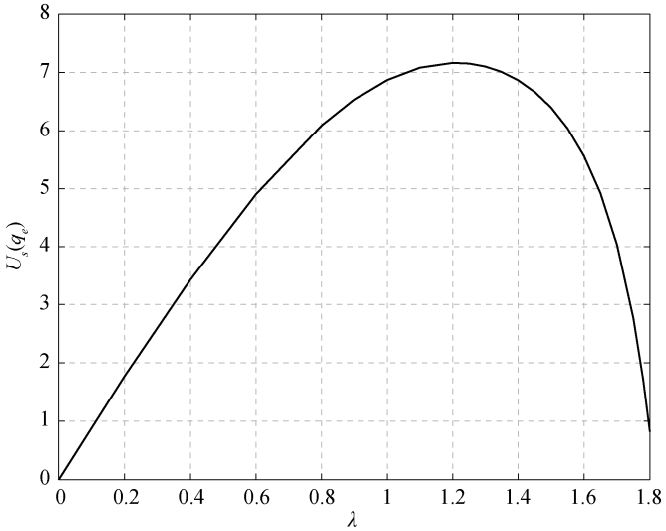

图 4-10　完全不可视情形下的均衡社会福利（$R = 10, C = 1, \mu_b = 2, \mu_v = 1, \theta = 0.1$）

从社会最优化的角度考虑，顾客的社会最优混合止步策略可记为 q^*，且 q^* 可通过 $U_s(q)$的一阶最优条件得到。类似于完全可视和几乎不可视情形，图 4-11 仍然表明 $q_e > q^*$。

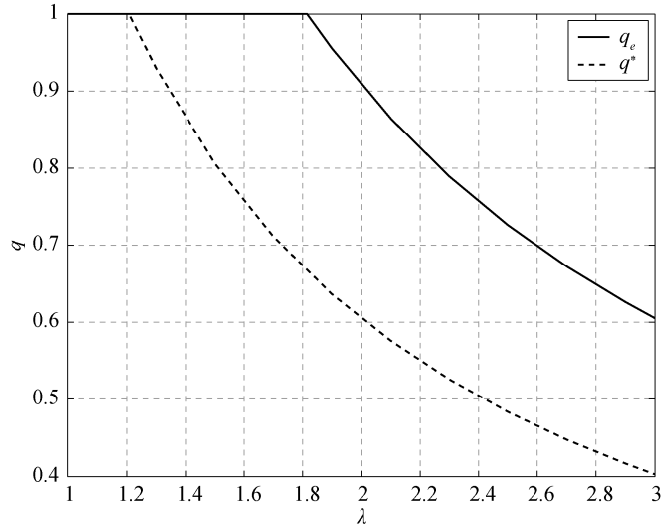

图 4-11　完全不可视情形下的均衡和最优混合止步策略
（$R = 10, C = 1, \mu_b = 2, \mu_v = 1, \theta = 0.1$）

4.3　具有双阶段工作休假策略的排队系统

考虑到实际问题中服务员既不进行单重休假也不进行多重休假，本节将对具有两阶段工作休假策略的完全可视、几乎不可视与完全不可视 Markov 排队系统进行研究，其中两阶段工作休假期间服务员的慢速服务率不同，分别得到并比较顾客的均衡止步策略及均衡和最优社会福利。

4.3.1　模型描述

考虑一些具有双阶段工作休假策略的 Markov 单服务员排队系统。假定顾客的潜在到达率为 λ，服务员的正常服务率为 μ_b。一旦系统空竭，服务员开始第一阶段工作休假 V_1，V_1 服从参数为 θ 的指数分布。在此期间，到达顾客以服务率 μ_{v_1} 接受服务，且 $\mu_{v_1} < \mu_b$。在第一阶段工作休假结束之后，无论系统中有无顾客，服务员将继续进行第二阶段工作休假 V_2，且 V_2 与 V_1 分布类型相同。在此期间，服务率由 μ_{v_1} 转换为 μ_{v_2}，且 $\mu_{v_2} < \mu_b$。当第二阶段工作休假结束后，如果有顾客等待，则开始一个新的忙期，否则，服务员处于闲期，直到有顾客到达为止。我们记该类排队系统为 M/M/1/(T-S)WV 排队系统。

令$(N(t), I(t))$表示时刻 t 时的系统状态，其中 $N(t)$表示时刻 t 时的系统队长，$I(t)$表示时刻 t 时的服务员状态，且

$$I(t)=\begin{cases}0, & \text{服务员处于忙期或闲期}\\ 1, & \text{服务员处于第一阶段工作休假期间}\\ 2, & \text{服务员处于第二阶段工作休假期间}\end{cases}$$

在此，标记一个到达顾客，服务完成可获得服务收益 R，但也需承担单位等待损耗 C。我们采用线性等待损耗函数，因此平均剩余效用 $U = R - CE[W]$，其中 $E[W]$表示顾客的平均逗留时间。假设到达过程、服务过程以及两阶段休假时间相互独立，服务顺序为先到先服务。

4.3.2　完全可视排队系统

首先，考虑完全可视情形下顾客的均衡止步行为，即时刻 t 到达的顾客可同时观测到 $N(t)$和 $I(t)$。定义标记顾客在观测到系统状态$(n, i)(i=0,1,2)$后的逗留时间为 $W_i(n)$，均值为 $E[W_i(n)]$，服务完成后的平均剩余效用为 $U_i(n)$。因此，在服务

员状态 i 下顾客的均衡止步阈值记为 $n_e(i)$，现在我们需要找到顾客的均衡整体止步策略$(n_e(0), n_e(1), n_e(2))$。

对于 $E[W_0(n)]$，显然它等于 n+1 个参数为 μ_b 的服务时间之和，即

$$E\left[W_0(n)\right]=\frac{n+1}{\mu_b} \tag{4-66}$$

另一方面，为了导出 $E[W_i(n)](i=1,2)$，需要讨论在状态$(n,1)$下的三种可能情况：第一阶段的剩余工作休假时间 V_{1R} 内足够服务完 n+1 个顾客，则标记顾客的逗留时间 $W_1(n)$等于 n+1 个参数为 μ_{v_1} 的服务时间之和；如果只有 $j(0\leq j\leq n)$个顾客在 V_{1R} 期间完成服务，而其他 n+1−j 个顾客在第二阶段工作休假期间完成服务，则 $W_1(n)$等于 V_{1R} 与 n+1−j 个参数为 μ_{v_2} 的服务时间之和；否则，如果有 $j(0\leq j\leq n)$个顾客在 V_{1R} 期间完成服务，有 $i(0\leq i\leq n-j)$个顾客在 V_2 期间完成服务，其他 n+1−i−j 个顾客在下个忙期完成服务，则 $W_1(n)$等于 V_{1R}、V_2 和 n+1−i−j 个参数为 μ_b 的服务时间之和。因此，得到 $W_1(n)$的 Laplace 变换为

$$\begin{aligned}
\widetilde{W}_1^*(s)=&P\left\{S_{v_1}^{(n+1)}<V_{1R}\right\}\left(\frac{\mu_{v_1}+\theta}{\mu_{v_1}+\theta+s}\right)^{n+1}+\sum_{j=0}^{n}P\left\{S_{v_1}^{(j)}<V_{1R}<S_{v_1}^{(j+1)}\cap S_{v_2}^{(n+1-j)}<V_2\right\}\\
&\times\left(\frac{\mu_{v_1}+\theta}{\mu_{v_1}+\theta+s}\right)^{j+1}\left(\frac{\mu_{v_2}+\theta}{\mu_{v_2}+\theta+s}\right)^{n+1-j}+\sum_{j=0}^{n}\sum_{i=0}^{n-j}P\left\{S_{v_1}^{(j)}<V_{1R}<S_{v_1}^{(j+1)}\cap S_{v_2}^{(i)}<V_2<S_{v_2}^{(i+1)}\right\}\\
&\times\left(\frac{\mu_{v_1}+\theta}{\mu_{v_1}+\theta+s}\right)^{j+1}\left(\frac{\mu_{v_2}+\theta}{\mu_{v_2}+\theta+s}\right)^{i+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1-i-j}\\
=&\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{n+1}+\frac{\theta}{\mu_{v_1}+\theta+s}\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{n+1}\sum_{j=0}^{n}\left(\frac{\mu_{v_1}(\mu_{v_2}+\theta+s)}{\mu_{v_2}(\mu_{v_1}+\theta+s)}\right)^{j}\\
&+\frac{\theta}{\mu_{v_1}+\theta+s}\frac{\theta}{\mu_{v_2}+\theta+s}\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1}\sum_{j=0}^{n}\sum_{i=0}^{n-j}\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{j}\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{i}\left(\frac{\mu_b}{\mu_b+s}\right)^{-(i+j)}\\
=&\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{n+1}+\left(\frac{\theta\mu_{v_2}}{(\theta+s)(\mu_{v_1}-\mu_{v_2})}\right)\left(\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{n+1}-\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{n+1}\right)\\
&+\frac{\mu_b\theta^2}{\mu_b\theta+(\mu_b-\mu_{v_2})s}\left(\frac{\mu_b}{\mu_b\theta+(\mu_b-\mu_{v_1})s}\left(\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1}-\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{n+1}\right)\right.\\
&\left.+\frac{\mu_{v_2}}{(\theta+s)(\mu_{v_1}-\mu_{v_2})}\left(\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{n+1}-\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{n+1}\right)\right)
\end{aligned} \tag{4-67}$$

然后得到 $E[W_1(n)]=-\widetilde{W}_1^{*\prime}(0)$，即

$$E\left[W_1(n)\right]=\frac{n+1}{\mu_b}+\frac{2\mu_b-\mu_{v_1}-\mu_{v_2}}{\mu_b\theta}+\frac{\mu_{v_1}^2-2\mu_b\mu_{v_1}+\mu_b\mu_{v_2}}{\mu_b\theta(\mu_{v_1}-\mu_{v_2})}\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta}\right)^{n+1}+\frac{\mu_{v_2}(\mu_b-\mu_{v_2})}{\mu_b\theta(\mu_{v_1}-\mu_{v_2})}\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta}\right)^{n+1} \tag{4-68}$$

对于系统状态$(n, 2)$，我们需要讨论两种情况：如果第二阶段工作休假的剩余时间 V_{2R} 内足够服务完 n+1 个顾客，则 $W_2(n)$等于 n+1 个参数为 μ_{v_2} 的服务时间之和，否则，如果只有 $j(0\leq j\leq n)$个顾客在 V_{2R} 期间完成服务而其他 $n+1-j$ 个顾客在下个忙期完成服务，则 $W_2(n)$等于 V_{2R} 和 $n+1-j$ 个参数为 μ_b 的服务时间之和。因此，得到 $W_2(n)$的 Laplace 变换为

$$\begin{aligned}\widetilde{W}_2^*(s)&=P\left\{S_{v_2}^{(n+1)}<V_{2R}\right\}\left(\frac{\mu_{v_2}+\theta}{\mu_{v_2}+\theta+s}\right)^{n+1}\\&\quad+\sum_{j=0}^{n}P\left\{S_{v_2}^{(j)}<V_{2R}<S_{v_2}^{(j+1)}\right\}\left(\frac{\mu_{v_2}+\theta}{\mu_{v_2}+\theta+s}\right)^{j+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1-j}\\&=\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta}\right)^{n+1}\left(\frac{\mu_{v_2}+\theta}{\mu_{v_2}+\theta+s}\right)^{n+1}\\&\quad+\sum_{j=0}^{n}\left(\frac{\theta}{\mu_{v_2}+\theta}\right)\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta}\right)^{j}\left(\frac{\mu_{v_2}+\theta}{\mu_{v_2}+\theta+s}\right)^{j+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1-j}\\&=\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{n+1}+\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1}\frac{\theta}{\mu_{v_2}+\theta+s}\sum_{j=0}^{n}\left(\frac{\mu_{v_2}(\mu_b+s)}{\mu_b(\mu_{v_2}+\theta+s)}\right)^{j}\\&=\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{n+1}+\frac{\mu_b\theta}{\mu_b\theta+(\mu_b-\mu_{v_2})s}\left(\left(\frac{\mu_b}{\mu_b+s}\right)^{n+1}-\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{n+1}\right)\end{aligned} \tag{4-69}$$

然后得到$E[W_2(n)]=-\widetilde{W}_2^{*\prime}(0)$，即

$$E\left[W_2(n)\right]=\frac{n+1}{\mu_b}+\frac{\mu_b-\mu_{v_2}}{\theta\mu_b}\left(1-\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta}\right)^{n+1}\right) \tag{4-70}$$

基于式（4-66）、式（4-68）和式（4-70），可分别得到 $U_i(n)=R-CE[W_i(n)]$ $(i=0,1,2)$。均衡状态下，标记顾客的平均剩余效用等于零，即 $U_i^e(n)=0$。因此，求解 $U_i^e(n)=0$，可得到唯一可行解 $n_e^*(i)$[①]。因此，均衡状态下，顾客的整体止步

① 在先到先服务的服务规则和到达过程与休假时间独立的假设下，在休假期内一个顾客的进入只会给后到达顾客带来负效应，因此 $U_1(n)$显然随 n 单调递减。

阈值为 $(n_e(0), n_e(1), n_e(2)) = (\lfloor n_e^*(0) \rfloor, \lfloor n_e^*(1) \rfloor, \lfloor n_e^*(2) \rfloor)$。

图 4-12 表明 $n_e(i)$ (i=0,1,2)随 μ_b 递增，且当 $\mu_{v_1} < \mu_{v_2}$ 时，$n_e(0) \geq n_e(2) \geq n_e(1)$。另一方面，如果 μ_{v_1} 相比 μ_{v_2} 很大且 μ_b 相对较小时，在第一阶段工作休假期到达的第 n 个顾客相比于在第二阶段工作休假期到达的第 n 个顾客更有可能在当前休假期内完成服务。然而，随着 μ_b 的增大，尽管 μ_{v_2} 较小，但在第二阶段工作休假期到达的顾客将面临更加乐观的形势，因此 $n_e(2)$的增速相比 $n_e(1)$更快。图 4-13 表明随着 μ_b 的增长，起初 $n_e(1) \geq n_e(2)$而后 $n_e(2) \geq n_e(1)$。总之，可观察到对于所有可行的 μ_b，当 $\mu_{v_1} < \mu_{v_2}$ 或 μ_{v_1} 超过 μ_{v_2} 不是很多时，$n_e(0) \geq n_e(2) \geq n_e(1)$。然而，当 μ_{v_1} 超过 μ_{v_2} 很多且 μ_{v_1} 接近 μ_b 时，此结论不成立。

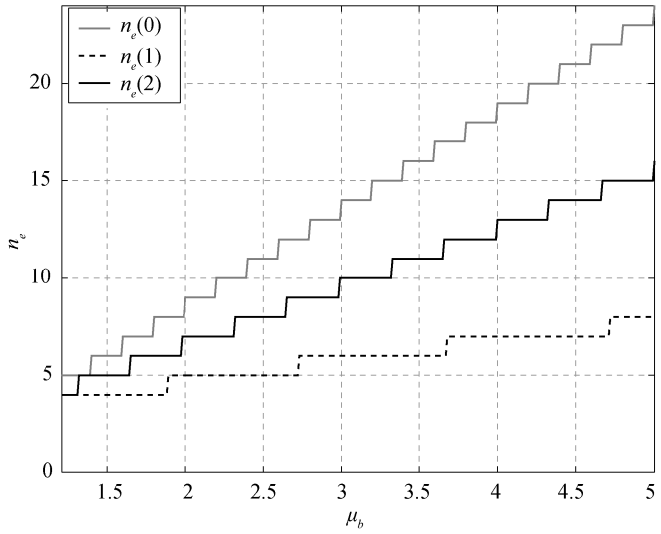

图 4-12　完全可视情形下的均衡止步阈值

（$R = 5$, $C = 1$, $\mu_{v_1} = 0.8$, $\mu_{v_2} = 1$, $\mu = 0.5$）

接下来，我们导出完全可视情形下系统的稳态队长分布。鉴于 $n_e(1)$与 $n_e(2)$间的大小关系，我们讨论两种情形：$n_e(1) > n_e(2)$与 $n_e(1) < n_e(2)$[①]。对于第一种情形，均衡状态转移图如图 4-14 所示。

因此，该 Markov 过程$\{N(t), I(t)\}$的状态空间为

$$\Omega_{ob_1} = \{(n,0): 0 \leq n \leq n_e(0)+1\} \cup \{(n,j): 0 \leq n \leq n_e(1)+1, j = 1,2\}$$

① 情形 $n_e(1) = n_e(2)$可被认为是 $n_e(1) > n_e(2)$或 $n_e(1) < n_e(2)$的特殊情况。

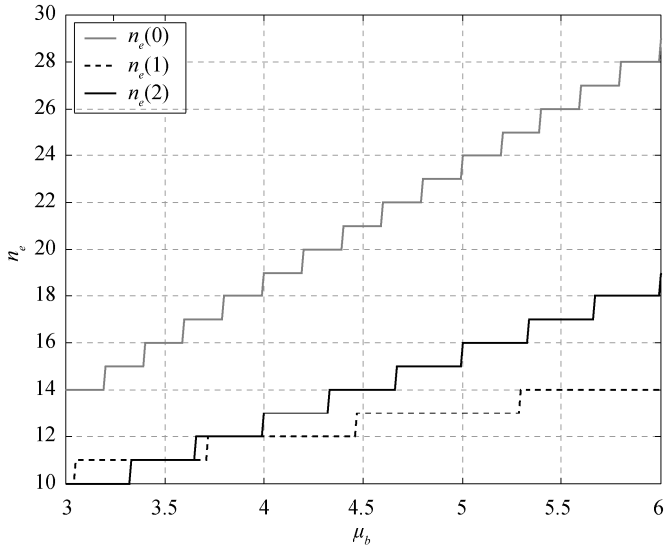

图 4-13　完全可视情形下的均衡止步阈值

（$R=5$, $C=1$, $\mu_{v_1}=3$, $\mu_{v_2}=1$, $\mu=0.5$）

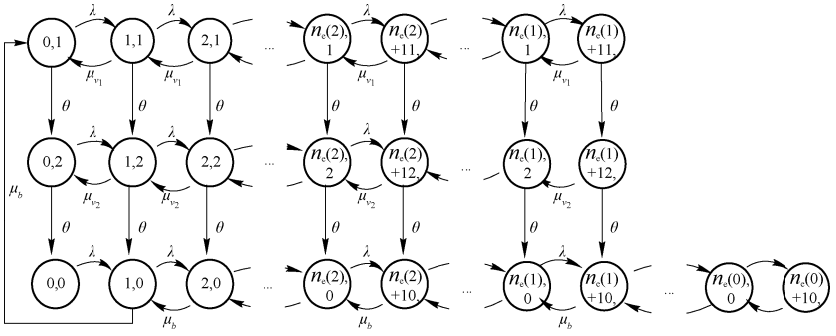

图 4-14　完全可视排队系统的均衡状态转移图[$n_e(1)>n_e(2)$]

记稳态队长分布为

$$\pi_{nj}=P\{N=n,I=j\}=\lim_{t\to\infty}P\{N(t)=n,I(t)=j\},\quad (n,j)\in\Omega_{ob_1}$$

状态转移概率方程为

$$\pi_{00}\lambda=\pi_{02}\theta \tag{4-71}$$

$$\pi_{n0}(\lambda+\mu_b)=\pi_{n-10}\lambda+\pi_{n+10}\mu_b+\pi_{n2}\theta,\ 1\le n\le n_e(1)+1 \tag{4-72}$$

$$\pi_{n0}(\lambda+\mu_b)=\pi_{n-10}\lambda+\pi_{n+10}\mu_b,\ n_e(1)+2\le n\le n_e(0) \tag{4-73}$$

$$\pi_{n_e(0)+10}\mu_b=\pi_{n_e(0)0}\lambda \tag{4-74}$$

$$\pi_{01}(\lambda+\theta)=\pi_{10}\mu_b+\pi_{11}\mu_{v_1} \tag{4-75}$$

$$\pi_{n1}(\lambda+\theta+\mu_{v_1})=\pi_{n-11}\lambda+\pi_{n+11}\mu_{v_1},\ 1\le n\le n_e(1) \tag{4-76}$$

$$\pi_{n_e(1)+11}(\theta+\mu_{v_1})=\pi_{n_e(1)1}\lambda \tag{4-77}$$

$$\pi_{02}(\lambda+\theta)=\pi_{01}\theta+\pi_{12}\mu_{v_2} \tag{4-78}$$

$$\pi_{n2}(\lambda+\theta+\mu_{v_2})=\pi_{n-12}\lambda+\pi_{n+12}\mu_{v_2}+\pi_{n1}\theta,\ 1\le n\le n_e(2) \tag{4-79}$$

$$\pi_{n_e(2)+12}(\theta+\mu_{v_2})=\pi_{n_e(2)2}\lambda+\pi_{n_e(2)+22}\mu_{v_2}+\pi_{n_e(2)+11}\theta \tag{4-80}$$

$$\pi_{n2}(\theta+\mu_{v_2})=\pi_{n+12}\mu_{v_2}+\pi_{n1}\theta,\ n_e(2)+2\le n\le n_e(1) \tag{4-81}$$

$$\pi_{n_e(1)+12}(\theta+\mu_{v_2})=\pi_{n_e(1)+11}\theta \tag{4-82}$$

首先考虑第一阶段工作休假期间的队长概率$\{\pi_{n1}\,|\,0\le n\le n_e(1)+1\}$。根据式(4-76)，注意到上述概率为以下齐次线性差分方程

$$\mu_{v_1}x_{n+1}-(\lambda+\theta+\mu_{v_1})x_n+\lambda x_{n-1}=0,\ 1\le n\le n_e(1) \tag{4-83}$$

的解，且相应的特征方程

$$\mu_{v_1}x^2-(\lambda+\theta+\mu_{v_1})x+\lambda=0$$

有两个解，分别为

$$x_{1,2}^{*}=\frac{(\lambda+\theta+\mu_{v_1})\pm\sqrt{(\lambda+\theta+\mu_{v_1})^2-4\lambda\mu_{v_1}}}{2\mu_{v_1}} \tag{4-84}$$

因此，方程（4-83）的通解为$x_n^{\hom}=A_1x_1^{*n}+B_1x_2^{*n}$（显然，$x_1^{*}\ne x_2^{*}$），其中$A_1$、$B_1$为未知系数。根据式（4-75）和式（4-77），得到

$$\begin{cases}A_1=\dfrac{x_2^{*n_e(1)}\left(\left(\theta+\mu_{v_1}\right)x_2^{*}-\lambda\right)\mu_b\pi_{10}}{x_2^{*n_e(1)}\left(\left(\theta+\mu_{v_1}\right)x_2^{*}-\lambda\right)\left(\lambda+\theta-\mu_{v_1}x_1^{*}\right)-x_1^{*n_e(1)}\left(\left(\theta+\mu_{v_1}\right)x_1^{*}-\lambda\right)\left(\lambda+\theta-\mu_{v_1}x_2^{*}\right)}\\ B_1=\dfrac{x_1^{*n_e(1)}\left(\left(\theta+\mu_{v_1}\right)x_1^{*}-\lambda\right)\mu_b\pi_{10}}{x_1^{*n_e(1)}\left(\left(\theta+\mu_{v_1}\right)x_1^{*}-\lambda\right)\left(\lambda+\theta-\mu_{v_1}x_2^{*}\right)-x_2^{*n_e(1)}\left(\left(\theta+\mu_{v_1}\right)x_2^{*}-\lambda\right)\left(\lambda+\theta-\mu_{v_1}x_1^{*}\right)}\end{cases} \tag{4-85}$$

因此，

$$\pi_{n1}=A_1x_1^{*n}+B_1x_2^{*n},\quad 0\le n\le n_e(1)+1 \tag{4-86}$$

对于第二阶段工作休假期间的队长分布，首先考虑概率$\{\pi_{n2}\,|\,n_e(2)+2\le n\le n_e(1)+1\}$。根据式（4-81），上述概率为以下非齐次线性差分方程

$$\begin{aligned}\mu_{v_2}x_{n+1}-\left(\theta+\mu_{v_2}\right)x_n&=-\pi_{n1}\theta\\&=-\theta\left(A_1x_1^{*n}+B_1x_2^{*n}\right),\ \ n_e(2)+2\le n\le n_e(1)\end{aligned} \tag{4-87}$$

的解。由于式（4-87）的齐次形式的通解为$x_n^{\hom}=B_3\hat{x}_2^{n}$，其中

$$\hat{x}_2=\frac{\theta+\mu_{v_2}}{\mu_{v_2}} \tag{4-88}$$

因此，方程（4-87）的通解为 $x_n^{\text{gen}} = x_n^{\text{hom}} + x_n^{\text{spec}}$，其中 x_n^{spec} 为方程（4-87）的一个特解。由于方程（4-87）的非齐次部分为参数 x_1^* 和 x_2^* 的几何形式，我们考虑一个形为 $x_n^{\text{spec}} = C_2 x_1^{*n} + D_2 x_2^{*n}$ 的特解。将其代入方程（4-87），得到

$$\begin{cases} C_2 = \dfrac{-\theta A_1}{\mu_{v_2} x_1^* - (\theta + \mu_{v_2})} \\ D_2 = \dfrac{-\theta B_1}{\mu_{v_2} x_2^* - (\theta + \mu_{v_2})} \end{cases} \tag{4-89}$$

因此，

$$x_n^{\text{gen}} = B_3 \hat{x}_2^n + C_2 x_1^{*n} + D_2 x_2^{*n}, \quad n_e(2) + 2 \leq n \leq n_e(1) + 1 \tag{4-90}$$

其中 B_3 为未知系数。再考虑式（4-82），得到

$$B_3 = \frac{\theta\left(A_1 x_1^{*n_e(1)+1} + B_1 x_2^{*n_e(1)+1}\right) - (\theta + \mu_{v_2})\left(C_2 x_1^{*n_e(1)+1} + D_2 x_2^{*n_e(1)+1}\right)}{(\theta + \mu_{v_2})\hat{x}_2^{n_e(1)+1}} \tag{4-91}$$

因此，根据式（4-90），可得到

$$\pi_{n2} = B_3 \hat{x}_2^n + C_2 x_1^{*n} + D_2 x_2^{*n}, \quad n_e(2) + 2 \leq n \leq n_e(1) + 1 \tag{4-92}$$

然后考虑概率 $\{\pi_{n2} \mid 0 \leq n \leq n_e(2) + 1\}$。根据式（4-79），上述概率为以下非齐次线性差分方程

$$\begin{aligned} \mu_{v_2} x_{n+1} - (\lambda + \theta + \mu_{v_2}) x_n + \lambda x_{n-1} = -\pi_{n1}\theta = -\theta\left(A_1 x_1^{*n} + B_1 x_2^{*n}\right), \\ 1 \leq n \leq n_e(2) \end{aligned} \tag{4-93}$$

的解。由于式（4-93）的齐次形式的通解为 $x_n^{\text{hom}} = A_2 \tilde{x}_1^n + B_2 \tilde{x}_2^n$，其中

$$\tilde{x}_{1,2} = \frac{(\lambda + \theta + \mu_{v_2}) \pm \sqrt{(\lambda + \theta + \mu_{v_2})^2 - 4\lambda\mu_{v_2}}}{2\mu_{v_2}} \tag{4-94}$$

因此方程（4-93）的通解为 $x_n^{\text{gen}} = x_n^{\text{hom}} + x_n^{\text{spec}}$。其中 $x_n^{\text{spec}} = C_2 x_1^{*n} + D_1 x_2^{*n}$。将其代入方程（4-93），得到

$$\begin{cases} C_1 = \dfrac{-\theta A_1 x_1^*}{\mu_{v_2} x_1^{*2} - (\theta + \lambda + \mu_{v_2}) x_1^* + \lambda} \\ D_1 = \dfrac{-\theta B_1 x_2^*}{\mu_{v_2} x_2^{*2} - (\theta + \lambda + \mu_{v_2}) x_2^* + \lambda} \end{cases} \tag{4-95}$$

因此，

$$x_n^{\text{gen}} = A_2 \tilde{x}_1^n + B_2 \tilde{x}_2^n + C_1 x_1^{*n} + D_1 x_2^{*n}, \quad 0 \leq n \leq n_e(2) + 1 \tag{4-96}$$

其中 A_2、B_2 为未知系数。再考虑式（4-78）、式（4-80）和式（4-82），得到

$$\begin{cases} A_2=\Big(\big((\lambda+\theta)-\mu_{v_2}\tilde{x}_2\big)\big(C_1x_1^{*n_e(2)}\big(\lambda-(\mu_{v_2}+\theta)x_1^*\big)+D_1x_2^{*n_e(2)}\big(\lambda-(\mu_{v_2}+\theta)x_2^*\big)\\ \quad+\theta\big(A_1x_1^{*n_e(2)+1}+B_1x_2^{*n_e(2)+1}\big)+\mu_{v_2}\big(B_3\hat{x}_2^{n_e(2)+2}+C_2x_1^{*n_e(2)+2}+D_2x_2^{*n_e(2)+2}\big)\big)\\ \quad-\tilde{x}_2^{n_e(2)}\big((\mu_{v_2}+\theta)\tilde{x}_2-\lambda\big)\big(\theta(A_1+B_1)+\mu_{v_2}(C_1x_1^*+D_1x_2^*)-(\lambda+\theta)(C_1+D_1)\big)\big)\\ \quad\times\big(\tilde{x}_1^{n_e(2)}\big((\mu_{v_2}+\theta)\tilde{x}_1-\lambda\big)\big((\lambda+\theta)-\mu_{v_2}\tilde{x}_2\big)-\tilde{x}_2^{n_e(2)}\big((\mu_{v_2}+\theta)\tilde{x}_2-\lambda\big)\big((\lambda+\theta)-\mu_{v_2}\tilde{x}_1\big)\big)^{-1}\\ B_2=\Big(\big((\lambda+\theta)-\mu_{v_2}\tilde{x}_1\big)\big(C_1x_1^{*n_e(2)}\big(\lambda-(\mu_{v_2}+\theta)x_1^*\big)+D_1x_2^{*n_e(2)}\big(\lambda-(\mu_{v_2}+\theta)x_2^*\big)\\ \quad+\theta\big(A_1x_1^{*n_e(2)+1}+B_1x_2^{*n_e(2)+1}\big)+\mu_{v_2}\big(B_3\hat{x}_2^{n_e(2)+2}+C_2x_1^{*n_e(2)+2}+D_2x_2^{*n_e(2)+2}\big)\big)\\ \quad-\tilde{x}_1^{n_e(2)}\big((\mu_{v_2}+\theta)\tilde{x}_1-\lambda\big)\big(\theta(A_1+B_1)+\mu_{v_2}(C_1x_1^*+D_1x_2^*)-(\lambda+\theta)(C_1+D_1)\big)\big)\\ \quad\times\big(\tilde{x}_2^{n_e(2)}\big((\mu_{v_2}+\theta)\tilde{x}_2-\lambda\big)\big((\lambda+\theta)-\mu_{v_2}\tilde{x}_1\big)-\tilde{x}_1^{n_e(2)}\big((\mu_{v_2}+\theta)\tilde{x}_1-\lambda\big)\big((\lambda+\theta)-\mu_{v_2}\tilde{x}_2\big)\big)^{-1} \end{cases} \tag{4-97}$$

因此，根据式（4-96），可得到

$$\pi_{n2}=A_2\tilde{x}_1^n+B_2\tilde{x}_2^n+C_1x_1^{*n}+D_1x_2^{*n},\quad 0\le n\le n_e(2)+1 \tag{4-98}$$

接下来分析服务员处于忙期的队长分布，首先考虑概率 $\{\pi_{n0}\,|\,0\le n\le n_e(2)+1\}$。根据式（4-72），上述概率为以下非齐次线性差分方程

$$\begin{aligned}\mu_bx_{n+1}-(\lambda+\mu_b)x_n+\lambda x_{n-1}&=-\pi_{n2}\theta\\&=-\theta\big(A_2\tilde{x}_1^n+B_2\tilde{x}_2^n+C_1x_1^{*n}+D_1x_2^{*n}\big),1\le n\le n_e(2)\end{aligned} \tag{4-99}$$

的解。因此，方程（4-99）的通解为 $x_n^{\text{gen}}=x_n^{\text{hom}}+x_n^{\text{spec}}$，其中 $x_n^{\text{hom}}=A_4+B_4\rho^n$ 且 $x_n^{\text{spec}}=E_1\tilde{x}_1^n+F_1\tilde{x}_2^n+C_3x_1^{*n}+D_3x_2^{*n}$。将其代入方程（4-99），得到

$$\begin{cases} E_1=\dfrac{-\theta A_2\tilde{x}_1}{\mu_b\tilde{x}_1^2-(\lambda+\mu_b)\tilde{x}_1+\lambda}\\ F_1=\dfrac{-\theta B_2\tilde{x}_2}{\mu_b\tilde{x}_2^2-(\lambda+\mu_b)\tilde{x}_2+\lambda}\\ C_3=\dfrac{-\theta C_1x_1^*}{\mu_bx_1^{*2}-(\lambda+\mu_b)x_1^*+\lambda}\\ D_3=\dfrac{-\theta D_1x_2^*}{\mu_bx_2^{*2}-(\lambda+\mu_b)x_2^*+\lambda} \end{cases} \tag{4-100}$$

因此，

$$x_n^{\text{gen}}=A_4+B_4\rho^n+E_1\tilde{x}_1^n+F_1\tilde{x}_2^n+C_3x_1^{*n}+D_3x_2^{*n},\quad 0\le n\le n_e(2)+1 \tag{4-101}$$

其中 A_4，B_4 为未知系数。再考虑式（4-71），得到

$$\begin{cases} A_4=\dfrac{(\theta(A_2+B_2+C_1+D_1)-\lambda(E_1+F_1+C_3+D_3))-\mu_b(\pi_{10}-(E_1\tilde{x}_1+F_1\tilde{x}_2+C_3x_1^*+D_3x_2^*))}{\lambda-\mu_b}\\ B_4=\dfrac{(\theta(A_2+B_2+C_1+D_1)-\lambda(E_1+F_1+C_3+D_3))-\lambda(\pi_{10}-(E_1\tilde{x}_1+F_1\tilde{x}_2+C_3x_1^*+D_3x_2^*))}{\lambda(1-\rho)} \end{cases} \tag{4-102}$$

因此，根据式（4-101），可得到

$$\pi_{n0}=A_4+B_4\rho^n+E_1\tilde{x}_1^n+F_1\tilde{x}_1^n+C_3x_1^{*n}+D_3x_2^{*n},\quad 0\le n\le n_e(2)+1 \tag{4-103}$$

然后考虑概率 $\{\pi_{n0}\mid n_e(2)+2\le n\le n_e(1)+1\}$，上述概率为以下非齐次线性差分方程

$$\begin{aligned}\mu_b x_{n+1}-(\lambda+\mu_b)x_n+\lambda x_{n-1}&=-\pi_{n2}\theta\\&=-\theta(B_3\tilde{x}_2^n+C_2x_1^{*n}+D_2x_2^{*n}),\ n_e(2)+2\le n\le n_e(1)+1\end{aligned} \tag{4-104}$$

的解。因此，方程（4-104）的通解为 $x_n^{\text{gen}}=x_n^{\text{hom}}+x_n^{\text{spec}}$，其中 $x_n^{\text{hom}}=A_5+B_5\rho^n$ 且 $x_n^{\text{spec}}=F_2\tilde{x}_2^n+C_4x_1^{*n}+D_4x_2^{*n}$。将其代入方程（4-104），得到

$$\begin{cases}F_2=\dfrac{-\theta B_3\tilde{x}_2}{\mu_b\tilde{x}_2^2-(\lambda+\mu_b)\tilde{x}_2+\lambda}\\[2ex] C_4=\dfrac{-\theta C_2x_1^*}{\mu_b x_1^{*2}-(\lambda+\mu_b)x_1^*+\lambda}\\[2ex] D_4=\dfrac{-\theta D_2x_2^*}{\mu_b x_2^{*2}-(\lambda+\mu_b)x_2^*+\lambda}\end{cases} \tag{4-105}$$

因此，

$$x_n^{\text{gen}}=A_5+B_5\rho^n+F_2\tilde{x}_2^n+C_4x_1^{*n}+D_4x_2^{*n},\qquad n_e(2)+2\le n\le n_e(1)+1 \tag{4-106}$$

其中 A_5、B_5 为未知系数。再考虑式（4-103），得到

$$\begin{cases}A_5=A_4+\dfrac{E_1\tilde{x}_1^{n_e(2)+1}(\tilde{x}_1-\rho)+\tilde{x}_2^{n_e(2)+1}(\tilde{x}_2-\rho)(F_1-F_2)}{1-\rho}\\[2ex]\qquad+\dfrac{x_1^{*n_e(2)+1}(x_1^*-\rho)(C_3-C_4)+x_2^{*n_e(2)+1}(x_2^*-\rho)(D_3-D_4)}{1-\rho}\\[2ex] B_5=B_4+\dfrac{E_1\tilde{x}_1^{n_e(2)+1}(1-\tilde{x}_1)+\tilde{x}_2^{n_e(2)+1}(1-\tilde{x}_2)(F_1-F_2)}{\rho^{n_e(2+1)}(1-\rho)}\\[2ex]\qquad+\dfrac{x_1^{*n_e(2)+1}(1-x_1^*)(C_3-C_4)+x_2^{*n_e(2)+1}(1-x_2^*)(D_3-D_4)}{\rho^{n_e(2+1)}(1-\rho)}\end{cases} \tag{4-107}$$

因此，根据式（4-106），可得到

$$\pi_{n0}=A_5+B_5\rho^n+F_2\tilde{x}_2^n+C_4x_1^{*n}+D_4x_2^{*n},\quad n_e(2)+2\le n\le n_e(1)+1 \tag{4-108}$$

最后，考虑概率 $\{\pi_{n0}\mid n_e(1)+2\le n\le n_e(0)+1\}$。根据式(4-73)，得到 $x_n^{\text{hom}}=A_6+B_6\rho^n$，其中 A_6、B_6 为未知系数。再考虑式（4-74）和式（4-108），得到 $A_6=0$ 和

$$B_6=\frac{A_5+B_5\rho^{n_e(1)+2}+F_2\tilde{x}_2^{n_e(1)+2}+C_4x_1^{*n_e(1)+2}+D_4x_2^{*n_e(1)+2}}{\rho^{n_e(1+2)}} \tag{4-109}$$

因此，

$$\pi_{n0}=B_6\rho^n,\qquad n_e(1)+2\le n\le n_e(0)+1 \tag{4-110}$$

总之，所有稳态状态概率 $\{\pi_{nj}\mid(n,j)\in\Omega_{ob_1}\}$ 均与 π_{10} 相关。应用平衡方程，得到如下定理中的结果。

定理 4.3.1 对于具有两阶段工作休假策略的完全可视 Markov 排队系统，状态空间为 $\Omega_{ob_1}=\{(n,0):0\le n\le n_e(0)+1\}\bigcup\{(n,j):0\le n\le n_e(1)+1,j=1,2\}$，其中$(n_e(0),n_e(1),n_e(2))$为顾客的均衡阈值止步策略，则系统稳态队长分布$\{\pi_{nj}\mid(n,j)\in\Omega_{ob_1}\}$为

$$\pi_{n1}=A_1x_1^{*n}+B_1x_2^{*n},\quad 0\le n\le n_e(1)+1 \tag{4-111}$$

$$\pi_{n2}=A_2\tilde{x}_1^{n}+B_2\tilde{x}_2^{n}+C_1x_1^{*n}+D_1x_2^{*n},\quad 0\le n\le n_e(2)+1 \tag{4-112}$$

$$\pi_{n2}=B_3\hat{x}_2^{n}+C_2x_1^{*n}+D_2x_2^{*n},\qquad n_e(2)+2\le n\le n_e(1)+1 \tag{4-113}$$

$$\pi_{n0}=A_4+B_4\rho^{n}+E_1\tilde{x}_1^{n}+F_1\tilde{x}_2^{n}+C_3x_1^{*n}+D_3x_2^{*n},\quad 0\le n\le n_e(2)+1 \tag{4-114}$$

$$\pi_{n0}=A_5+B_5\rho^{n}+F_2\tilde{x}_2^{n}+C_4x_1^{*n}+D_4x_2^{*n},\qquad n_e(2)+2\le n\le n_e(1)+1 \tag{4-115}$$

$$\pi_{n0}=B_6\rho^{n},\qquad n_e(1)+2\le n\le n_e(0)+1 \tag{4-116}$$

其中 $\rho=\lambda/\mu_b\ne1$ 且 $x_{1,2}^{*}$、$\tilde{x}_{1,2}$、$\hat{x}_2$、$A_i(i=1,2,4,5)$、$B_j(1\le j\le6)$、C_k、$D_k(1\le k\le4)$、E_1、F_1、F_2 由式（4-84）、式（4-85）、式（4-88）、式（4-89）、式（4-91）、式（4-94）、式（4-95）、式（4-97）、式（4-100）、式（4-102）、式（4-105）、式（4-107）、式（4-109）分别给出，且 π_{10} 可由平衡方程

$$\sum_{(n,j)\in\Omega_{ob_1}}\pi_{nj}=1 \tag{4-117}$$

求得。

基于图 4-14 和定理 4.3.1，顾客止步状态为 $(n_e(0)+1,0),(n_e(1)+1,1)$ 或 $\{(n,2):n_e(2)+1\le n\le n_e(1)+1\}$。因此，均衡状态下的单位社会福利为

$$\begin{aligned}U_s(n_e(0),n_e(1),n_e(2))=\lambda R\left(1-\pi_{n_e(0)+10}-\pi_{n_e(1)+11}-\sum_{n=n_e(2)+1}^{n_e(1)+1}\pi_{n2}\right)\\-C\left(\sum_{n=1}^{n_e(0)+1}n\pi_{n0}+\sum_{n=1}^{n_e(1)+1}n(\pi_{n1}+\pi_{n2})\right)\end{aligned} \tag{4-118}$$

对于第二种情形 $n_e(1)<n_e(2)$，状态转移图如图 4-15 所示。

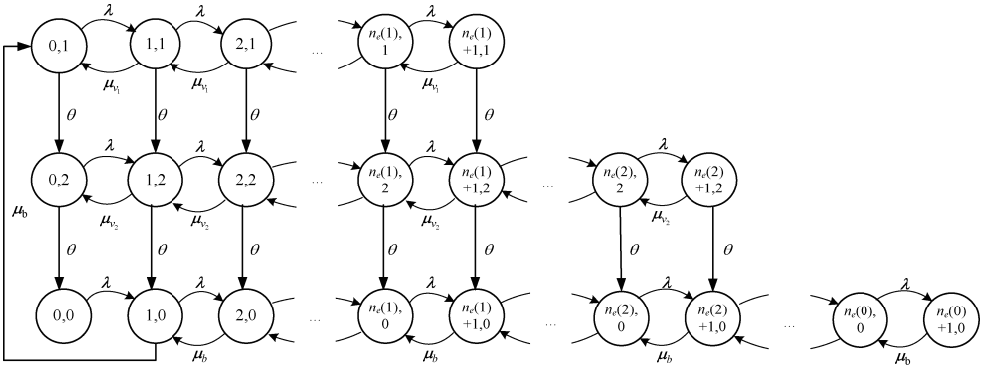

图 4-15 完全可视排队系统的状态转移图[$n_e(1)<n_e(2)$]

因此，Markov 过程$\{N(t), I(t)\}$的状态空间为

$$\Omega_{ob_2} = \{(n, j) : 0 \le n \le n(j) + 1, j = 0,1,2\}$$

状态转移概率方程为

$$\pi_{00}\lambda = \pi_{02}\theta \tag{4-119}$$

$$\pi_{n0}(\lambda + \mu_b) = \pi_{n-10}\lambda + \pi_{n+10}\mu_b + \pi_{n2}\theta, \quad 1 \le n \le n_e(2) + 1 \tag{4-120}$$

$$\pi_{n0}(\lambda + \mu_b) = \pi_{n-10}\lambda + \pi_{n+10}\mu_b, \qquad n_e(2) + 2 \le n \le n_e(0) \tag{4-121}$$

$$\pi_{n_e(0)+10}\mu_b = \pi_{n_e(0)0}\lambda \tag{4-122}$$

$$\pi_{01}(\lambda + \theta) = \pi_{10}\mu_b + \pi_{11}\mu_{v_1} \tag{4-123}$$

$$\pi_{n1}(\lambda + \theta + \mu_{v_1}) = \pi_{n-11}\lambda + \pi_{n+11}\mu_{v_1}, \quad 1 \le n \le n_e(1) \tag{4-124}$$

$$\pi_{n_e(1)+11}(\theta + \mu_{v_1}) = \pi_{n_e(1)1}\lambda \tag{4-125}$$

$$\pi_{02}(\lambda + \theta) = \pi_{01}\theta + \pi_{12}\mu_{v_2} \tag{4-126}$$

$$\pi_{n2}(\lambda + \theta + \mu_{v_2}) = \pi_{n-12}\lambda + \pi_{n+12}\mu_{v_2} + \pi_{n1}\theta, \quad 1 \le n \le n_e(1) + 1 \tag{4-127}$$

$$\pi_{n2}(\lambda + \theta + \mu_{v_2}) = \pi_{n-12}\lambda + \pi_{n+12}\mu_{v_2}, \qquad n_e(1) + 2 \le n \le n_e(2) \tag{4-128}$$

$$\pi_{n_e(2)+12}\mu_{v_2} = \pi_{n_e(2)2}\lambda \tag{4-129}$$

与第一种情形类似，我们也可得到第二种情形下的稳态队长分布$\{\pi_{nj} \mid (n, j) \in \Omega_{ob_2}\}$。根据图 4-15，可知顾客的止步状态为($n_e$(0)+1,0),($n_e$(1)+1,1)或($n_e$(2)+1,2)。因此，均衡状态下的单位社会福利为

$$\begin{aligned} U_s(n_e(0), n_e(1), n_e(2)) = {} & \lambda R(1 - \pi_{n_e(0)+10} - \pi_{n_e(1)+11}) \\ & - C\left(\sum_{n=1}^{n_e(0)+1} n\pi_{n0} + \sum_{n=1}^{n_e(1)+1} n\pi_{n1} + \sum_{n=1}^{n_e(2)+1} n\pi_{n2}\right) \end{aligned} \tag{4-130}$$

针对 n_e(1)>n_e(2)和 n_e(1)<n_e(2)两种情形，图 4-16 和图 4-17 分别展示了均衡社会福利关于 λ 的敏感度。我们观察到 U_s(n_e(0), n_e(1), n_e(2))都是先增后减的变化趋势，说明当 λ 较小时，顾客无须等待太长时间，但随着 λ 的增大，顾客面临越来越大的等待损耗，从而使均衡社会福利持续减少。

4.3.3　几乎不可视排队系统

接下来，考虑几乎不可视排队系统，即顾客在时刻 t 只可观测到服务员状态 $I(t)$，但观测不到队长 $N(t)$。因此，顾客的决策问题其实是确定一组进入概率 (q_0, q_1, q_2) $(0 \le q_i \le 1, i = 0,1,2)$，且均衡状态下，这组进入概率表示为 $(q_e(0), q_e(1), q_e(2))$。

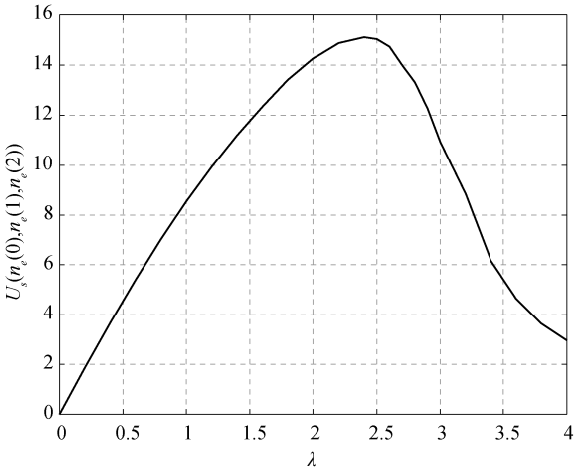

图 4-16　当 $n_e(1)>n_e(2)$ 时的均衡社会福利
（$R=10,\ C=1,\ \mu_b=3,\ \mu_{v_1}=2.5,\ \mu_{v_2}=1,\ \theta=0.2$）

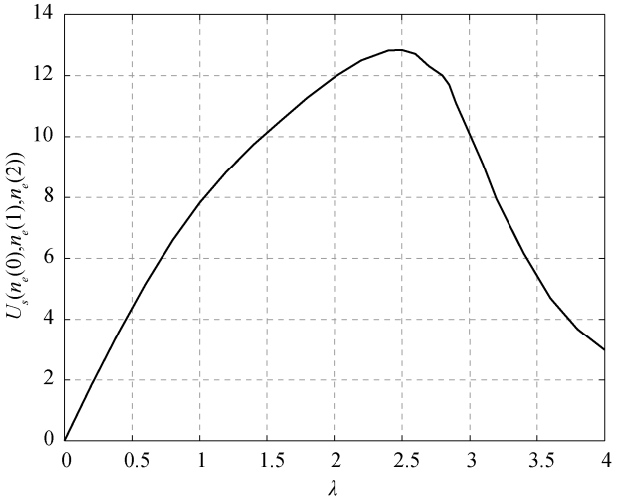

图 4-17　当 $n_e(1)<n_e(2)$ 时的均衡社会福利
（$R=10,\ C=1,\ \mu_b=3,\ \mu_{v_1}=1.2,\ \mu_{v_2}=1,\ \theta=0.2$）

为了导出 $(q_e(0),q_e(1),q_e(2))$，首先导出稳态队长分布。状态转移图如图 4-18 所示。

根据图 4-18，可知过程 $\{N(t),I(t)\}$ 为一个拟生灭过程，状态空间为

$$\Omega_{po}=\{(k,j):k\geq 0,j=0,1,2\}$$

如果 $\rho=\lambda q_0/\mu_b<1$，令(N,I)为过程 $\{N(t),I(t)\}$ 的稳态极限。记稳态队长分布为

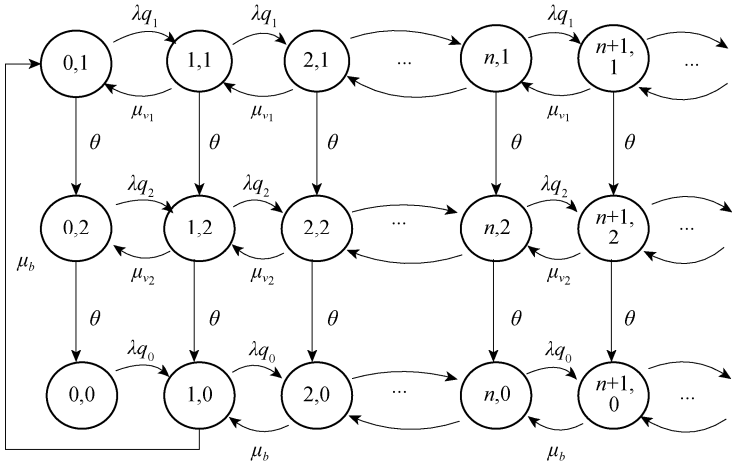

图 4-18　几乎不可视排队系统的状态转移图

$$\boldsymbol{\pi}=(\boldsymbol{\pi}_0,\boldsymbol{\pi}_1,\boldsymbol{\pi}_2,\cdots),\quad \boldsymbol{\pi}_k=(\pi_{k0},\pi_{k1},\pi_{k2}),\qquad k\geq 0$$

$$\boldsymbol{\pi}_{kj}=P\{N=k,I=j\}=\lim_{t\to\infty},P\{N(t)=k,I(t)=j\},\quad (k,j)\in\Omega_{po}$$

该过程的无限生成元记为 $\boldsymbol{Q}$，因而得到稳态状态转移概率方程为 $\boldsymbol{\pi Q}=\boldsymbol{0}$，即

$$\pi_{00}\lambda q_0=\pi_{02}\theta \tag{4-131}$$

$$\pi_{k0}(\lambda q_0+\mu_b)=\pi_{k-10}\lambda q_0+\pi_{k2}\theta+\pi_{k+10\mu}\mu_b,\qquad k\geq 1 \tag{4-132}$$

$$\pi_{01}(\lambda q_1+\theta)=\pi_{10}\mu_b+\pi_{11}\mu_{v_1} \tag{4-133}$$

$$\pi_{k1}(\lambda q_1+\theta+\mu_{v_1})=\pi_{k-11}\lambda q_1+\pi_{k+11}\mu_{v_1},\qquad k\geq 1 \tag{4-134}$$

$$\pi_{02}(\lambda q_2+\theta)=\pi_{01}\theta+\pi_{12}\mu_{v_2} \tag{4-135}$$

$$\pi_{k2}(\lambda q_2+\theta+\mu_{v_2})=\pi_{k-12}\lambda q_2+\pi_{k+12}\mu_{v_2}+\pi_{k1}\theta,\qquad k\geq 1 \tag{4-136}$$

应用词典序，$\boldsymbol{Q}$ 可写成分块矩阵，即

$$\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{A}_{00} & \boldsymbol{C} & & & \\ \boldsymbol{B}_{10} & \boldsymbol{A} & \boldsymbol{C} & & \\ & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} & \\ & & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} \\ & & & \ddots & \ddots & \ddots\end{pmatrix} \tag{4-137}$$

其中

$$\boldsymbol{A}_{00}=\begin{pmatrix}-\lambda q_0 & 0 & 0\\ 0 & -(\theta+\lambda q_1) & \theta\\ \theta & 0 & -(\theta+\lambda q_2)\end{pmatrix},\qquad \boldsymbol{B}_{10}=\begin{pmatrix}0 & \mu_b & 0\\ 0 & \mu_{v_1} & 0\\ \theta & 0 & \mu_{v_2}\end{pmatrix}$$

$$\boldsymbol{B}=\begin{pmatrix}\mu_b & 0 & 0\\ 0 & \mu_{v_1} & 0\\ 0 & 0 & \mu_{v_2}\end{pmatrix},\qquad \boldsymbol{C}=\begin{pmatrix}\lambda q_0 & 0 & 0\\ 0 & \lambda q_1 & 0\\ 0 & 0 & \lambda q_2\end{pmatrix}$$

$$\boldsymbol{A}=\begin{pmatrix}-(\lambda q_0+\mu_b) & 0 & 0\\ 0 & -(\lambda q_1+\theta+\mu_{v_1}) & \theta\\ \theta & 0 & -(\lambda q_2+\theta+\mu_{v_2})\end{pmatrix}$$

现在，需要求解下列矩阵二次方程

$$\boldsymbol{R}^2\boldsymbol{B}+\boldsymbol{R}\boldsymbol{A}+\boldsymbol{C}=0 \tag{4-138}$$

的最小非负解，即率阵 $\boldsymbol{R}$。鉴于 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 的特殊结构，可知 $\boldsymbol{R}$ 具有类似的结构。因此，假设

$$\boldsymbol{R}=\begin{pmatrix}r_{11} & 0 & 0\\ r_{21} & r_{22} & r_{23}\\ r_{31} & 0 & r_{33}\end{pmatrix} \tag{4-139}$$

将其代入式（4-138），可得到如下方程组

$$\begin{cases}r_{11}^2\mu_b-r_{11}(\lambda q_0+\mu_b)+\lambda q_0=0\\ (r_{21}r_{11}+r_{22}r_{21}+r_{23}r_{31})\mu_b-r_{21}(\lambda q_0+\mu_b)+r_{23}\theta=0\\ r_{22}^2\mu_{v_1}-r_{22}(\theta+\lambda q_1+\mu_{v_1})+\lambda q_1=0\\ (r_{22}r_{23}+r_{23}r_{33})\mu_{v_2}+r_{22}\theta-r_{23}(\theta+\lambda q_2+\mu_{v_2})=0\\ (r_{31}r_{11}+r_{33}r_{31})\mu_b-r_{31}(\lambda q_0+\mu_b)+r_{33}\theta=0\\ r_{33}^2\mu_{v_2}-r_{33}(\theta+\lambda q_2+\mu_{v_2})+\lambda q_2=0\end{cases} \tag{4-140}$$

求解式（4-140），得到

$$\begin{cases}r_{11}=\dfrac{\lambda q_0}{\mu_b}\\ r_{21}=\dfrac{2\mu_{v_1}\theta^2 r_{q_1}^-}{\mu_b((\theta+\lambda q_1-\mu_{v_1})(r_{q_1}^-r_{q_2}^--2\mu_{v_2}r_{q_1}^--2\mu_{v_1}r_{q_2}^-+\mu_{v_2})+\theta(2\mu_{v_1}r_{q_2}^--\mu_{v_1}-\mu_{v_2}))}\\ r_{22}=\dfrac{r_{q_1}^-}{2\mu_{v_1}}\\ r_{23}=\dfrac{\theta r_{q_1}^-}{\mu_{v_1}r_{q_2}^+-\mu_{v_2}r_{q_1}^-}\\ r_{31}=\dfrac{\theta(2\mu_{v_1}\lambda(q_1r_{q_2}^-+q_2r_{q_1}^-)-4\mu_{v_1}^2\lambda q_2-(\theta+\lambda q_1)r_{q_1}^-r_{q_2}^-)}{\mu_b((\theta+\lambda q_1-\mu_{v_1})(r_{q_1}^-r_{q_2}^--2\mu_{v_2}r_{q_1}^--2\mu_{v_1}r_{q_2}^-+\mu_{v_2})+\theta(2\mu_{v_1}r_{q_2}^--\mu_{v_1}-\mu_{v_2}))}\\ r_{33}=\dfrac{r_{q_2}^-}{2\mu_{v_2}}\end{cases} \tag{4-141}$$

其中

$$r_{q_1}^{-}=\theta+\lambda q_1+\mu_{v_1}-\sqrt{(\theta+\lambda q_1+\mu_{v_1})^2-4\lambda q_1\mu_{v_1}}$$

$$r_{q_2}^{-}=\theta+\lambda q_2+\mu_{v_2}-\sqrt{(\theta+\lambda q_2+\mu_{v_2})^2-4\lambda q_2\mu_{v_2}}$$

$$r_{q_2}^{+}=\theta+\lambda q_2+\mu_{v_2}+\sqrt{(\theta+\lambda q_2+\mu_{v_2})^2-4\lambda q_2\mu_{v_2}}$$

应用矩阵解析方法，得到

$$\boldsymbol{\pi}_k=(\pi_{k0},\pi_{k1},\pi_{k2})=(\pi_{00},\pi_{01},\pi_{02})\boldsymbol{R}^k,\qquad k\geq 0 \tag{4-142}$$

且$(\pi_{00},\pi_{01},\pi_{02})$满足

$$(\pi_{00},\pi_{01},\pi_{02})(\boldsymbol{RB}_{10}+\boldsymbol{A}_{00})=\boldsymbol{0} \tag{4-143}$$

平衡条件为

$$(\pi_{00},\pi_{01},\pi_{02})(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e}=1 \tag{4-144}$$

其中 $\boldsymbol{I}$ 为单位阵，$\boldsymbol{e}$ 是单位向量，且

$$\boldsymbol{RB}_{10}+\boldsymbol{A}_{00}=\begin{pmatrix}-\lambda q_0 & r_{11}\mu_b & 0\\ 0 & r_{21}\mu_b+r_{22}\mu_{v_1}-\theta-\lambda q_1 & \theta+r_{23}\mu_{v_2}\\ \theta & r_{31}\mu_b & r_{33}\mu_{v_2}-\theta-\lambda q_2\end{pmatrix} \tag{4-145}$$

将式（4-145）代入式（4-143），得到

$$\begin{cases}-\lambda q_0\pi_{00}+\theta\pi_{02}=0\\ r_{11}\mu_b\pi_{00}+(r_{21}\mu_b+r_{22}\mu_{v_1}-\theta-\lambda q_1)\pi_{01}+r_{31}\mu_b\pi_{02}=0\\ (\theta+r_{23}\mu_{v_2})\pi_{01}+(r_{33}\mu_{v_2}-\theta-\lambda q_2)\pi_{02}=0\end{cases} \tag{4-146}$$

求解式（4-146），得到

$$(\pi_{00},\pi_{01},\pi_{02})=\pi_{00}\left(1,\frac{\lambda q_0(\theta+\lambda q_2-r_{33}\mu_{v_2})}{\theta(\theta+r_{23}\mu_{v_2})},\frac{\lambda q_0}{\theta}\right)$$

而且,

$$\boldsymbol{R}^k=\begin{pmatrix}r_{11}^k & 0 & 0\\ r_{21}\sum\limits_{j=0}^{k-1}r_{11}^j r_{22}^{k-1-j}+r_{23}r_{31}\sum\limits_{i=0}^{k-2k-2-i}\sum\limits_{j=0}r_{11}^i r_{22}^j r_{33}^{k-2-j-i} & r_{22}^k & r_{23}\sum\limits_{j=0}^{k-1}r_{22}^j r_{33}^{k-1-j}\\ r_{31}\sum\limits_{j=0}^{k-1}r_{11}^j r_{33}^{k-1-j} & 0 & r_{33}^k\end{pmatrix},\qquad k\geq 1$$

将$(\pi_{00},\pi_{01},\pi_{02})$和 $\boldsymbol{R}^k$ 代入式（4-142），可得到稳态队长分布为

$$\begin{cases}\pi_{k0}=\dfrac{\pi_{00}}{\theta\left(\theta+r_{23}\mu_{v_2}\right)}r_{11}^{k}+\dfrac{\lambda q_0 r_{31}\pi_{00}}{\theta\left(r_{11}-r_{33}\right)}\left(r_{11}^{k}-r_{33}^{k}\right)+\dfrac{\lambda q_0\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)\pi_{00}}{\theta\left(\theta+r_{23}\mu_{v_2}\right)}\left(\dfrac{r_{21}}{r_{11}-r_{22}}\left(r_{11}^{k}-r_{22}^{k}\right)\right.\\ \qquad\left.+\dfrac{r_{23}r_{31}}{\left(r_{11}-r_{22}\right)\left(r_{11}-r_{33}\right)}r_{11}^{k}-\dfrac{r_{23}r_{31}}{\left(r_{11}-r_{22}\right)\left(r_{22}-r_{33}\right)}r_{22}^{k}+\dfrac{r_{23}r_{31}}{\left(r_{11}-r_{33}\right)\left(r_{22}-r_{33}\right)}r_{33}^{k}\right),\ k\geq 0\\ \pi_{k1}=\dfrac{\lambda q_0\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)\pi_{00}}{\theta\left(\theta+r_{23}\mu_{v_2}\right)}r_{22}^{k},\ k\geq 0\\ \pi_{k2}=\dfrac{\lambda q_0\pi_{00}}{\theta}\left(\dfrac{\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{22}-r_{33}\right)}\left(r_{22}^{k}-r_{33}^{k}\right)+r_{33}^{k}\right),\ k\geq 0\end{cases}\tag{4-147}$$

其中π_{00}可由式（4-144）求得。

根据式（4-147），系统处于状态$i(i=0,1,2)$的概率p_i分别为

$$\begin{aligned}p_0&=P\{I=0\}=\sum_{k=0}^{\infty}\pi_{k0}\\&=\pi_{00}\left(\frac{\theta-r_{23}\mu_{v_2}}{\left(1-r_{11}\right)\left(\theta+r_{23}\mu_{v_2}\right)}-\frac{\lambda q_0 r_{21}\left(\theta+r_{33}\mu_{v_2}-\lambda q_2\right)}{\left(1-r_{11}\right)\left(1-r_{22}\right)\left(\theta+r_{23}\mu_{v_2}\right)\theta}\right.\\&\quad\left.+\frac{\lambda q_0 r_{31}}{\left(1-r_{11}\right)\left(1-r_{33}\right)\theta}+\frac{\lambda q_0 r_{23}r_{31}\left(\theta-r_{33}\mu_{v_2}+\lambda q_2\right)}{\left(1-r_{11}\right)\left(1-r_{22}\right)\left(1-r_{33}\right)\left(\theta+r_{23}\mu_{v_2}\right)\theta}\right)\end{aligned}\tag{4-148}$$

$$p_1=P\{I=1\}=\sum_{k=0}^{\infty}\pi_{k1}=\frac{\lambda q_0\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)\pi_{00}}{\theta\left(1-r_{22}\right)\left(\theta+r_{23}\mu_{v_2}\right)}\tag{4-149}$$

$$\begin{aligned}p_2&=P\{I=2\}=\sum_{k=0}^{\infty}\pi_{k2}\\&=\frac{\lambda q_0\pi_{00}}{\theta}\left(\frac{1}{1-r_{33}}+\frac{r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{22}-r_{33}\right)}\left(\frac{1}{1-r_{22}}-\frac{1}{1-r_{33}}\right)\right)\end{aligned}\tag{4-150}$$

因此，可得到系统处于状态i ($i=0, 1, 2$)的条件平均队长为

$$\begin{aligned}E[L_0]&=\frac{\sum_{k=1}^{\infty}k\pi_{k0}}{\sum_{k=0}^{\infty}\pi_{k0}}\\&=\left(\frac{\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)\left(1-r_{11}r_{22}\right)r_{21}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{11}\right)^2\left(1-r_{22}\right)^2}+\frac{r_{31}\left(1-r_{11}r_{33}\right)}{\left(1-r_{11}\right)^2\left(1-r_{33}\right)^2}+\frac{\theta r_{11}}{\lambda q_0\left(1-r_{11}\right)^2}\right.\\&\quad\left.+\frac{r_{31}r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{22}-r_{33}\right)}\left(\frac{r_{21}\left(2-\left(r_{11}+r_{22}\right)\right)}{\left(1-r_{11}\right)^2\left(1-r_{22}\right)^2}-\frac{r_{33}\left(2-\left(r_{11}+r_{33}\right)\right)}{\left(1-r_{11}\right)^2\left(1-r_{33}\right)^2}\right)\right)\end{aligned}$$

$$\times\left(\frac{\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{21}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{11}\right)\left(1-r_{22}\right)}+\frac{r_{31}}{\left(1-r_{11}\right)\left(1-r_{33}\right)}+\frac{\theta}{\lambda q_0\left(1-r_{11}\right)}\right.$$
$$\left.-\frac{r_{31}r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{11}-r_{33}\right)\left(r_{22}-r_{33}\right)}\frac{1+r_{33}-\left(r_{11}+r_{22}\right)}{\left(1-r_{11}\right)\left(1-r_{22}\right)}\right)^{-1} \tag{4-151}$$

$$E\left[L_1\right]=\frac{\sum_{k=1}^{\infty}k\pi_{k1}}{\sum_{k=0}^{\infty}\pi_{k1}}$$
$$=\frac{\pi_{00}\lambda q_0\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{22}}{\theta\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{22}\right)^2}\left(\frac{\pi_{00}\lambda q_0\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\theta\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{22}\right)}\right)^{-1} \tag{4-152}$$
$$=\frac{r_{22}}{1-r_{22}}$$

$$E\left[L_2\right]=\frac{\sum_{k=1}^{\infty}k\pi_{k2}}{\sum_{k=0}^{\infty}\pi_{k2}}$$
$$=\frac{\lambda q_0\pi_{00}}{\theta}\left(\frac{r_{33}}{\left(1-r_{33}\right)^2}+\frac{r_{22}r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{22}-r_{33}\right)\left(1-r_{22}\right)^2}-\frac{r_{23}r_{33}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{22}-r_{33}\right)\left(1-r_{33}\right)^2}\right)$$
$$\times\left(\frac{\lambda q_0\pi_{00}}{\theta}\left(\frac{1}{1-r_{33}}+\frac{r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(r_{22}-r_{33}\right)\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{22}\right)}-\frac{r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(r_{22}-r_{33}\right)\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{33}\right)}\right)\right)$$
$$=\frac{r_{33}}{1-r_{33}}+\frac{1}{1-r_{22}}\frac{r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(1-r_{22}\right)\left(\theta+r_{23}\mu_{v_2}\right)+r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)} \tag{4-153}$$

将式（4-66）中的 n 替换为 $E[L_0]$，可得到系统处于忙期时顾客的条件平均逗留时间，记为 $E[W_0]$，为

$$E\left[W_0\right]=\frac{E\left[L_0\right]+1}{\mu_b}$$
$$=\left(\frac{\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)\left(2-\left(r_{11}+r_{22}\right)\right)r_{21}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{11}\right)^2\left(1-r_{22}\right)^2}+\frac{r_{31}\left(2-\left(r_{11}+r_{33}\right)\right)}{\left(1-r_{11}\right)^2\left(1-r_{33}\right)^2}+\frac{\theta}{\lambda q_0\left(1-r_{11}\right)^2}\right.$$
$$+\frac{r_{31}r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{11}-r_{33}\right)\left(r_{22}-r_{33}\right)}\left(\frac{r_{11}-r_{33}}{\left(1-r_{11}\right)^2}\left(\frac{r_{21}\left(2-\left(r_{11}+r_{22}\right)\right)}{\left(1-r_{22}\right)^2}-\frac{r_{33}\left(2-\left(r_{11}+r_{33}\right)\right)}{\left(1-r_{33}\right)^2}\right)\right.$$
$$\left.\left.-\frac{1+r_{33}-\left(r_{11}+r_{22}\right)}{\left(1-r_{11}\right)\left(1-r_{22}\right)}\right)\right)\left(\frac{\mu_b\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{21}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{11}\right)\left(1-r_{22}\right)}+\frac{\mu_b r_{31}}{\left(1-r_{11}\right)\left(1-r_{33}\right)}+\frac{\mu_b\theta}{\lambda q_0\left(1-r_{11}\right)}\right.$$

$$-\frac{\mu_b r_{31} r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(r_{11}-r_{33}\right)\left(r_{22}-r_{33}\right)}\frac{1+r_{33}-\left(r_{11}+r_{22}\right)}{\left(1-r_{11}\right)\left(1-r_{22}\right)}\Bigg)^{-1} \tag{4-154}$$

对于系统处于第一阶段工作休假期时顾客的条件平均逗留时间，记为 $E[W_1]$，我们可类似于完全可视情形进行讨论，从而得到顾客逗留时间的 Laplace 变换 $W_1^*(s)$为

$$\begin{aligned}W_1^*(s)=&\sum_{k=0}^{\infty}\pi_{k1}\Bigg(\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{k+1}+\frac{\theta\mu_{v_2}}{(\theta+s)(\mu_{v_1}-\mu_{v_2})}\left(\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{k+1}-\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{k+1}\right)\\&+\frac{\mu_b\theta^2}{\mu_b\theta+(\mu_b-\mu_{v_2})s}\left(\frac{\mu_b}{\mu_b\theta+(\mu_b-\mu_{v_1})s}\left(\left(\frac{\mu_b}{\mu_b+s}\right)^{k+1}-\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{k+1}\right)\right.\\&\left.+\frac{\mu_{v_2}}{(\theta+s)(\mu_{v_1}-\mu_{v_2})}\left(\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{k+1}-\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta+s}\right)^{k+1}\right)\right)\Bigg)\\=&\frac{\lambda q_0(\theta+\lambda q_2-r_{33}\mu_{v_2})\pi_{00}}{\theta(\theta+r_{23}\mu_{v_2})(\mu_{v_1}(1-r_{22})+\theta+s)}\left(\frac{\theta^2\mu_b}{\mu_b((1-r_{22})+s)(\mu_{v_2}(1-r_{22})+\theta+s)}\right.\\&\left.+\frac{\theta\mu_{v_2}}{\mu_{v_2}(1-r_{22})+\theta+s}+\mu_{v_1}\right)\end{aligned} \tag{4-155}$$

因此，系统处于第一阶段工作休假期时顾客的条件逗留时间的 Laplace 变换为

$$\begin{aligned}\widetilde{W}_1^*(s)=&\frac{1}{p_1}W_1^*(s)\\=&\frac{1-r_{22}}{\mu_{v_1}(1-r_{22})+\theta+s}\left(\frac{\theta^2\mu_b}{(\mu_b(1-r_{22})+s)(\mu_{v_2}(1-r_{22})+\theta+s)}\right.\\&\left.+\frac{\theta\mu_{v_2}}{\mu_{v_2}(1-r_{22})+\theta+s}+\mu_{v_1}\right)\end{aligned} \tag{4-156}$$

从而得到 $E\left[W_1\right]=-\widetilde{W}_1^*\left(s\right)$，即

$$E\left[W_1\right]=\frac{\left(\mu_b\left(1-r_{22}\right)+\theta\right)^2-\mu_b\left(1-r_{22}\right)^2\left(\mu_b-\mu_{v_2}\right)}{\left(\theta+\mu_{v_1}\left(1-r_{22}\right)\right)\left(\theta+\mu_{v_2}\left(1-r_{22}\right)\right)\left(1-r_{22}\right)\mu_b} \tag{4-157}$$

另一方面，对于系统处于第二阶段工作休假期时顾客的条件平均逗留时间，记为 $E[W_2]$，也可类似于完全可视情形进行讨论，从而得到顾客逗留时间的 Laplace 变换为

$$
\begin{aligned}
W_2^*(s)&=\sum_{k=0}^{\infty}\pi_{k2}\left(\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{k+1}+\frac{\mu_b\theta}{\mu_b\theta+(\mu_b-\mu_{v_2})s}\left(\left(\frac{\mu_b}{\mu_b+s}\right)^{k+1}-\left(\frac{\mu_{v_2}}{\mu_{v_2}+\theta+s}\right)^{k+1}\right)\right)\\
&=\frac{\lambda q_0\pi_{00}}{\theta}\left(\frac{\mu_b\left(\mu_{v_2}(1-r_{33})+\theta\right)+\mu_{v_2}s}{\left(\mu_{v_2}(1-r_{33})+\theta+s\right)\left(\mu_b(1-r_{33})+s\right)}-\frac{\mu_{v_2}^2\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(\mu_{v_2}s-\mu_b(\theta+s)\right)}\right.\\
&\quad\times\frac{(\mu_b-\mu_{v_2})s}{\left((1-r_{22})\mu_{v_2}+\theta+s\right)\left((1-r_{33})\mu_{v_2}+\theta+s\right)}-\frac{\mu_b^2\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(\mu_{v_2}s-\mu_b(\theta+s)\right)}\\
&\quad\left.\times\frac{\mu_b\theta}{\left((1-r_{22})\mu_b+s\right)\left((1-r_{33})\mu_b+s\right)}\right)
\end{aligned}
\tag{4-158}
$$

因此，系统处于第二阶段工作休假期时顾客的条件逗留时间的 Laplace 变换为

$$
\begin{aligned}
\widetilde{W}_2^*(s)&=\frac{1}{p_2}W_2^*(s)\\
&=\left(\frac{\mu_b\left(\mu_{v_2}(1-r_{33})+\theta\right)+\mu_{v_2}s}{\left(\mu_{v_2}(1-r_{33})+\theta+s\right)\left(\mu_b(1-r_{33})+s\right)}-\frac{\mu_{v_2}^2\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(\mu_{v_2}s-\mu_b(\theta+s)\right)}\right.\\
&\quad\times\frac{(\mu_b-\mu_{v_2})s}{\left((1-r_{22})\mu_{v_2}+\theta+s\right)\left((1-r_{33})\mu_{v_2}+\theta+s\right)}-\frac{\mu_b^2\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}}{\left(\theta+r_{23}\mu_{v_2}\right)\left(\mu_{v_2}s-\mu_b(\theta+s)\right)}\\
&\quad\left.\times\frac{\mu_b\theta}{\left((1-r_{22})\mu_b+s\right)\left((1-r_{33})\mu_b+s\right)}\right)\frac{\left(\theta+r_{23}\mu_{v_2}\right)(1-r_{22})(1-r_{33})}{\left(\theta+r_{23}\mu_{v_2}\right)(1-r_{22})+r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}
\end{aligned}
\tag{4-159}
$$

从而得到 $E[W_2]=-\widetilde{W}_2^{*\prime}(0)$

$$
\begin{aligned}
E[W_2]&=\left(\frac{\mu_{v_2}\left(\theta+r_{23}\mu_{v_2}\right)(1-r_{22})(1-r_{33})}{\left(\mu_{v_2}(1-r_{33})+\theta\right)^2}+\frac{\theta\left(\theta+r_{23}\mu_{v_2}\right)(1-r_{22})\left((1-r_{33})(\mu_b+\mu_{v_2})+\theta\right)}{\left(\mu_{v_2}(1-r_{33})+\theta\right)^2\left(\mu_b(1-r_{33})\right)}\right.\\
&\quad-\frac{\mu_{v_2}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}(\mu_b-\mu_{v_2})(1-r_{22})(1-r_{33})}{\mu_b\theta\left(\mu_{v_2}(1-r_{22})+\theta\right)\left(\mu_{v_2}(1-r_{33})+\theta\right)}+\frac{\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}(\mu_b-\mu_{v_2})}{\mu_b\theta}\\
&\quad\left.+\frac{\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)r_{23}(2-r_{22}-r_{33})}{\mu_b(1-r_{22})(1-r_{33})}\right)\left((1-r_{22})\left(\theta+r_{23}\mu_{v_2}\right)+r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)\right)^{-1}
\end{aligned}
\tag{4-160}
$$

因此，根据式（4-157），可知在第一阶段工作休假期内到达的顾客的平均剩余效用为 $U(1; q_1) = R-CE[W_1]$，且与 q_0 和 q_2 无关。根据式（4-160），可知在第二阶段工作休假期到达的顾客的平均剩余效用为 $U(2; q_1, q_2) = R-CE[W_2]$，且只与 q_1 和 q_2 有关，而与 q_0 无关。再根据式（4-154），可知在忙期到达的顾客的平均剩余效用为 $U(0; q_0, q_1, q_2) = R-CE[W_0]$，且与 q_0, q_1 和 q_2 均相关。求解方程组 $U(1; q_1) = 0$, $U(2; q_1, q_2) = 0$ 和 $U(0; q_0, q_1, q_2) = 0$，可得到其根为$(q^*(0), q^*(1), q^*(2))$。至于$(q^*(0), q^*(1), q^*(2))$的唯一性问题，由于其表达式的复杂性，我们只通过数值方法观察到对于固定的 λ，的确只存在唯一的可行非负解。这说明大批顾客的到

达只具有负效应。因此, 在几乎不可视情形下, 顾客的均衡止步策略为$(q_e(0), q_e(1), q_e(2)) = (\min\{q^*(0), 1\}, \min\{q^*(1), 1\}, \min\{q^*(2), 1\})$。

虽然$\mu_{v_1} < \mu_b$且$\mu_{v_2} < \mu_b$，但是图 4-19 表明$q_e(2)$最低，且$q_e(0)$不一定比$q_e(1)$高。原因在于μ_{v_1}相比μ_b并不算太低，因此能够满足顾客服务需求。而且，尽管处于休假期，较短的队长也会吸引顾客进入。然而，在第二阶段工作休假期内，由于μ_{v_2}太低，且队列中可能包括在上一阶段工作休假期内未被服务的顾客，所以$q_e(2)$最低。当然, $q_e(2) > q_e(0)$或$q_e(1) < q_e(2)$也可能发生, 这取决于μ_b，μ_{v_1}和μ_{v_2}的取值。

除了图 4-19，图 4-20 至图 4-22 分别表示的是顾客的均衡止步策略关于参数θ、μ_{v_1}和μ_{v_2}的敏感度。图 4-20 表明$q_e(1)$和$q_e(2)$均随θ递增，但$q_e(1)$的增速远高于$q_e(2)$(<0.5)，这是由$\mu_{v_1} = 2\mu_{v_2}$以及第一阶段工作休假期内积压的顾客导致的。图 4-21 表明$q_e(1)$随μ_{v_1}递增，但$q_e(2)$随μ_{v_1}递减。而且，虽然$\mu_{v_1} < \mu_{v_2}$，但$q_e(1)$可能大于$q_e(2)$，这是因为在第一阶段工作休假期内具有较短的平均队长。类似于图 4-20，图 4-22 表明$q_e(1)$和$q_e(2)$均随μ_{v_2}递增。至于$q_e(0)$，图 4-20 至图 4-22 均表明$q_e(0)$几乎不变，且如果休假环境足够好的话，$q_e(0)$不一定大于$q_e(1)$或$q_e(2)$。

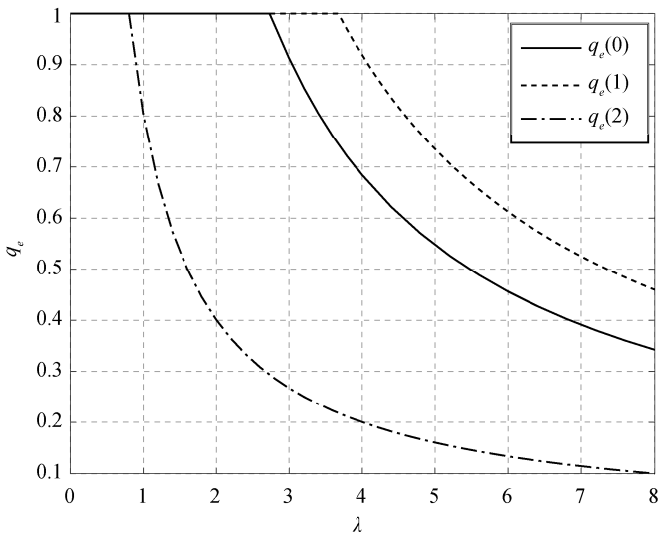

图 4-19　几乎不可视情形下的均衡混合止步策略

（$R = 10,\ C = 1,\ \mu_b = 3,\ \mu_{v_1} = 2,\ \mu_{v_2} = 1,\ \theta = 0.1$）

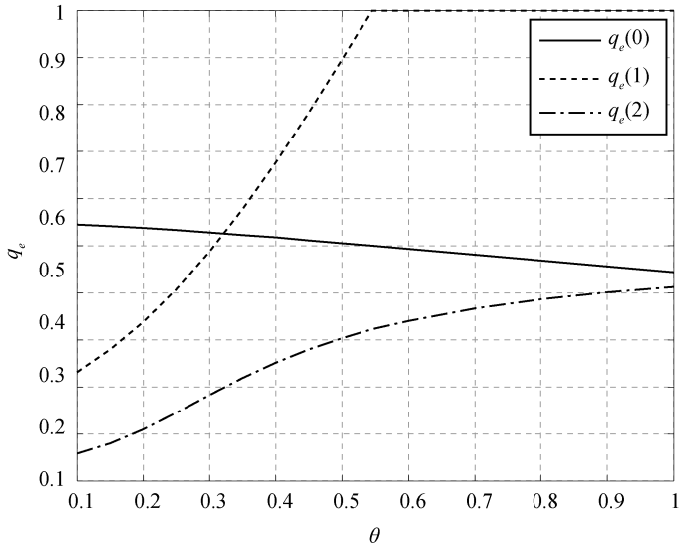

图 4-20　几乎不可视情形下的均衡混合止步策略关于 θ 的敏感度

（$R=5, C=1, \mu_b=3, \lambda=5, \mu_{v_1}=1, \mu_{v_2}=0.5$）

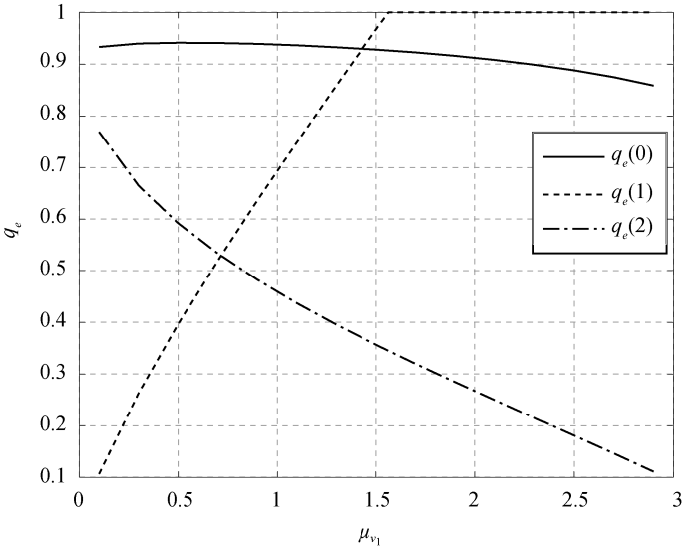

图 4-21　几乎不可视情形下的均衡混合止步策略关于 μ_{v_1} 的敏感度

（$R=10, C=1, \mu_b=3, \lambda=3, \mu_{v_1}=1, \theta=0.1$）

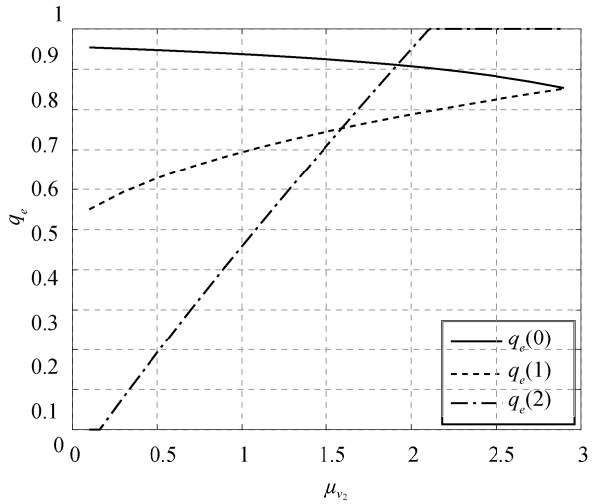

图 4-22　几乎不可视情形下的均衡混合止步策略关于 μ_{v_2} 的敏感度

（$R=10, C=1, \mu_b=3, \lambda=3, \mu_{v_2}=1, \theta=0.1$）

然后，考虑均衡状态下的单位社会福利。根据式（4-147），系统无条件平均队长为

$$
\begin{aligned}
E\left[L_{po}\right] &= \sum_{k=0}^{\infty} k\left(\pi_{k0}+\pi_{k1}+\pi_{k2}\right) \\
&= \frac{\lambda q_0 \pi_{00}}{\theta}\left(\frac{\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{11}\right)^2\left(1-r_{22}\right)^2}\left(r_{23}r_{31}r_{22}\left(2-r_{11}-r_{22}\right)\left(2-r_{11}-r_{22}-r_{33}\right)\right.\right. \\
&\quad \left.+r_{21}\left(1-r_{11}r_{22}\right)\right)+\frac{\theta r_{11}}{\lambda q_0\left(1-r_{11}\right)^2}+\frac{r_{31}\left(1-r_{11}r_{33}\right)}{\left(1-r_{11}\right)^2\left(1-r_{33}\right)^2}+\frac{r_{22}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{22}\right)^2} \\
&\quad \left.+\frac{r_{33}}{\left(1-r_{33}\right)^2}+\frac{r_{23}\left(\theta+\lambda q_2-r_{33}\mu_{v_2}\right)\left(1-r_{22}r_{33}\right)}{\left(\theta+r_{23}\mu_{v_2}\right)\left(1-r_{22}\right)^2\left(1-r_{33}\right)^2}\right)
\end{aligned}
\tag{4-161}
$$

因此，单位社会福利为 $U_s(q_0, q_1, q_2)=\lambda(p_0q_0+p_1q_1+p_2q_2)R-CE[L_{po}]$。在均衡止步策略$(q_e(0), q_e(1), q_e(2))$下，均衡社会福利为 $U_s(q_e(0), q_e(1), q_e(2))$。与 $U_s(n_e(0), n_e(1), n_e(2))$类似，图 4-23 表明 $U_s(q_e(0), q_e(1), q_e(2))$仍随 λ 先增后减。

4.3.4　完全不可视排队系统

最后，讨论完全不可视排队系统，即顾客在时刻 t 观测不到服务员状态 $I(t)$

和队长 $N(t)$。因此，顾客的决策相当于选择一个进入概率 q $(0\leq q\leq 1)$，且顾客的均衡混合止步策略可记为 q_e。状态转移图如图 4-24 所示。

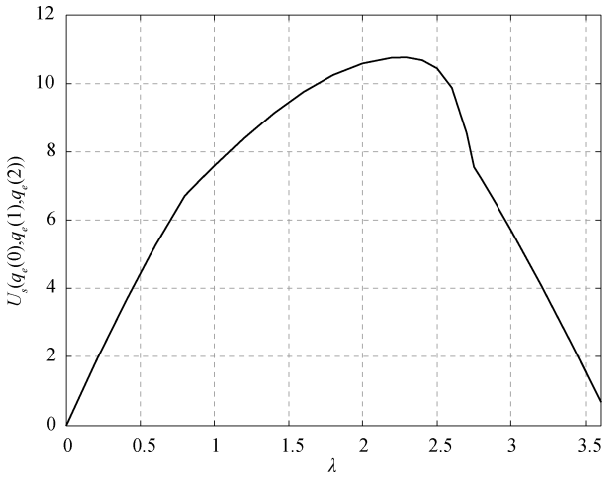

图 4-23　几乎不可视情形下的均衡社会福利 $(R=10,C=1,\mu_b=3,\mu_{v_1}=2,\mu_{v_2}=1,\theta=0.1)$

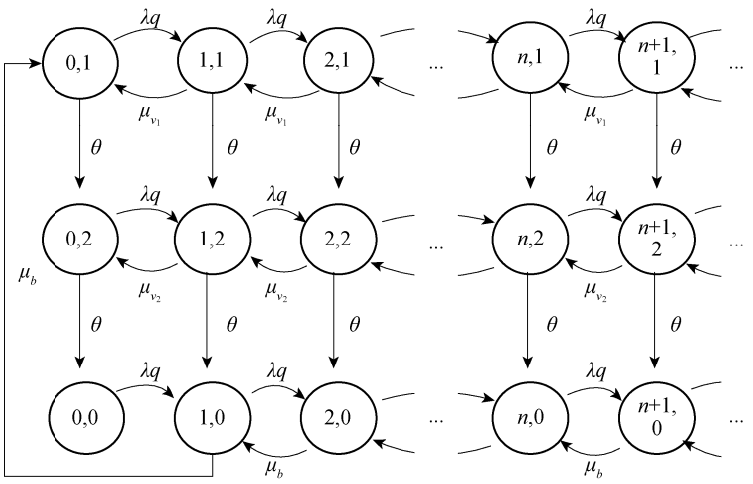

图 4-24　完全不可视排队系统的状态转移图

在式（4-161）中，将所有的 $q_i(i = 0, 1, 2)$替换为 q，可得到完全不可视情形下的稳态平均队长 $E[L_{uo}]$。因此，顾客的稳态平均逗留时间为

$$
\begin{aligned}
E[W] &= \frac{E[L_{uo}]}{\lambda q} \\
&= \frac{\pi_{00}}{\theta}\left(\frac{(\theta+\lambda q-r_{33}\mu_{v_2})}{(\theta+r_{23}\mu_{v_2})(1-r_{11})^2(1-r_{22})^2}(r_{23}r_{31}r_{22}(2-r_{11}-r_{22})(2-r_{11}-r_{22}-r_{33})\right. \\
&\quad \left.+r_{21}(1-r_{11}r_{22}))+\frac{\theta r_{11}}{\lambda q(1-r_{11})^2}+\frac{r_{31}(1-r_{11}r_{33})}{(1-r_{11})^2(1-r_{33})^2}+\frac{r_{22}(\theta+\lambda q-r_{33}\mu_{v_2})}{(\theta+r_{23}\mu_{v_2})(1-r_{22})^2}\right. \\
&\quad \left.+\frac{r_{33}}{(1-r_{33})^2}+\frac{r_{33}(\theta+\lambda q-r_{33}\mu_{v_2})(1-r_{22}r_{33})}{(\theta+r_{23}\mu_{v_2})(1-r_{22})^2(1-r_{33})^2}\right)
\end{aligned}
\tag{4-162}
$$

因此，顾客的平均剩余效用为 $U(q)=R\text{-}CE[W]$。求解 $U(q)=0$，可得到其根为 q^*，因此顾客的均衡止步策略为 $q_e=\min\{q^*,1\}$。类似于几乎不可视情形，如果 $\rho=\lambda q/\mu_b<1$，我们也观察到 $U(q)=0$ 具有唯一正根。图 4-25 表明 q_e 随 λ 递减，而图 4-26 至图 4-28 表明 q_e 均随 θ、μ_{v_1} 和 μ_{v_2} 递增。

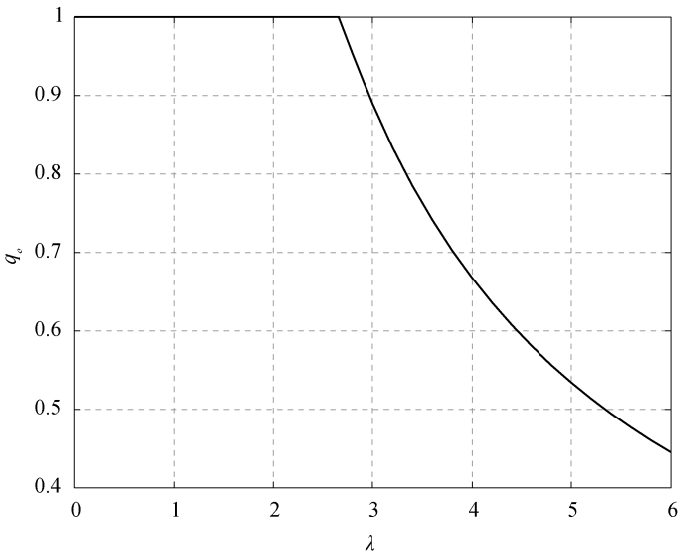

图 4-25　完全不可视情形下的均衡混合策略
（$R=10, C=1, \mu_b=3, \mu_{v_1}=2, \mu_{v_2}=1, \theta=0.1$）

根据式（4-162），得到单位社会福利为 $U_s(q)=\lambda q(R\text{-}CE[W])$。在均衡止步策略 q_e 下，均衡社会福利为 $U_s(q_e)$。类似于 $U_s(n_e(0), n_e(1), n_e(2))$和 $U_s(q_e(0), q_e(1), q_e(2))$，图 4-29 表明 $U_s(q_e)$仍随 λ 先增后减。

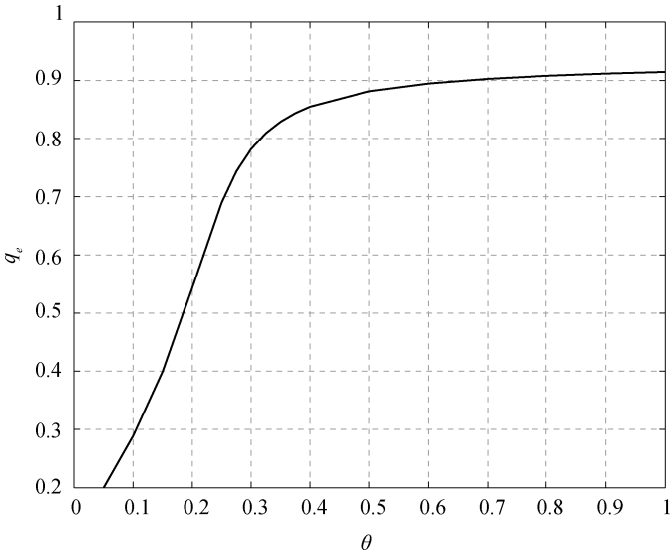

图 4-26　完全不可视情形下的均衡混合策略关于 θ 的敏感度
（$R=5, C=1, \mu_b=3, \lambda=3, \mu_{v_1}=1, \mu_{v_2}=0.5$）

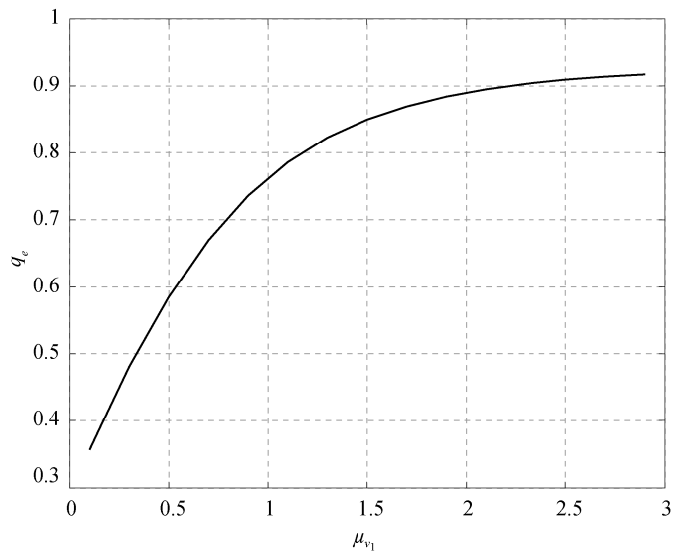

图 4-27　完全不可视情形下的均衡混合策略关于 μ_{v_1} 的敏感度
（$R=10, C=1, \mu_b=3, \lambda=3, \mu_{v_2}=1, \theta=0.1$）

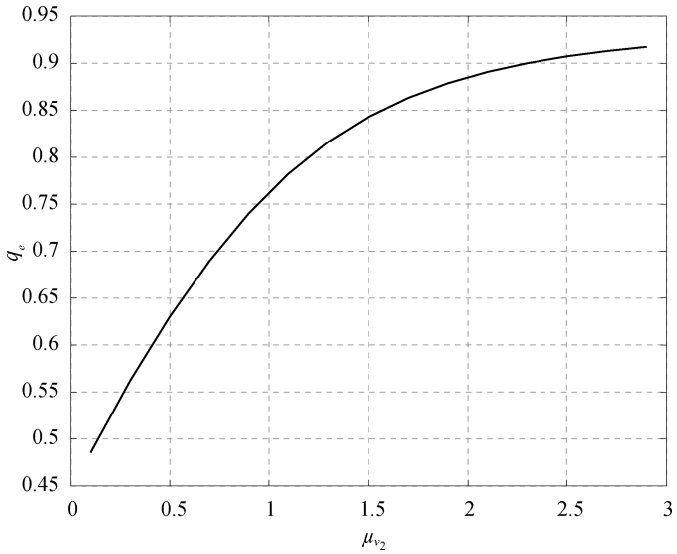

图 4-28　完全不可视情形下的均衡混合策略关于 μ_{v_2} 的敏感度
$(R=10, C=1, \mu_b=3, \lambda=3, \mu_{v_2}=1, \theta=0.1)$

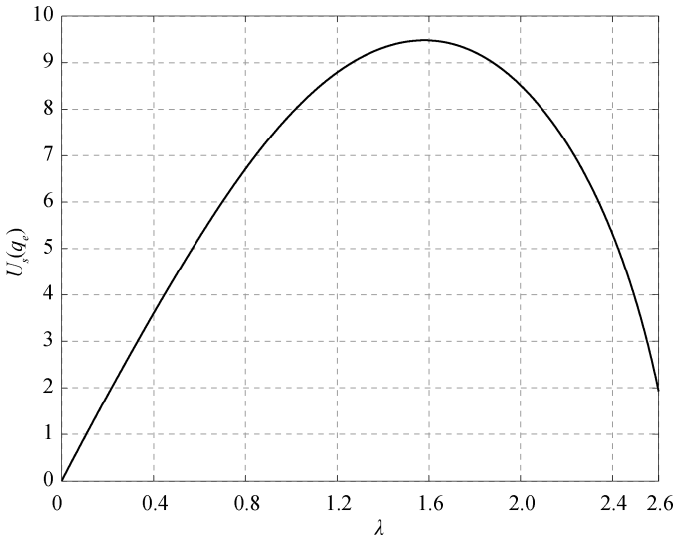

图 4-29　完全不可视情形下的均衡社会福利
$(R=10, C=1, \mu_b=3, \mu_{v_1}=2, \mu_{v_2}=1, \theta=0.1)$

4.4　具有双级工作休假策略的排队系统

本节将研究具有双级工作休假策略的几乎不可视与完全不可视排队系统中顾客的均衡和最优止步策略，并导出队长和等待时间的随机分解结果以及均衡和最优社会福利。

4.4.1　模型描述

考虑一些具有双级工作休假策略的 Markov 单服务员排队系统。假定顾客的潜在到达率为 Λ，服务员的正常服务率为 μ_b。一旦系统空竭，服务员开始第一级工作休假期 V_1，V_1 服从参数为 θ 的指数分布。在此期间，顾客以服务率 μ_{v_1} 接受服务，且 $\mu_{v_1} < \mu_b$。休假结束后，如仍有顾客等待，则开始一个新的忙期。否则，进行第二级工作休假 V_2，V_2 与 V_1 具有相同的分布，且服务率由 μ_{v_1} 转换为 μ_{v_2}，一般 $\mu_{v_2} < \mu_{v_1}$。第二级工作休假结束后，如果有顾客等待，则进入忙期，否则进入闲期等待顾客到达。因此，我们记该类排队系统为 M/M/1/(D-A)WV 排队系统。

令$(N(t), I(t))$表示时刻 t 时的系统状态, 其中 $N(t)$表示时刻 t 时的系统队长, $I(t)$表示时刻 t 时的服务员状态，且

$$I(t)\begin{cases}0, & \text{服务员处于忙期或闲期}\\ 1, & \text{服务员处于第一级工作休假期间}\\ 2, & \text{服务员处于第二级工作休假期间}\end{cases}$$

在此，标记一个到达顾客，服务完成可获得服务收益 R，但也需承担单位等待损耗 C。我们采用线性等待损耗函数，且到达过程、服务过程以及两级休假时间相互独立，服务顺序为先到先服务。

4.4.2　几乎不可视排队系统

首先，考虑几乎不可视排队系统中顾客的均衡和最优止步策略。由于顾客可观测到服务员状态，因此，他们的决策问题可表示为一组实际到达率 $(\lambda_0, \lambda_1, \lambda_2)$ $(\lambda_i \le \Lambda, i = 0,1,2)$。在此，分别记顾客的均衡止步策略为 $(\lambda_e(0), \lambda_e(1), \lambda_e(2))$，最优止步策略为 $(\lambda_s(0), \lambda_s(1), \lambda_s(2))$。

为了得到 $(\lambda_e(0), \lambda_e(1), \lambda_e(2))$，首先导出稳态队长分布。状态转移图如图 4-30 所示。根据图 4-30，可知过程$\{N(t), I(t)\}$为一个拟生灭过程，且状态空间为

$$\Omega=\{(k,j):k\geq 0,j=0,1,2\}$$

如果 $\rho=\lambda_0/\mu_b<1$，令(N,I)为过程$\{N(t),I(t)\}$的稳态极限。记稳态队长分布为

$$\boldsymbol{\pi}=(\boldsymbol{\pi}_0,\boldsymbol{\pi}_1,\boldsymbol{\pi}_2\cdots),\quad \boldsymbol{\pi}_k=(\pi_{k0},\pi_{k1},\pi_{k2}),\quad k\geq 0$$

$$\pi_{kj}=P\{N=k,I=j\}=\lim_{t\to\infty}P\{N(t)=k,I(t)=j\},\ (k,j)\in\Omega$$

最小生成元为 $\boldsymbol{Q}$，从而得到状态转移概率方程为 $\boldsymbol{\pi Q}=\boldsymbol{0}$，即

$$\pi_{00}\lambda_0=\pi_{02}\theta \tag{4-163}$$

$$\pi_{k0}(\lambda_0+\mu_b)=\pi_{k-10}\lambda_0+\pi_{k+10}\mu_b+\pi_{k1}\theta+\pi_{k2}\theta,\ k\geq 1 \tag{4-164}$$

$$\pi_{01}(\lambda_1+\theta)=\pi_{10}\mu_b+\pi_{11}\mu_{v_1} \tag{4-165}$$

$$\pi_{k1}(\lambda_1+\theta+\mu_{v_1})=\pi_{k-11}\lambda_1+\pi_{k+11}\mu_{v_1},\ k\geq 1 \tag{4-166}$$

$$\pi_{02}(\lambda_2+\theta)=\pi_{01}\theta+\pi_{12}\mu_{v_2} \tag{4-167}$$

$$\pi_{k2}(\lambda_2+\theta+\mu_{v_2})=\pi_{k-12}\lambda_2+\pi_{k+12}\mu_{v_2},\ k\geq 1 \tag{4-168}$$

应用字典序，$\boldsymbol{Q}$ 可表示为分块矩阵的形式：

$$\boldsymbol{Q}\begin{pmatrix}\boldsymbol{A}_{00} & \boldsymbol{C} & & \\ \boldsymbol{B}_{10} & \boldsymbol{A} & \boldsymbol{C} & \\ & \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{C} \\ & & \ddots & \ddots & \ddots\end{pmatrix} \tag{4-169}$$

其中

$$\boldsymbol{A}_{00}=\begin{pmatrix}-\lambda_0 & 0 & 0\\ 0 & -(\theta+\lambda_1) & \theta\\ \theta & 0 & -(\theta+\lambda_2)\end{pmatrix},\quad \boldsymbol{B}_{10}=\begin{pmatrix}0 & \mu_b & 0\\ 0 & \mu_{v_1} & 0\\ 0 & 0 & \mu_{v_2}\end{pmatrix}$$

$$\boldsymbol{B}=\begin{pmatrix}\mu_b & 0 & 0\\ 0 & \mu_{v_1} & 0\\ 0 & 0 & \mu_{v_2}\end{pmatrix},\quad \boldsymbol{C}=\begin{pmatrix}\lambda_0 & 0 & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_2\end{pmatrix}$$

$$\boldsymbol{A}=\begin{pmatrix}-(\lambda_0+\mu_b) & 0 & 0\\ \theta & -(\lambda_1+\theta+\mu_{v_1}) & 0\\ \theta & 0 & -(\lambda_2+\theta+\mu_{v_2})\end{pmatrix}$$

现在需要求解以下矩阵二次方程

$$\boldsymbol{R}^2\boldsymbol{B}+\boldsymbol{R}\boldsymbol{A}+\boldsymbol{C}=\boldsymbol{0} \tag{4-170}$$

的最小非负解，即率阵 $\boldsymbol{R}$。鉴于 $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$ 的特殊结构，$\boldsymbol{R}$ 必将具有类似的结构。因此，假设

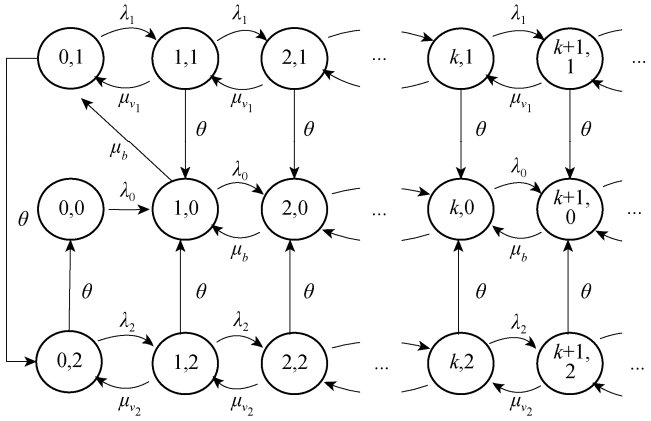

图 4-30　几乎不可视排队系统的状态转移图

$$\boldsymbol{R}=\begin{pmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \tag{4-171}$$

将其代入式（4-170），得到方程组

$$\begin{cases} r_{11}^2\mu_b - r_{11}(\lambda_0+\mu_b) +\lambda_0 = 0 \\ (r_{21}r_{11}+r_{22}r_{21})\mu_b - r_{21}(\lambda_0+\mu_b)+r_{22}\theta = 0 \\ r_{22}^2\mu_{v_1} - r_{22}(\lambda_1+\theta+\mu_{v_1})+\lambda_1 = 0 \\ (r_{11}r_{31}+r_{32}r_{21}+r_{33}r_{31})\mu_b - r_{31}(\lambda_0+\mu_b)+r_{32}\theta + r_{33}\theta = 0 \\ (r_{32}r_{22}+r_{33}r_{32})\mu_{v_1} - r_{32}(\lambda_1+\theta+\mu_{v_1})=0 \\ r_{33}^2\mu_{v_2} - r_{33}(\lambda_2+\theta+\mu_{v_2})+\lambda_2 = 0 \end{cases} \tag{4-172}$$

求解方程组（4-172），得到

$$\boldsymbol{R}=\begin{pmatrix} \rho & 0 & 0 \\ \dfrac{\theta r_1}{\mu_b(1-r_1)} & r_1 & 0 \\ \dfrac{\theta r_2}{\mu_b(1-r_2)} & 0 & r_2 \end{pmatrix} \tag{4-173}$$

其中

$$\begin{cases} r_1 = \dfrac{\lambda_1+\theta+\mu_{v_1}-\sqrt{(\lambda_1+\theta+\mu_{v_1})^2-4\lambda_1\mu_{v_1}}}{2\mu_{v1}} \\ r_2 = \dfrac{\lambda_2+\theta+\mu_{v_2}-\sqrt{(\lambda_2+\theta+\mu_{v_2})^2-4\lambda_2\mu_{v_2}}}{2\mu_{v_2}} \end{cases} \tag{4-174}$$

显然，$0<r_i<1(i=1,2)$。应用矩阵解析方法，得到

$$\boldsymbol{\pi}_k=(\pi_{k0},\pi_{k1},\pi_{k2})=(\pi_{00},\pi_{01},\pi_{02})\boldsymbol{R}^k,k\geq 0 \tag{4-175}$$

且$(\pi_{00},\pi_{01},\pi_{02})$满足

$$(\pi_{00},\pi_{01},\pi_{02})(\boldsymbol{RB}_{10}+\boldsymbol{A}_{00})=\boldsymbol{0} \tag{4-176}$$

平衡方程为

$$(\pi_{00},\pi_{01},\pi_{02})(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e}=1 \tag{4-177}$$

其中 $\boldsymbol{I}$ 为单位阵，$\boldsymbol{e}$ 为单位向量，且

$$\boldsymbol{RB}_{10}+\boldsymbol{A}_{00}=\begin{pmatrix}-\lambda_0 & r_{11}\mu_b & 0\\ 0 & r_{21}\mu_b+r_{22}\mu_{v_1}-(\theta+\lambda_1) & \theta\\ \theta & r_{31}\mu_b+r_{32}\mu_{v_1} & r_{33}\mu_{v_2}-(\theta+\lambda_2)\end{pmatrix} \tag{4-178}$$

将式（4-178）代入式（4-176），得到

$$\begin{cases}-\lambda_0\pi_{00}+\theta\pi_{02}=0\\ r_{11}\mu_b\pi_{00}+(r_{21}\mu_b+r_{22}\mu_{v_1}-(\theta+\lambda_1))\pi_{01}+(r_{31}\mu_b+r_{32}\mu_{v_1})\pi_{02}=0\\ \theta\pi_{01}+(r_{33}\mu_{v_2}-(\theta+\lambda_2))\pi_{02}=0\end{cases} \tag{4-179}$$

根据式（4-173），求解方程组（4-179）得到

$$(\pi_{00},\pi_{01},\pi_{02})=\pi_{00}\left(1,\frac{\lambda_0(\lambda_2+\theta-r_2\mu_{v_2})}{\theta^2},\frac{\lambda_0}{\theta}\right)$$

而且，

$$\boldsymbol{R}^k=\begin{pmatrix}r_{11}^k & 0 & 0\\ r_{21}\sum_{j=0}^{k-1}r_{11}^j r_{22}^{k-1-j} & r_{22}^k & 0\\ r_{31}\sum_{j=0}^{k-1}r_{11}^j r_{33}^{k-1-j}+r_{21}r_{32}\sum_{i=0}^{k-2}\sum_{j=0}^{k-2-i}r_{11}^i r_{22}^j r_{33}^{k-2-j-i} & r_{32}\sum_{j=0}^{k-1}r_{22}^j r_{33}^{k-1-j} & r_{33}^k\end{pmatrix},k\geq 1$$

将$(\pi_{00},\pi_{01},\pi_{02})$和 $\boldsymbol{R}^k$ 代入式（4-175），得到稳态队长分布为

$$\begin{cases}\pi_{k0}=\left(\rho^k+\dfrac{\rho r_1(\lambda_2+\theta-r_2\mu_{v_2})(\rho^k-r_1^k)}{\theta(1-r_1)(\rho-r_1)}+\dfrac{\rho r_2(\rho^k-r_2^k)}{(1-r_2)(\rho-r_2)}\right)\pi_{00},\ k\geq 0\\ \pi_{k1}=\dfrac{\lambda_0(\lambda_2+\theta-r_2\mu_{v_2})}{\theta^2}r_1^k\pi_{00},\ k\geq 0\\ \pi_{k2}=\dfrac{\lambda_0}{\theta}r_2^k\pi_{00},\ k\geq 0\end{cases} \tag{4-180}$$

其中 π_{00} 可通过求解方程（4-177）得到，且

$$\pi_{00}=(1-r_1)(1-r_2)(1-\rho)\theta^2\left(\left((1-r_1)(1-r_2)+\frac{\rho r_2(1-r_1)}{(1-r_2)}+\frac{\rho r_1(1-r_2)}{1-r_1}\right)\theta^2\right. \tag{4-181}$$
$$\left.+\left(\lambda_0(1-\rho)(2-r_1-r_2)+\frac{\rho r_1(1-r_2)}{1-r_1}(\lambda_2-r_2\mu_{v_2})\right)\theta+\lambda_0(1-r_2)(1-\rho)(\lambda_2-r_2\mu_{v_2})\right)^{-1}$$

根据式（4-180），系统处于状态 $i(i=0,1,2)$的概率 p_i 分别为

$$\begin{cases}p_0=\sum_{k=0}^{\infty}\pi_{k0}=\dfrac{1}{1-\rho}\left(1+\dfrac{\rho r_1(\lambda_2+\theta-r_2\mu_{v_2})}{\theta(1-r_1)^2}+\dfrac{\rho r_2}{(1-r_2)^2}\right)\pi_{00}\\ p_1=\sum_{k=0}^{\infty}\pi_{k1}=\dfrac{\lambda_0(\lambda_2+\theta-r_2\mu_{v_2})}{\theta^2(1-r_1)}\pi_{00}\\ p_2=\sum_{k=0}^{\infty}\pi_{k2}=\dfrac{\lambda_0}{\theta(1-r_2)}\pi_{00}\end{cases} \tag{4-182}$$

为了得到顾客在忙期的条件平均逗留时间，记为 $E[W_0]$，首先得到顾客逗留时间的 Laplace 变换为

$$W_0^*(s)=\sum_{k=0}^{\infty}\pi_{k0}\left(\frac{\mu_b}{\mu_b+s}\right)^{k+1}=\left(\frac{1}{\mu_b(1-\rho)+s}+\frac{\lambda_0 r_1(\lambda_2+\theta-r_2\mu_{v_2})}{\theta(1-r_1)(\mu_b(1-\rho)+s)(\mu_b(1-r_1)+s)}\right.$$
$$\left.+\frac{\lambda_0 r_2}{(1-r_2)(\mu_b(1-\rho)+s)(\mu_b(1-r_2)+s)}\right)\mu_b\pi_{00} \tag{4-183}$$

因此，顾客在忙期的条件逗留时间的 Laplace 变换为

$$\tilde{W}_0^*(s)=\frac{1}{p_0}W_0^*(s) \tag{4-184}$$

从而得到 $E[W_0]-\tilde{W}_0^{*\prime}(0)$，即

$$E[W_0]=\frac{1}{1-\rho}\left(1+\frac{\rho r_1(\lambda_2+\theta-r_2\mu_{v_2})(2-\rho-r_1)}{\theta(1-r_1)^3}+\frac{\rho r_2(2-\rho-r_2)}{(1-r_2)^3}\right)$$
$$\times\left(\mu_b\left(1+\frac{\rho r_1(\lambda_2+\theta-r_2\mu_{v_2})}{\theta(1-r_1)^2}+\frac{\rho r_2}{(1-r_2)^2}\right)\right)^{-1} \tag{4-185}$$

对于顾客在第一级工作休假期内的条件平均逗留时间，记为 $E[W_1]$，我们需要讨论在状态$(k, 1)(k\geq 0)$下的两种情形：首先标记一个在状态$(k, 1)$到达的顾客。如果剩余休假时间内足够服务完 $k+1$（包括他自己）个顾客，顾客的逗留时间为 $k+1$ 个参数为 μ_{v_1} 的服务时间之和；否则，如果只有 $j(0\leq j\leq k)$个顾客在剩余休假期内完成服务，顾客的逗留时间为剩余休假时间与 $k+1-j$ 个参数为 μ_b 的服务时间

之和。因此，得到顾客逗留时间的 Laplace 变换为

$$\begin{aligned} W_1^*(s) &= \frac{\theta}{\mu_{v_1}+\theta}\sum_{k=0}^{\infty}\pi_{k0}\sum_{j=0}^{k}\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta}\right)^j\left(\frac{\mu_{v_1}+\theta}{\mu_{v_1}+\theta+s}\right)^{j+1}\left(\frac{\mu_b}{\mu_b+s}\right)^{k+1-j} \\ &\quad +\sum_{k=0}^{\infty}\pi_{k0}\left(\frac{\mu_{v_1}}{\mu_{v_1}+\theta}\right)^{k+1}\left(\frac{\mu_{v_1}+\theta}{\mu_{v_1}+\theta+s}\right)^{k+1} \\ &= \frac{\pi_{00}\lambda_0(\lambda_2+\theta-r_2\mu_{v_2})}{\theta^2(\mu_{v_1}(1-r_1)+\theta+s)}\left(\mu_{v_1}+\frac{\theta\mu_b}{\mu_b(1-r_1)+s}\right) \end{aligned} \tag{4-186}$$

因此，顾客在第一级工作休假期内的条件逗留时间的 Laplace 变换为 $\tilde{W}_1^*(s)=W_1^*(s)/p_1$。从而得到 $E[W_1]=-\tilde{W}_1^{*\prime}(0)$，即

$$E[W_1]=\frac{\mu_b(1-r_1)+\theta}{\mu_b(1-r_1)(\mu_{v_1}(1-r_1)+\theta)} \tag{4-187}$$

另一方面，对于顾客在第二级工作休假期内的条件平均逗留时间，记为 $E[W_2]$，我们可类似讨论在状态$(k,\ 2)(k\geq 0)$下的两种情形。因此，顾客逗留时间的 Laplace 变换为

$$W_2^*(s)=\frac{\pi_{00}\lambda_0}{\theta(\mu_{v_2}(1-r_2)+\theta+s)}\left(\mu_{v_2}+\frac{\theta\mu_b}{\mu_b(1-r_2)+s}\right) \tag{4-188}$$

顾客在第二级工作休假期内的条件逗留时间的 Laplace 变换为 $\tilde{W}_2^*(s)=W_2^*(s)/p_2$。从而得到 $E[W_2]=-\tilde{W}_1^{*\prime}(0)$，即

$$E[W_2]=\frac{\mu_b(1-r_2)+\theta}{\mu_b(1-r_2)(\mu_{v_2}(1-r_2)+\theta)} \tag{4-189}$$

因此，根据式（4-187），顾客在第一级工作休假期内的平均剩余效用为 $U_1(\lambda_1)=R-CE[W_1]$，即

$$U_1(\lambda_1)=R-C\left(\frac{\mu_b(1-r_1)+\theta}{\mu_b(1-r_1)(\mu_{v_1}(1-r_1)+\theta)}\right) \tag{4-190}$$

可见其只与 λ_1 有关。类似地，根据式（4-189），顾客在第二级工作休假期内的平均剩余效用为 $U_2(\lambda_2)=R-CE[W_2]$，即

$$U_2(\lambda_2)=R-C\left(\frac{\mu_b(1-r_2)+\theta}{\mu_b(1-r_2)(\mu_{v_2}(1-r_2)+\theta)}\right) \tag{4-191}$$

可见其只与 λ_2 有关。再根据式（4-185），顾客在忙期的平均剩余效用为 $U_0(\lambda_0,\lambda_1,\lambda_2)=R-CE[W_0]$，即

$$U_0(\lambda_0,\lambda_1,\lambda_2)=R-\frac{C}{\mu_b-\lambda_0}\left(1+\frac{\rho r_1(\lambda_2+\theta-r_2\mu_{v_2})(2-\rho-r_1)}{\theta(1-r_1)^3}+\frac{\rho r_2(2-\rho-r_2)}{(1-r_2)^3}\right)\times\left(1+\frac{\rho r_1(\lambda_2+\theta-r_2\mu_{v_2})}{\theta(1-r_1)^2}+\frac{\rho r_2}{(1-r_2)^2}\right)^{-1} \tag{4-192}$$

可见其与和 λ_0，λ_1 和 λ_2 均相关。求解方程组 $U_1(\lambda_1)=0, U_2(\lambda_2)=0$ 和 $U_0(\lambda_0,\lambda_1,\lambda_2)=0$，可得正可行解为 $(\lambda_0^*,\lambda_1^*,\lambda_2^*)$。现在，我们检验 λ_1^* 和 λ_2^* 的唯一性，即 $E[W_i](i=1,2)$ 是否随 λ_i 单调递增。实际上，

$$r_i'(\lambda_i)=\frac{1}{2\mu_{v_i}}\left(1-\frac{\lambda_i+\theta-\mu_{v_i}}{\sqrt{(\lambda_i+\theta+\mu_{v_i})^2-4\lambda_i\mu_{v_i}}}\right)$$

和

$$E[W_i'(r_i)]=\frac{\mu_{v_i}}{(\mu_{v_i}(1-r_i)+\theta)^2}+\frac{\theta(\mu_b\theta+2\mu_b\mu_{v_i}(1-r_i))}{\mu_b^2(1-r_i)^2(\mu_{v_i}(1-r_i)+\theta)^2}$$

都为正，因此 $E[W_i'(\lambda_i)]=E[W_i'(r_i)]r_i'(\lambda_i)>0(i=1,2)$，这说明 λ_1^* 和 λ_2^* 具有唯一性。对于 λ_0^*，鉴于 $E[W_0]$ 表达式的复杂性，我们只做一些数值分析，且观察到 $\lambda_0^*(0<\lambda_0^*<\min\{\mu_b,\Lambda\})$ 不唯一且一般具有两个值，即 $E[W_0]$ 关于 λ_0 严格凸。原因在于大批的顾客到达可以刺激服务员以高服务率提供服务，从而避免休假，因此 λ_0 的增大同时具有正负效应。但是很显然，只有 λ_0 取较大值时才稳定。因此我们只考虑顾客唯一稳定的均衡止步策略 $(\lambda_e(0),\lambda_e(1),\lambda_e(2))$，即 $(\lambda_e(0),\lambda_e(1),\lambda_e(2))=(\min\{\lambda_0^*,\Lambda\},\min\{\lambda_1^*,\Lambda\},\min\{\lambda_2^*,\Lambda\})$。

然后，我们考虑顾客的最优止步策略，以及顾客的均衡和最优社会福利。根据式（4-180），可知无条件平均队长为

$$\begin{aligned}E[L]&=\sum_{k=0}^{\infty}k(\pi_{k0}+\pi_{k1}+\pi_{k2})\\&=\left(\frac{\rho}{(1-\rho)^2}+\frac{\lambda_0 r_1(\lambda_2+\theta-r_2\mu_{v_2})(1-\rho r_1)}{\theta\mu_b(1-r_1)^3(1-\rho)^2}+\frac{\lambda_0 r_2(1-\rho r_2)}{\mu_b(1-r_2)^3(1-\rho)^2}\right.\\&\quad\left.+\frac{\lambda_0}{\theta}\left(\frac{(\lambda_2+\theta-r_2\mu_{v_2})r_1}{\theta(1-r_1)^2}+\frac{r_2}{(1-r_2)^2}\right)\right)\pi_{00}\end{aligned} \tag{4-193}$$

因此，对于实际到达率 $(\lambda_0,\lambda_1,\lambda_2)$ 来说，单位社会福利为

$$U_s(\lambda_0,\lambda_1,\lambda_2)=\overline{\lambda}R-CE[L]=(\lambda_0 p_0+\lambda_1 p_1+\lambda_2 p_2)R-CE[L] \tag{4-194}$$

其中，$\overline{\lambda}=\lambda_0 p_0+\lambda_1 p_1+\lambda_2 p_2$。

在均衡止步策略 $(\lambda_e(0),\lambda_e(1),\lambda_e(2))$ 下，均衡社会福利可表示为 $U_s(\lambda_e(0),$

$\lambda_e(1),\lambda_e(2))$。虽然$\mu_{v_1}<\mu_b$，图 4-31 和图 4-32 表明$\lambda_e(1)$不一定小于$\lambda_e(0)$，甚至$\lambda_e(1)>\mu_b$。原因在于第一级工作休假期间的队长比忙期时的要短。图 4-33 说明$U_s(\lambda_e(0),\lambda_e(1),\lambda_e(2))$是$\Lambda$的凹函数，直至$U_s(\lambda_e(0),\lambda_e(1),\lambda_e(2))$降为零。

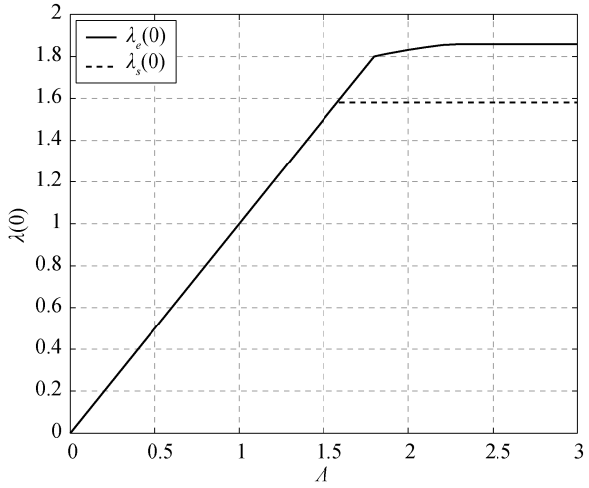

图 4-31　几乎不可视情形下顾客在忙期的均衡和最优到达率
（$R=10,\ C=1,\ \mu_b=2,\ \mu_{v_1}=1,\ \mu_{v_2}=0.5,\ \theta=0.1$）

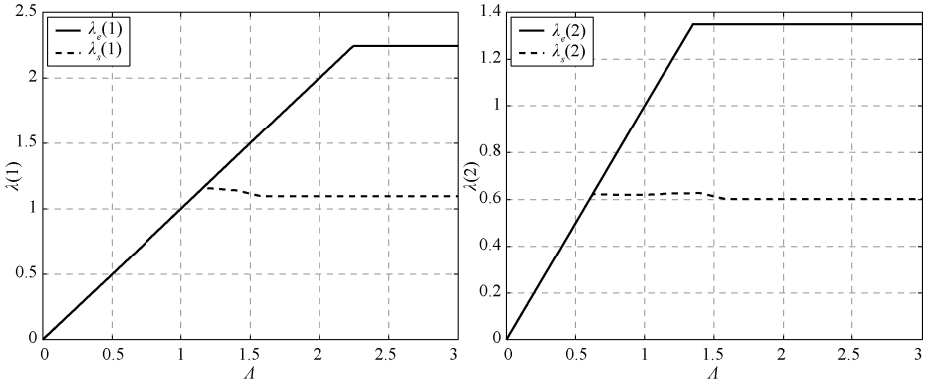

图 4-32　几乎不可视情形下顾客在休假期的均衡和最优到达率
（$R=10,\ C=1,\ \mu b=2,\ \mu_{v1}=1,\ \mu_{v2}=0.5,\ \theta=0.1$）

另一方面，社会计划者希望顾客采用最优止步策略，记为$(\lambda_s(0),\lambda_s(1),\lambda_s(2))$，以最大化社会福利，记为$U_s(\lambda_s(0),\lambda_s(1),\lambda_s(2))$，且$(\lambda_s(0),\lambda_s(1),\lambda_s(2))$可通过求解无约束非线性规划$\max U_s(\lambda_0,\lambda_1,\lambda_2)$（当然，$\lambda_0,\lambda_1,\lambda_2\le\Lambda$）得到。图 4-31 和图 4-32 表明个人最优偏离社会最优。特别地，$\lambda_e(0)$与$\lambda_s(0)$间的差距最小，这是由

于在忙期顾客到达具有正效应的原因。对于最优社会福利来说，图 4-34 表明 $U_s(\lambda_s(0),\lambda_s(1),\lambda_s(2))$ 单增且最终保持恒定。

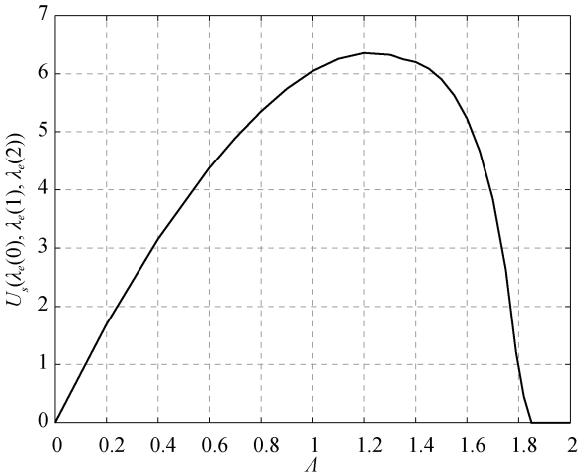

图 4-33　几乎不可视情形下的均衡社会福利
（$R=10$, $C=1$, $\mu_b=2$, $\mu_{v_1}=1.0$, $\mu_{v_2}=0.5$, $\theta=0.1$）

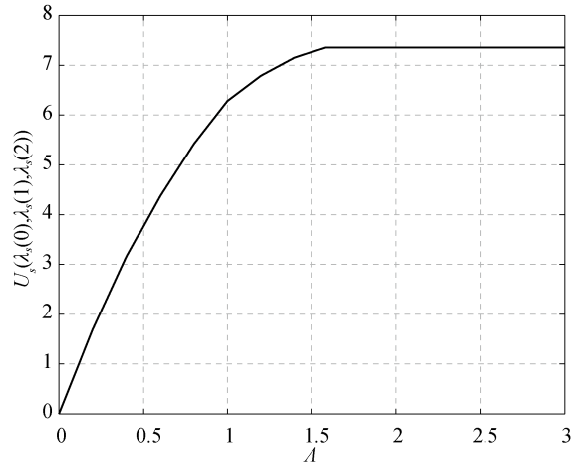

图 4-34　几乎不可视情形下的最优社会福利
（$R=10$, $C=1$, $\mu_b=2$, $\mu_{v_1}=1.0$, $\mu_{v_2}=0.5$, $\theta=0.1$）

4.4.3　完全不可视排队系统

接下来，继续讨论完全不可视情形下顾客的均衡和最优止步策略。由于顾客观测不到系统状态，他们的决策问题相当于选择一个统一的实际进入率

$\lambda(\lambda \le \Lambda)$。他们的均衡和最优止步策略分别记为 λ_e 和 λ_s，状态转移图如图 4-35 所示。

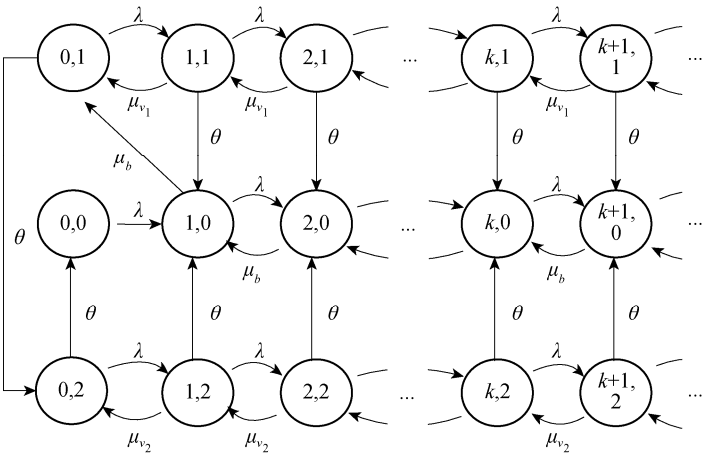

图 4-35　完全不可视排队系统的状态转移图

现在，首先导出系统稳态平均队长和顾客的平均逗留时间。同时，队长和逗留时间的随机分解结果如以下定理所示。

定理 4.4.1　对于具有双级工作休假策略的完全不可视 M/M/1 排队系统，如果 $\rho = \lambda / \mu_b < 1$ 且 $\mu_b > \mu_{v_1}$，$\mu_b > \mu_{v_2}$，则稳态平均队长可分解为两部分：$E[L]=E[L_0]+E[L_d]$，其中

$$E[L_0] = \frac{\rho}{1-\rho} \tag{4-195}$$

为经典 M/M1 排队系统的稳态平均队长，$E[L_d]$为附加队长，且

$$\begin{aligned} E[L_d] = {} & \frac{\lambda(\lambda+\theta-r_2\mu_{v_2})\pi_{00}}{\theta(1-r_1)^2(1-\rho)}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right) \\ & +\frac{\lambda\pi_{00}}{(1-r_2)^2(1-\rho)}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right) \end{aligned} \tag{4-196}$$

其中

$$r_i = \frac{\lambda+\theta+\mu_{v_i}-\sqrt{(\lambda+\theta+\mu_{v_i})^2-4\lambda\mu_{v_i}}}{2\mu_{v_i}}, i=1,2 \tag{4-197}$$

且

$$\pi_{00}=(1-r_1)(1-r_2)(1-\rho)\theta^2\left(\left((1-r_1)(1-r_2)+\frac{\rho r_2(1-r_1)}{1-r_2}+\frac{\rho r_1(1-r_2)}{1-r_1}\right)\theta^2\right.$$
$$\left.+\left(\lambda(1-\rho)(2-r_1-r_2)+\frac{\rho r_1(1-r_2)}{1-r_1}(\lambda-r_2\mu_{v_2})\right)\theta+\lambda(1-r_2)(1-\rho)(\lambda-r_2\mu_{v_2})\right)^{-1} \tag{4-198}$$

证明　在式（4-180）和式（4-181）中，令 $\lambda_0=\lambda_1=\lambda_2=\lambda$,得到稳态队长分布为

$$\begin{cases}\pi_{k0}=\left(\rho^k+\dfrac{\rho r_1(\lambda_2+\theta-r_2\mu_{v_2})(\rho^k-r_1^k)}{\theta(1-r_1)(\rho-r_1)}+\dfrac{\rho r_2(\rho^k-r_2^k)}{(1-r_2)(\rho-r_2)}\right)\pi_{00},k\geq 0\\ \pi_{k1}=\dfrac{\lambda(\lambda+\theta-r_2\mu_{v_2})}{\theta^2}r_1^k\pi_{00},k\geq 0\\ \pi_{k2}=\dfrac{\lambda}{\theta}r_2^k\pi_{00},k\geq 0\end{cases} \tag{4-199}$$

其中 $\rho=\lambda/\mu_b$。因此，得到队长 L 的母函数为

$$\begin{aligned}L(z)&=\sum_{k=0}^{\infty}z^k(\pi_{k0}+\pi_{k1}+\pi_{k2})\\&=\left(\frac{1}{1-\rho z}+\frac{\lambda r_1(\lambda+\theta-r_2\mu_{v_2})}{\theta\mu_b(1-r_1)(\rho-r_1)}\left(\frac{1}{1-\rho z}-\frac{1}{1-r_1z}\right)\right.\\&\quad\left.+\frac{\lambda r_2}{\mu_b(1-r_2)(\rho-r_2)}\left(\frac{1}{1-\rho z}-\frac{1}{1-r_2z}\right)+\frac{\lambda}{\theta}\left(\frac{(\lambda+\theta-r_2\mu_{v_2})}{\theta(1-r_1z)}+\frac{1}{1-r_2z}\right)\right)\pi_{00}\\&=\frac{1-\rho}{1-\rho z}\frac{\pi_{00}}{1-\rho}\left(1+\frac{\lambda(\lambda+\theta-r_2\mu_{v_2})}{\theta^2}+\frac{\lambda}{\theta}\right.\\&\quad+\frac{\lambda(\lambda+\theta-r_2\mu_{v_2})}{\theta(1-r_1)}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)\frac{(1-r_1)z}{1-r_1z}\\&\quad\left.+\frac{\lambda}{1-r_2}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right)\frac{(1-r_2)z}{1-r_2z}\right)\end{aligned} \tag{4-200}$$

显然，$L(z)$可分解为两部分，其中

$$L_0(z)=\frac{1-\rho}{1-\rho z} \tag{4-201}$$

为经典 M/M/1 排队系统队长的母函数，服从参数为 $1-\rho$ 的几何分布，且 $E[L_0]=L_0'(1)$。另一部分

$$L_d(z)=\frac{\pi_{00}}{1-\rho}\left(1+\frac{\lambda(\lambda+\theta-r_2\mu_{v_2})}{\theta^2}+\frac{\lambda}{\theta}+\frac{\lambda(\lambda+\theta-r_2\mu_{v_2})}{\theta(1-r_1)}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)\frac{(1-r_1)z}{1-r_1z}\right.$$
$$\left.+\frac{\lambda}{1-r_2}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right)\frac{(1-r_2)z}{1-r_2z}\right) \tag{4-202}$$

为由双级工作休假策略导致的附加队长的母函数。记

$$\begin{cases}\delta_0=\dfrac{\pi_{00}}{1-\rho}\left(1+\dfrac{\lambda(\lambda+\theta-r_2\mu_{v_2})}{\theta^2}+\dfrac{\lambda}{\theta}\right)\\ \delta_1=\dfrac{\pi_{00}}{1-\rho}\dfrac{\lambda(\lambda+\theta-r_2\mu_{v_2})}{\theta(1-r_1)}\left(\dfrac{r_1}{\mu_b(1-r_1)}-\dfrac{\rho-r_1}{\theta}\right)\\ \delta_2=\dfrac{\pi_{00}}{1-\rho}\dfrac{\lambda}{1-r_2}\left(\dfrac{r_2}{\mu_b(1-r_2)}-\dfrac{\rho-r_2}{\theta}\right)\end{cases}$$

因此，$L_d(z)$可表示为

$$L_d(z)=\delta_0+\delta_1\frac{(1-r_1)z}{1-r_1z}+\delta_2\frac{(1-r_2)z}{1-r_2z} \tag{4-203}$$

根据π_{00}，容易证明$\delta_0+\delta_1+\delta_2=1$。$L_d(z)$的结构说明附加队长$L_d$服从修正的几何分布。因此，$L_d$以概率$\delta_0$等于零，以概率$\delta_1$服从一个参数为$1-r_1$的几何分布，以概率$\delta_2$服从一个参数为$1-r_2$的几何分布。所以，$E[L_d]=L_d'(1)$。总之，$E[L]=L'(1)=L_0'(1)L_d(1)+L_d'(1)L_0(1)=E[L_0]+E[L_d]$。

定理 4.4.2 对于具有双级工作休假策略的完全不可视 M/M/1 排队系统，如果 $\rho=\lambda/\mu_b<1$ 且 $\mu_b>\mu_{v_1}$，$\mu_b>\mu_{v_2}$，则顾客平均逗留时间可分解为两部分：$E[W]=E[W_0]+E[W_d]$，其中

$$E[W_0]=\frac{1}{\mu_b(1-\rho)} \tag{4-204}$$

为经典 M/M/1 排队系统的稳态平均逗留时间，$E[W_d]$为附加延迟，且

$$E[W_d]=\frac{(\lambda+\theta-r_2\mu_{v_2})\pi_{00}}{\theta(1-r_1)^2(1-\rho)}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)$$
$$+\frac{\pi_{00}}{(1-r_2)^2(1-\rho)}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right) \tag{4-205}$$

证明 事实上，当一个顾客服务完成离开系统时，队长 L 为在他的逗留时间 W 内到达的顾客数。因此，基于定理 4.4.1 以及 L 的母函数和 W 的 Laplace 变换之间的关系式：$L(z)=\tilde{W}^*(\lambda(1-z))$，得到

$$\begin{aligned}\tilde{W}^*(s)&=L\left(1-\frac{s}{\lambda}\right)\\&=\frac{\mu_b(1-\rho)}{\mu_b(1-\rho)+s}\frac{\pi_{00}}{1-\rho}\left(1+\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta^2}+\frac{\lambda}{\theta}\right.\\&\quad+\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta(1-r_1)}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)\left(1-\frac{s}{\lambda}\right)\frac{\eta_1}{\eta_1+s}\\&\quad\left.+\frac{\lambda}{1-r_2}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right)\left(1-\frac{s}{\lambda}\right)\frac{\eta_2}{\eta_2+s}\right)\\&=\frac{\mu_b(1-\rho)}{\mu_b(1-\rho)+s}\frac{\pi_{00}}{1-\rho}\left(1+\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta^2}+\frac{\lambda}{\theta}\right.\\&\quad-\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta r_1}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)-\frac{\lambda}{r_2}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right)\\&\quad+\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta r_1(1-r_1)}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)\frac{\eta_1}{\eta_1+s}\\&\quad\left.+\frac{\lambda}{r_2(1-r_2)}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right)\frac{\eta_2}{\eta_2+s}\right)\end{aligned}\tag{4-206}$$

其中 $\eta_1=\lambda(1-r_1)/r_1$，$\eta_2=\lambda(1-r_2)/r_2$。类似于 $L(z)$，$\tilde{W}^*(s)$ 也可分解为两部分，其中

$$\tilde{W}_0^*(s)=\frac{\mu_b(1-\rho)}{\mu_b(1-\rho)+s}\tag{4-207}$$

为经典 M/M/1 排队系统的顾客逗留时间的 Laplace 变换，说明 W_0 服从参数为 $\mu_b(1-\rho)$ 的指数分布，且 $E[W_0]=-\tilde{W}_0^{*\prime}(0)$。另一部分

$$\begin{aligned}\tilde{W}_d^*(s)&=\frac{\pi_{00}}{1-\rho}\left(1+\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta^2}+\frac{\lambda}{\theta}-\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta r_1}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)\right.\\&\quad-\frac{\lambda}{r_2}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right)+\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta r_1(1-r_1)}\left(\frac{r_1}{\mu_b(1-r_1)}-\frac{\rho-r_1}{\theta}\right)\frac{\eta_1}{\eta_1+s}\\&\quad\left.+\frac{\lambda}{r_2(1-r_2)}\left(\frac{r_2}{\mu_b(1-r_2)}-\frac{\rho-r_2}{\theta}\right)\frac{\eta_2}{\eta_2+s}\right)\end{aligned}\tag{4-208}$$

为由双级工作休假策略导致的附加延迟的 Laplace 变换。记

$$\begin{cases}\sigma_0=\dfrac{\pi_{00}}{1-\rho}\left(1+\dfrac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta^2}+\dfrac{\lambda}{\theta}-\dfrac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta r_1}\left(\dfrac{r_1}{\mu_b\left(1-r_1\right)}-\dfrac{\rho-r_1}{\theta}\right)\right.\\ \qquad\left.-\dfrac{\lambda}{r_2}\left(\dfrac{r_2}{\mu_b\left(1-r_2\right)}-\dfrac{\rho-r_2}{\theta}\right)\right)\\ \sigma_1=\dfrac{\pi_{00}}{1-\rho}\dfrac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)}{\theta r_1\left(1-r_1\right)}\left(\dfrac{r_1}{\mu_b\left(1-r_1\right)}-\dfrac{\rho-r_1}{\theta}\right)\\ \sigma_1=\dfrac{\pi_{00}}{1-\rho}\dfrac{\lambda}{r_2\left(1-r_2\right)}\left(\dfrac{r_2}{\mu_b\left(1-r_2\right)}-\dfrac{\rho-r_2}{\theta}\right)\end{cases}$$

得到

$$\tilde{W}_d^*(s)=\sigma_0+\sigma_1\frac{\eta_1}{\eta_1+s}+\sigma_2\frac{\eta_2}{\eta_2+s} \tag{4-209}$$

因此，类似于 L_d 的分解结果，附加延迟 W_d 服从一个修正的指数分布。也就是说，W_d 以概率 σ_0 等于零，以概率 σ_1 服从参数为 η_1 的指数分布，以概率 σ_2 服从参数为 η_2 的指数分布。因此，$E[W_d]=-\tilde{W}_d^{*\prime}(0)$。总之，$E[W]=-\tilde{W}^{*\prime}(0)=-\tilde{W}_0^{*\prime}(0)\tilde{W}_d^*(0)-\tilde{W}_d^{*\prime}(0)\tilde{W}_0^*(0)=E[W_0]+E[W_d]$。

根据定理4.4.2，顾客的平均剩余效用为

$$U(\lambda)=R-CE[W]=R-\frac{C}{1-\rho}\left(\frac{1}{\mu_b}+\frac{\left(\lambda+\theta-r_2\mu_{v_2}\right)\pi_{00}}{\theta\left(1-r_1\right)^2}\left(\frac{r_1}{\mu_b\left(1-r_1\right)}-\frac{\rho-r_1}{\theta}\right)\right.$$
$$\left.+\frac{\pi_{00}}{\left(1-r_2\right)^2}\left(\frac{r_2}{\mu_b\left(1-r_2\right)}-\frac{\rho-r_2}{\theta}\right)\right) \tag{4-210}$$

求解 $U(\lambda)=0$，可得到正可行解 λ^*。至于 λ^* 的唯一性，我们可通过数值方法验证 $E[W]$是先减后增的凸函数还是严格增函数。结果表明，$E[W_d]$ 为凹函数且始终严格递增 $(\lambda<\mu_b)$，且由于 $E[W_0]$ 为凸函数且也是始终严格递增，所以 $E[W]$必为严格增函数。这说明顾客到达的正效应几乎可以被忽略。因此，我们推定顾客的正稳定均衡止步策略 λ_e 是唯一的，且 $\lambda_e=\min\{\lambda^*,\Lambda\}$。

再根据式（4-210），对于实际到达率 λ 来说，单位社会福利为

$$U_s(\lambda)=\lambda R-CE[L]=\lambda R-\frac{C}{1-\rho}\left(\rho+\frac{\lambda\left(\lambda+\theta-r_2\mu_{v_2}\right)\pi_{00}}{\theta\left(1-r_1\right)^2}\left(\frac{r_1}{\mu_b\left(1-r_1\right)}-\frac{\rho-r_1}{\theta}\right)\right.$$
$$\left.+\frac{\lambda\pi_{00}}{\left(1-r_2\right)^2}\left(\frac{r_2}{\mu_b\left(1-r_2\right)}-\frac{\rho-r_2}{\theta}\right)\right) \tag{4-211}$$

在均衡止步策略 λ_e 下，均衡社会福利可记为 $U_s(\lambda_e)$，而最优止步策略 λ_s 可通过求解最优问题 $\max U_s(\lambda)(\lambda \le \Lambda)$ 得到，且最优社会福利可记为 $U_s(\lambda_s)$。类似于几乎不可视情形，图 4-36 表明个人最优偏离社会最优，图 4-37 表明 $U_s(\lambda_e)$ 具有先增后减的变化趋势，而图 4-38 表明 $U_s(\lambda_s)$ 单增且最终恒定。

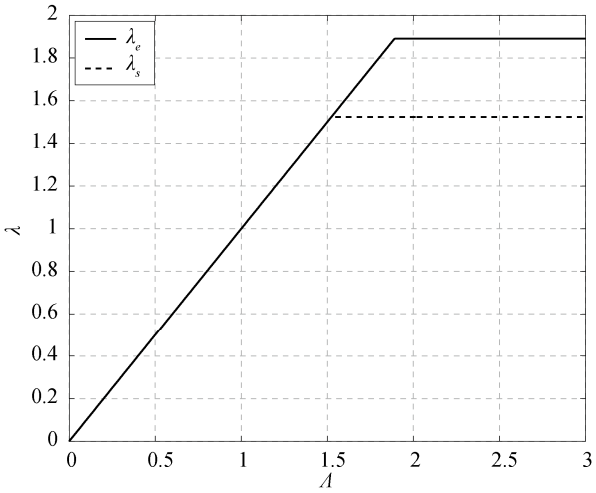

图 4-36 完全不可视情形下的均衡和最优混合止步策略
（$R=10, C=1, \mu_b=2, \mu_{v_1}=1.2, \mu_{v_2}=0.8, \mu=0.5$）

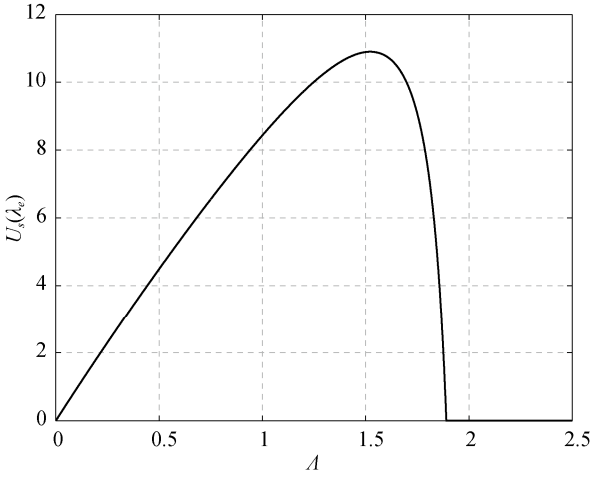

图 4-37　完全不可视情形下的均衡社会福利
（$R=10, C=1, \mu_b=2, \mu_{v_1}=1.2, \mu_{v_2}=0.8, \theta=0.5$）

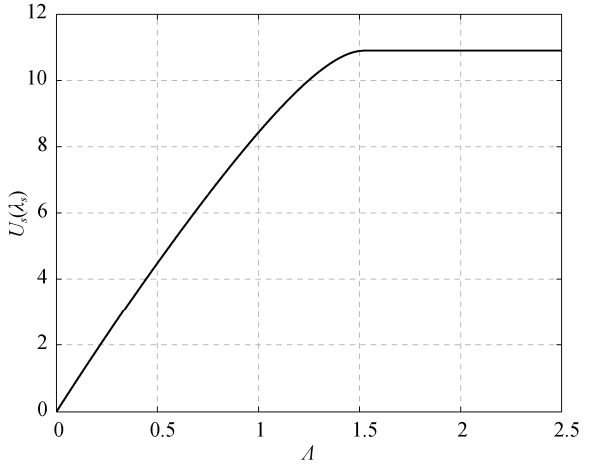

图 4-38 完全不可视情形下的最优社会福利
（$R=10, C=1, \mu_b=2$, $\mu_{v_1}=1.2$, $\mu_{v_2}=0.8$, $\theta=0.5$）

4.5 具有阈值工作休假策略的排队系统

本节将研究具有工作休假和阈值策略的有限容量排队系统中顾客的均衡和最优止步策略，并研究系统费用对顾客止步行为和最优社会福利的影响，以及顾客和社会计划者对空竭与非空竭服务规则的偏好。

4.5.1 模型描述

分析两类具有阈值工作休假策略的单服务员且容量有限的 Markov 排队系统，它们的区别在于服务规则不同，一类为空竭服务系统，一类为非空竭服务系统。假设服务需求很大，即顾客潜在到达率 Λ 足够大，顾客的实际到达率为 λ，服务员在忙期的服务率为 μ_b。对于空竭排队系统，一旦系统空竭，服务员则开始一个服务率为 $\mu_v\left(\mu_v<\mu_b\right)$ 的工作休假，当队长达到 T_r 时，服务员开始一个新忙期。然而，对于非空竭排队系统，服务员根据 T_r 阈值自动变换服务率。也就是说，当顾客数少于 T_r 时，服务率设为 μ_v，否则，服务率转为 μ_b。因此，我们记这两类排队系统分别为 $\mathrm{M/M/1/}N\left(T_r,E\right)$ 和 $\mathrm{M/M/1/}N\left(T_r,NE\right)$ 排队系统，其中 N 表示系统容量。

令 $(L_i(t),I_i(t))(i=1,2)$ 分别表示 $\mathrm{M/M/1/}N(T_r,E)$ 系统和 $\mathrm{M/M/1/}N(T_r,NE)$ 系统中时刻 t 的系统状态，其中 $L_i(t)$ 表示系统队长，$I_i(t)$ 表示时刻 t 时的服务员

状态，且

$$I_i(t)=\begin{cases}0, \text{服务员处于工作休假期}\\ 1, \text{服务员处于忙期}\end{cases}$$

我们考虑完全不可视情形，即到达顾客观测不到系统状态 $L_i(t)$ 和 $L_i(t)$。

在此，标记一个到达顾客，服务完成可获得服务收益 R，但也需承担单位等待损耗 c_w。我们采用线性等待损耗函数，因此平均剩余效用为 $U=R-c_wE[W]$，其中 $E[W]$表示他的平均逗留时间。假定系统总费用由状态转换费用和系统运行费用两部分构成，且单次转换费用为 c_s，单位时间运行费用为 $c_r\ (c_r\le c_s)$。而且，系统运行费用除了和运行时间 t_r 有关，还与服务率 μ 的大小成正比。因此，我们采用一个二元系统运行费用函数 $C_r(\mu,t_r)$ 且 $C_r(\mu,t_r)=c_r\mu t_r$。假设服务顺序为先到先服务，不允许顾客插队或中途退出系统。

4.5.2　队长与忙循环

4.5.2.1　空竭排队系统

对于空竭排队系统，首先导出稳态队长分布。显然，过程 $\{L_1(t),I_1(t)\}$ 为拟生灭过程，且状态空间为

$$\Omega_1=\{(k,0):0\le k\le T_r-1\}\cup\{(k,1):1\le k\le N\}$$

令(L_1, I_1)为过程 $\{L_1(t),I_1(t)\}$ 的稳态极限。记稳态队长分布为

$$\boldsymbol{\pi}=(\boldsymbol{\pi}_0,\boldsymbol{\pi}_1),\quad \boldsymbol{\pi}_0=(\pi_{0,0},\pi_{1,0},\cdots,\ \pi_{T_r-1,0}),\boldsymbol{\pi}_1=(\pi_{1,1},\pi_{2,1},\cdots,\pi_{N,1})$$

$$\pi_{k,j}=P\{L_1=k,I_1=j\}=\lim_{t\to\infty}P\{L_1(t)=k,I_1(t)=j\},(k,j)\in\Omega_1$$

最小生成元为 $\boldsymbol{Q}$，且状态转移概率方程 $\pi\boldsymbol{Q}=\boldsymbol{0}$ 为

$$\lambda\pi_{0,0}=\mu_b\pi_{1,1}+\mu_v\pi_{1,0} \tag{4-212}$$

$$(\lambda+\mu_v)\pi_{k,0}=\lambda\pi_{k-1,0}+\mu_v\pi_{k+1,0},1\le k\le T_r-2 \tag{4-213}$$

$$(\lambda+\mu_v)\pi_{T_r-1,0}=\lambda\pi_{T_r-2,0} \tag{4-214}$$

$$(\lambda+\mu_b)\pi_{1,1}=\mu_b\pi_{2,1} \tag{4-215}$$

$$(\lambda+\mu_b)\pi_{k,1}=\lambda\pi_{k-1,1}+\mu_b\pi_{k+1,1},2\le k\le N-1,k\ne T_r \tag{4-216}$$

$$(\lambda+\mu_b)\pi_{T_r,1}=\lambda\pi_{T_r-1,1}+\mu_b\pi_{T_r+1,1}+\lambda\pi_{T_r-1,0} \tag{4-217}$$

$$\lambda\pi_{N-1,1}=\mu_b\pi_{N,1} \tag{4-218}$$

令 $\boldsymbol{0}_{i\times j}$ 为 $i\times j$ 阶的零向量，$\varepsilon_j(1\le j\le T_r-1$ 或$1\le j\le N)$ 为单位列向量，ε_j^T 为它的转置。应用字典序，$\boldsymbol{Q}$ 可写成分块矩阵的形式：

$$\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{B} & \boldsymbol{D}\end{pmatrix} \tag{4-219}$$

其中

$$\boldsymbol{A}=\begin{pmatrix}-\lambda & \lambda & & & \\ \mu_v & -(\lambda+\mu_v) & \lambda & & \\ & \ddots & \ddots & \ddots & \\ & & \mu_v & -(\lambda+\mu_v) & \lambda \\ & & & \mu_v & -(\lambda+\mu_v)\end{pmatrix}_{T_r\times T_r}$$

$$\boldsymbol{B}=\begin{pmatrix}\mu_b & \boldsymbol{0}_{1\times(T_r-1)}\\ & \\ \boldsymbol{0}_{(N-1)\times 1} & \boldsymbol{0}_{(N-1)\times(T_r-1)}\end{pmatrix}_{N\times T_r},\qquad \boldsymbol{C}=\begin{pmatrix}\boldsymbol{O}_{(T_r-1)\times N}\\ \\ \lambda\varepsilon_{T_r}^{T}\end{pmatrix}_{T_r\times N}$$

$$\boldsymbol{D}=\begin{pmatrix}-(\lambda+\mu_b) & \lambda & & & & \\ \mu_b & -(\lambda+\mu_b) & \lambda & & & \\ & \mu_b & -(\lambda+\mu_b) & \lambda & & \\ & & \ddots & \ddots & \ddots & \\ & & & \mu_b & -(\lambda+\mu_b) & \lambda \\ & & & & \mu_b & -\mu_b\end{pmatrix}_{N\times N}$$

且 $\boldsymbol{A},\boldsymbol{D}$ 为三对角矩阵。为了以下证明方便，我们将 $\boldsymbol{A}$ 写成分块形式：

$$\boldsymbol{A}=\begin{pmatrix}-\lambda & \lambda\varepsilon_1^{T}\\ & \\ \mu_v\varepsilon_1 & \boldsymbol{A}_{0(T_r-1)\times(T_r-1)}\end{pmatrix}$$

引理 4.5.1　$\boldsymbol{A}_0$ 和 $\boldsymbol{D}$ 均为可逆阵，且它们的行列式分别为

$$\left|\boldsymbol{A}_0\right|=\begin{cases}(-1)^{T_r-1}\dfrac{\lambda^{T_r}-\mu_v^{T_r}}{\lambda-\mu_v}, & \lambda\neq\mu_v\\ T_r(-\lambda)^{T_r-1}, & \lambda\neq\mu_v\end{cases} \tag{4-220}$$

和

$$\left|\boldsymbol{D}\right|=(-\mu_b)^{N} \tag{4-221}$$

证明　根据$|\boldsymbol{A}_0|$的结构，得到递推公式：

$$\left|\boldsymbol{A}_0\right| = -\left(\lambda+\mu_v\right)\left|\boldsymbol{A}_0\right|^{(-1)} - \lambda\mu_v\left|\boldsymbol{A}_0\right|^{(-2)}$$

由此得到

$$\begin{aligned}\left|\boldsymbol{A}_0\right| + \mu_v\left|\boldsymbol{A}_0\right|^{(-1)} &= -\lambda\left(\left|\boldsymbol{A}_0\right|^{(-1)} + \mu_v\left|\boldsymbol{A}_0\right|^{(-2)}\right)\\ &= \left(-\lambda\right)^{T_r-3}\left(\left|\boldsymbol{A}_0\right|^{(2)} + \mu_v\left|\boldsymbol{A}_0\right|^{(1)}\right)\\ &= \left(-\lambda\right)^{T_r-1}\end{aligned}$$

从而，当 $\lambda \neq \mu_v$ 时，

$$\left|\boldsymbol{A}_0\right| = -\mu_v\left|\boldsymbol{A}_0\right|^{(-1)} + \left(-\lambda\right)^{T_r-1} = \sum_{k=0}^{T_r-1}\left(-\mu_v\right)^k\left(-\lambda\right)^{T_r-k-1} = \left(-1\right)^{T_r-1}\frac{\lambda^{T_r}-\mu_v^{T_r}}{\lambda-\mu_v}$$

特别地，如果 $\lambda = \mu_v$，我们有

$$\left|\boldsymbol{A}_0\right| = -\lambda\left|\boldsymbol{A}_0\right|^{(-1)} + \left(-\lambda\right)^{T_r-1} = \sum_{k=0}^{T_r-1}\left(-\lambda\right)^{T_r-1} = T_r\left(-\lambda\right)^{T_r-1}$$

类似地，通过初等变换，也可得到 $\boldsymbol{D}$ 的行列式的值。将 $\boldsymbol{D}$ 的后 N-1 列加到第一列上，然后将 $\boldsymbol{D}$ 对应的行列式按第一列展开，得到

$$\left|\boldsymbol{D}\right| = -\mu_b\left|\boldsymbol{D}\right|^{(-1)} = \left(-\mu_b\right)^N$$

定理 4.5.1　对于具有 T_r 阈值工作休假策略的空竭服务 M/M/1/N 排队系统，稳态系统队长分布为

$$\begin{cases}\pi_{j,0} = \dfrac{\lambda^j\left(\mu_v^{T_r-j}-\lambda^{T_r-j}\right)}{\mu_v^{T_r}-\lambda^{T_r}}\pi_{0,0}, & 1\le j\le T_r-1\\ \pi_{j,1} = \dfrac{\lambda^{T_r}\left(\mu_v-\lambda\right)}{\mu_v^{T_r}-\lambda^{T_r}}\varepsilon_{T_r}^T\boldsymbol{D}^{-1}\varepsilon_j\pi_{0,0}, & 1\le j\le N\end{cases} \tag{4-222}$$

其中

$$\pi_{0,0} = \left[1+\frac{T_r\lambda^{T_r}\left(\lambda-\mu_v\right)-\lambda\left(\lambda^{T_r}-\mu_v^{T_r}\right)}{\left(\mu_v-\lambda\right)\left(\mu_v^{T_r}-\lambda^{T_r}\right)} - \frac{\lambda^{T_r}\left(\mu_v-\lambda\right)}{\mu_v^{T_r}-\lambda^{T_r}}\varepsilon_{T_r}^T\boldsymbol{D}^{-1}\boldsymbol{e}_N\right]^{-1} \tag{4-223}$$

且 $\boldsymbol{e}_N = \left(1,1,\cdots,1\right)_{1\times N}^T$。

证明　由于 $\boldsymbol{\pi}$ 满足 $\boldsymbol{\pi Q} = \boldsymbol{0}_{1\times(T_r+N)}$，因此得到

$$\begin{cases}\boldsymbol{\pi}_0\boldsymbol{A}+\boldsymbol{\pi}_1\boldsymbol{B} = \boldsymbol{0}_{1\times T_r}\\ \boldsymbol{\pi}_0\boldsymbol{C}+\boldsymbol{\pi}_1\boldsymbol{D} = \boldsymbol{0}_{1\times N}\end{cases} \tag{4-224}$$

令 $\boldsymbol{\pi}_0 = (\pi_{0,0},\boldsymbol{\pi}_0'), \boldsymbol{\pi}_1 = (\pi_{1,1},\boldsymbol{\pi}_1')$，所以

$$\left(\pi_{0,0},\boldsymbol{\pi}_0'\right)\begin{pmatrix}-\lambda & \lambda\boldsymbol{\varepsilon}_1^T\\ \mu_v\boldsymbol{\varepsilon}_1 & \boldsymbol{A}_0\end{pmatrix}+\left(\pi_{1,1},\boldsymbol{\pi}_1'\right)\begin{pmatrix}\mu_b & \boldsymbol{0}_{1\times(T_r-1)}\\ \boldsymbol{0}_{(N-1)\times 1} & \boldsymbol{0}_{(N-1)\times(T_r-1)}\end{pmatrix}_{N\times T_r}=\boldsymbol{0}_{1\times T_r}$$

由此得到

$$\begin{cases}\boldsymbol{\pi}_0=\left(1,-\lambda\boldsymbol{\varepsilon}_1^T\boldsymbol{A}_0^{-1}\right)\pi_{0,0}\\ \boldsymbol{\pi}_1=\left(-1,\lambda\boldsymbol{\varepsilon}_1^T\boldsymbol{A}_0^{-1}\right)\boldsymbol{C}\boldsymbol{D}^{-1}\pi_{0,0}\end{cases}\tag{4-225}$$

令 $\lambda\boldsymbol{\varepsilon}_1^T\boldsymbol{A}_0^{-1}=\boldsymbol{X}$ 且向量 $\boldsymbol{X}=(x_1,x_2,\cdots,x_{T_r-1})$，则 $\boldsymbol{X}$ 可通过求解 $\boldsymbol{A}_0{}^T\boldsymbol{X}^T=\lambda\boldsymbol{\varepsilon}_1$ 得到。

根据克莱姆法则（Cramer's Rule），得到方程组的解为

$$x_j=(-1)^{T_r}\lambda^j\frac{\mu_v^{T_r-j}-\lambda^{T_r-j}}{\mu_v-\lambda}\Big/\left|\boldsymbol{A}_0{}^T\right|,\ 1\le j\le T_r-1\tag{4-226}$$

将式（4-220）代入式（4-226），得到

$$x_{j,0}=-\lambda\boldsymbol{\varepsilon}_1^T\boldsymbol{A}_0^{-1}\varepsilon_j\pi_{0,0}\frac{\lambda^j\left(\mu_v^{T_r-j}-\lambda^{T_r-j}\right)}{\mu_v^{T_r}-\lambda^{T_r}}\pi_{0,0},\ 1\le j\le T_r-1$$

根据 $\boldsymbol{C}$ 的特殊结构，还可得到

$$\boldsymbol{\pi}_1=-\lambda\pi_{T_r-1,0}\boldsymbol{\varepsilon}_{T_r}^T\boldsymbol{D}^{-1}=-\frac{\lambda^{T_r}(\mu_v-\lambda)}{\mu_v^{T_r}-\lambda^{T_r}}\boldsymbol{\varepsilon}_{T_r}^T\boldsymbol{D}^{-1}\pi_{0,0}$$

因此，

$$\pi_{j,1}=-\frac{\lambda^{T_r}(\mu_v-\lambda)}{\mu_v^{T_r}-\lambda^{T_r}}\boldsymbol{\varepsilon}_{T_r}^T\boldsymbol{D}^{-1}\varepsilon_j\pi_{0,0},\ 1\le j\le N$$

根据平衡方程 $\boldsymbol{\pi}\boldsymbol{e}_{T_r+N}=1$，其中 $\boldsymbol{e}_{T_r+N}=(1,1,\cdots,1)_{1\times(T_r+N)}^T$，可得到 $\pi_{0,0}$ 的表达式。

令 $\boldsymbol{D}^{-1}\boldsymbol{e}_N=\boldsymbol{Y}$，可通过求解 $\boldsymbol{D}\boldsymbol{Y}=\boldsymbol{e}_N$ 得到 $\boldsymbol{Y}$。

鉴于定理 4.5.1 中的结果，系统处于状态 $i(i=0,1)$的概率 p_i 分别为

$$\begin{cases}p_0=\sum_{j=0}^{T_r-1}\pi_{j,0}=\left(\dfrac{T_r\lambda^{T_r}}{\lambda^{T_r}-\mu_v^{T_r}}-\dfrac{\mu_v}{\lambda-\mu_v}\right)\pi_{0,0}\\ p_1=\sum_{j=1}^{N}\pi_{j,1}=-\dfrac{\lambda^{T_r}\left(\mu_v-\lambda\right)}{\mu_v^{T_r}-\lambda^{T_r}}\boldsymbol{\varepsilon}_{T_r}^T\boldsymbol{D}^{-1}\boldsymbol{e}_N\pi_{0,0}\end{cases}\tag{4-227}$$

平均队长为

$$\begin{aligned}E[L_1]&=\sum_{j=1}^{T_r-1}j\pi_{j,0}+\sum_{j=0}^{N}j\pi_{j,1}\\&=\left(\frac{\lambda^{T_r}\mu_vT_r\left(\lambda-\mu_v\right)-\lambda\mu_v\left(\lambda^{T_r}-1\right)}{\left(\mu_v^{T_r}-\lambda^{T_r}\right)\left(\lambda-\mu_v\right)^2}-\frac{\lambda^{T_r}T_r\left(T_r-1\right)}{2\left(\mu_v^{T_r}-\lambda^{T_r}\right)}-\frac{\lambda^{T_r}\left(\mu_v-\lambda\right)}{\mu_v^{T_r}-\lambda^{T_r}}\boldsymbol{\varepsilon}_{T_r}^T\boldsymbol{D}^{-1}\eta_N\right)\pi_{0,0}\end{aligned}\tag{4-228}$$

其中 $\eta_N=(1,2,\cdots,N)^T$。因此，顾客的平均逗留时间为 $E[W_1]=E[L_1]/(\lambda(1-\pi_{N,1}))$，

且顾客的均衡到达率 λ_1^e 可通过求解 $R-c_\omega E[W_1]=0$ 得到。如果不考虑系统费用，平均社会福利为

$$S_1(\lambda)=(1-\pi_{N,1})\lambda R-c_\omega E[L_1] \tag{4-229}$$

现在，为了分析系统费用以及考虑系统费用条件下的社会福利，继续推导空竭服务系统的平均忙循环。鉴于阈值策略和空竭服务规则，在整个工作休假期间，到达顾客数与服务完成顾客数之差应为阈值 T_r。因此，如果在休假期间有 $j(j\geq 0)$ 个顾客完成服务且到达顾客数不少于 j 个，则休假长度应等于 T_r+j 个到达间隔。令 $v_1(t),\overline{v_1}$ 分别为空竭服务系统中休假长度的概率密度函数和均值，得到

$$v_1(t)=\sum_{j=0}^{\infty}\frac{\lambda(\lambda t)^{T_r+j-1}}{(T_r+j-1)!}\frac{(\mu_v t)^j}{j!}e^{-(\lambda+\mu_v)t}\left(1-\sum_{k=0}^{j-1}\frac{(\lambda t)^k}{k!}e^{-\lambda t}\right) \tag{4-230}$$

从而得到平均休假时间为

$$\overline{v_1}=\sum_{j=0}^{\infty}\frac{(T_r+2j)!}{(T_r+j-1)!j!}\frac{\lambda^{T_r+j}\mu_v^j}{(\lambda+\mu_v)^{T_r+2j+1}}-\sum_{j=1}^{\infty}\sum_{k=0}^{j-1}\frac{(T_r+2j+k)!}{(T_r+j-1)!j!k!}\frac{\lambda^{T_r+j+k}\mu_v^j}{(2\lambda+\mu_v+s)^{T_r+2j+k+1}} \tag{4-231}$$

根据式（4-227），得到空竭服务系统的平均忙循环，记为 B_c^e，并且 $B_c^e=\overline{v_1}/p_0$。因此，单位时间内系统总费用 C_T^e 为

$$C_T^e=c_r(p_0\mu_v+p_1\mu_b)+\left(2c_s\left\lfloor\frac{1}{B_c^e}\right\rfloor+c_s(p_0+2p_1)\right) \tag{4-232}$$

其中$\lfloor\cdot\rfloor$是向下取整函数。不失一般性，假定一个忙循环开始于一个工作休假期开始的时刻。因此，考虑系统费用条件下的社会福利为

$$S_1(\lambda)=(1-\pi_{N,1})\lambda R-c_\omega E[L_1]-c_r(p_0\mu_v+p_1\mu_b)-\left(2c_s\left\lfloor\frac{1}{B_c^e}\right\rfloor+c_s(p_0+2p_1)\right) \tag{4-233}$$

4.5.2.2　非空竭排队系统

对于非空竭排队系统，过程$\{L_2(t), I_2(t)\}$仍为一个拟生灭过程，且状态空间为

$$\Omega_2=\{(k,0):0\leq k\leq T_r-1\}\cup\{(k,1):T_r\leq k\leq N\}$$

记稳态队长分布为

$$\boldsymbol{\pi}=(\boldsymbol{\pi}_0,\boldsymbol{\pi}_1),\quad \boldsymbol{\pi}_0=(\pi_{0,0},\pi_{1,0},\cdots,\pi_{T_r-1,0}),\quad \boldsymbol{\pi}_1=(\pi_{T_r,1},\pi_{T_r+1,1},\cdots,\pi_{N,1})$$

从而得到状态转移概率方程为

$$\lambda\pi_{0,0}=\mu_v\pi_{1,0} \tag{4-234}$$

$$(\lambda+\mu_v)\pi_{j,0}=\lambda\pi_{j-1,0}+\mu_v\pi_{j+1,0},\quad 1\leq j\leq T_r-2 \tag{4-235}$$

$$(\lambda+\mu_v)\pi_{T_r-1,0}=\lambda\pi_{T_r-2,0}+\mu_b\pi_{T_r,1} \tag{4-236}$$

$$(\lambda+\mu_b)\pi_{T_r,1}=\lambda\pi_{T_r-1,0}+\mu_b\pi_{T_r+1,1} \tag{4-237}$$

$$(\lambda+\mu_b)\pi_{j,1}=\lambda\pi_{j-1,1}+\mu_b\pi_{j+1,1},\quad T_r+1\le j\le N-1 \tag{4-238}$$

$$\mu_b\pi_{N,1}=\lambda\pi_{N-1,1} \tag{4-239}$$

定理 4.5.2 对于具有 T_r 阈值工作休假策略的非空竭服务 M/M/1/N 排队系统，稳态系统队长分布为

$$\begin{cases}\pi_{j,0}=\left(\dfrac{\lambda}{\mu_v}\right)^j\pi_{0,0},\ 0\le j\le T_r-1\\ \pi_{j,0}=\left(\dfrac{\lambda}{\mu_b}\right)^{j-T_r+1}\left(\dfrac{\lambda}{\mu_v}\right)^{T_r-1}\pi_{0,0},\quad T_r\le j\le N_r\end{cases} \tag{4-240}$$

其中

$$\pi_{0,0}=\mu_v^{T_r-1}\left[\frac{\mu_v^{T_r}-\lambda^{T_r}}{\mu_v-\lambda}+\frac{\lambda^{T_r}\left(\mu_b^{N-T_r+1}-\lambda^{N-T_r+1}\right)}{\mu_b^{N-T_r+1}\left(\mu_b-\lambda\right)}\right]^{-1} \tag{4-241}$$

证明 求解式（4-235），可得到递推公式：

$$\pi_{j+1,0}-\pi_{j,0}=\frac{\lambda}{\mu_v}\left(\pi_{j,0}-\pi_{j-1,0}\right)=\left(\frac{\lambda}{\mu_v}\right)^j\left(\frac{\lambda}{\mu_v}-1\right)\pi_{0,0}$$

考虑式（4-234），得到

$$\begin{aligned}\pi_{j,0}&=\pi_{j-1,0}+\left(\frac{\lambda}{\mu_v}\right)^{j-1}\left(\frac{\lambda}{\mu_v}-1\right)\pi_{0,0}\\&=\pi_{0,0}+\sum_{i=0}^{j-1}\left(\frac{\lambda}{\mu_v}\right)^i\left(\frac{\lambda}{\mu_v}-1\right)\pi_{0,0}\\&=\left(\frac{\lambda}{\mu_v}\right)^j\pi_{0,0},0\le j\le T_r-1\end{aligned}$$

将式（4-234）至式（4-236）加在一起，可得到关系式

$$\pi_{T_r,1}=\frac{\lambda}{\mu_b}\pi_{T_r-1,0}=\left(\frac{\lambda}{\mu_b}\right)\left(\frac{\lambda}{\mu_v}\right)^{T_r-1}\pi_{0,0}$$

再考虑式（4-237），可得

$$\pi_{T_r+1,1}=\frac{\lambda+\mu_b}{\mu_b}\pi_{T_r,1}-\frac{\lambda}{\mu_b}\pi_{T_r-1,0}=\left(\frac{\lambda}{\mu_b}\right)^2\left(\frac{\lambda}{\mu_v}\right)^{T_r-1}\pi_{0,0}$$

因此，根据式（4-238），得到递推公式：

$$\pi_{j+1,1}-\pi_{j,1}=\left(\frac{\lambda}{\mu_b}\right)^{j-T_r}\left(\pi_{T_r+1,1}-\pi_{T_r,1}\right)=\left(\frac{\lambda}{\mu_b}\right)^{j-T_r+1}\left(\frac{\lambda}{\mu_v}\right)^{T_r-1}\left(\frac{\lambda}{\mu_b}-1\right)\pi_{0,0}$$

从而得到

$$\begin{aligned}\pi_{j,1}&=\pi_{j-1,1}+\left(\frac{\lambda}{\mu_b}\right)^{j-T_r}\left(\frac{\lambda}{\mu_v}\right)^{T_r-1}\left(\frac{\lambda}{\mu_b}-1\right)\pi_{0,0}\\&=\pi_{T_r,1}+\sum_{i=T_r+1}^{j}\left(\frac{\lambda}{\mu_b}\right)^{i-T_r}\left(\frac{\lambda}{\mu_v}\right)^{T_r-1}\left(\frac{\lambda}{\mu_b}-1\right)\pi_{0,0}\\&=\left(\frac{\lambda}{\mu_b}\right)^{j-T_r+1}\left(\frac{\lambda}{\mu_v}\right)^{T_r-1}\pi_{0,0},T_r\le j\le N-1\end{aligned}$$

方程（4-239）说明 $\pi_{N,1}$ 具有与 $\{\pi_{j,1}:T_r\le j\le N-1\}$ 一样的形式。根据平衡方程 $\boldsymbol{\pi e}_N=1$，得到 $\pi_{0,0}$。

基于定理 4.5.2 中的结果，系统处于状态 i ($i=0,1$)的概率为

$$\begin{cases}p_0=\sum_{j=0}^{T_r-1}\pi_{j,0}=\dfrac{\mu_v}{\mu_v-\lambda}\left(1-\left(\dfrac{\lambda}{\mu_v}\right)^{T_r}\right)\pi_{0,0}\\p_1=\sum_{j=T_r}^{N}\pi_{j,1}=\dfrac{\lambda}{\mu_b-\lambda}\left(\dfrac{\lambda}{\mu_v}\right)^{T_r-1}\left(1-\left(\dfrac{\lambda}{\mu_b}\right)^{N-T_r+1}\right)\pi_{0,0}\end{cases}\tag{4-242}$$

平均队长为

$$\begin{aligned}E[L_2]&=\sum_{j=1}^{T_r-1}j\pi_{j,0}+\sum_{j=T_r}^{N}j\pi_{j,1}\\&=\left(\frac{\lambda^{T_r}\left(\lambda T_r-\mu_vT_r-\lambda\right)+\lambda\mu_v}{\mu_v^{T_r-1}\left(\lambda-\mu_v\right)^2}+\frac{\mu_v^{T_r-1}\lambda^{N+1}\left(N\lambda-(N+1)\mu_b\right)}{\mu_b^{N+T_r-1}\left(\lambda-\mu_b\right)^2}-\frac{\mu_v^{T_r-1}\lambda^{T_r}\left(\lambda T_r-\mu_bT_r-\lambda\right)}{\mu_b^{2(T_r-1)}\left(\lambda-\mu_b\right)^2}\right)\end{aligned}\tag{4-243}$$

因此，顾客的平均逗留时间为 $E[W_2]=E[L_2]/(\lambda(1-\pi_{N,1}))$，且顾客的均衡到达率 λ_2^e 可通过求解 $R-c_\omega E[W_2]=0$ 得到。如果不考虑系统费用，平均社会福利为

$$S_2(\lambda)=(1-\pi_{N,1})\lambda R-c_\omega E[L_2]\tag{4-244}$$

接下来，导出非空竭服务系统的平均忙循环。类似地，鉴于阈值策略和非空竭服务规则，在整个工作休假期间，到达顾客数与服务完成顾客数之差应为 1。因此，如果在休假期间有 $j(0\le j\le T_r-1)$ 个顾客完成服务，则休假长度应等于 j+1 个到达间隔。如果在 T_r-1 个顾客的服务期间（且只有 j 个顾客完成服务）有至少 $j-T_r+1(j\ge T_r)$ 个顾客到达，则休假长度应等于 $j+1$ 个到达间隔。令 $v_2(t),\overline{v}_2$ 分别为非空竭服务系统中休假长度的概率密度函数和均值，得到

$$v_2(t)=\sum_{j=0}^{T_r-1}\frac{\lambda(\lambda t)^j}{j!}\frac{(\mu_v t)^j}{j!}e^{-(\lambda+\mu_v)t}+\sum_{j=T_r}^{\infty}\frac{\lambda(\lambda t)^j}{j!}\frac{(\mu_v t)^j}{j!}e^{-(\lambda+\mu_v)t}\int_0^{\infty}\sum_{k=0}^{T_r-2}\frac{\lambda(\lambda u)^{j-T_r}}{(j-T_r)!}\frac{(\mu_v u)^k}{k!}e^{-(\lambda+\mu_v)u}\mathrm{d}u \tag{4-245}$$

从而得到平均休假时间为

$$\overline{v}_2=\sum_{j=0}^{T_r-1}\frac{(2j+1)!}{j!\,j!}\frac{\lambda^{j+1}\mu_v^j}{(\lambda+\mu_v)^{2j+2}}+\sum_{j=T_r}^{\infty}\frac{(2j+1)!}{j!\,j!(j-T_r)!}\frac{\lambda^{2j-T_r+2}\mu_v^j}{(\lambda+\mu_v)^{2j+2}}\sum_{k=0}^{T_r-2}\frac{\mu_v^k}{k!}\frac{(j+k-T_r)!}{(\lambda+\mu_v)^{j+k-T_r+1}} \tag{4-246}$$

根据式（4-242），得到非空竭服务系统的平均忙循环，记为 B_c^n，并且 $B_c^n=\overline{v}_2/p_0$。因此，单位时间内系统总费用 C_T^n 为

$$C_T^n=c_r\left(p_0\mu_v+p_1\mu_b\right)+\left(2c_s\left\lfloor\frac{1}{B_c^n}\right\rfloor+c_s\left(p_0+2p_1\right)\right) \tag{4-247}$$

因此，考虑系统费用条件下的社会福利为

$$S_2(\lambda)=\left(1-\pi_{N,1}\right)\lambda R-c_\omega E\left[L_2\right]-c_r\left(p_0\mu_v+p_1\mu_b\right)-\left(2c_s\left\lfloor\frac{1}{B_c^n}\right\rfloor+c_s\left(p_0+2p_1\right)\right) \tag{4-248}$$

4.5.3 数值比较

4.5.3.1 忙循环

首先，我们比较空竭服务系统和非空竭服务系统的平均忙循环。从图 4-39 和图 4-40 观察到，空竭系统的平均忙循环总是比非空竭系统的更长，而且空竭系统的平均忙循环对各类系统参数变化的敏感度更高。

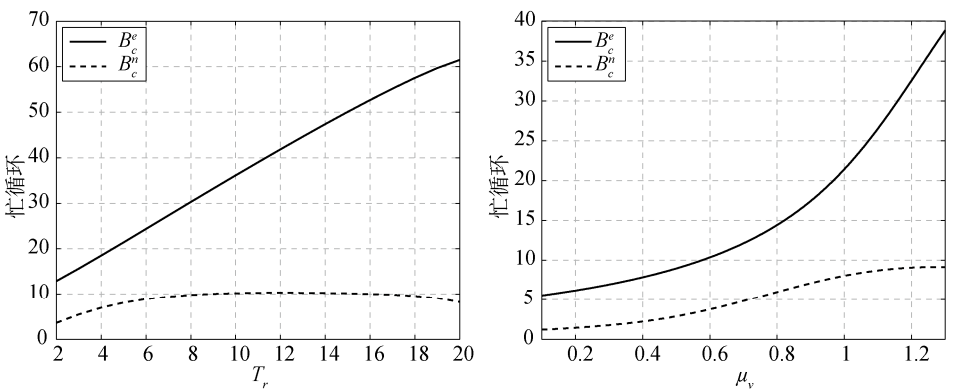

图 4-39　空竭与非空竭排队系统的忙循环关于(1) T_r 的敏感度（$\lambda=2,\mu_b=4,\mu_v=1,N=20$）、(2) μ_v 的敏感度（$\lambda=2,\mu_b=4,N=20,T_r=5$）

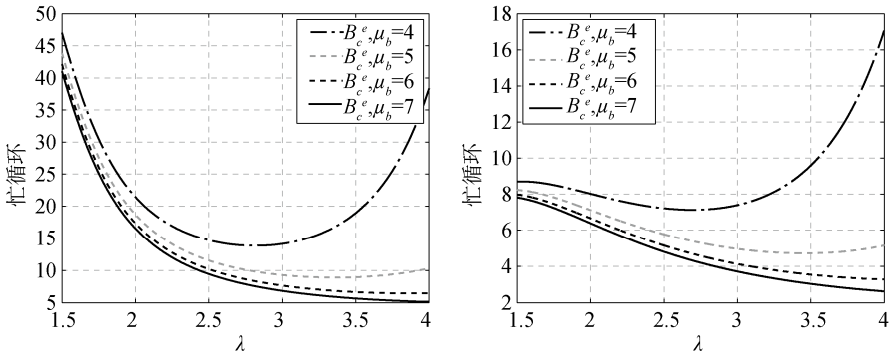

图 4-40　空竭与非空竭排队系统的忙循环关于 λ 的敏感度

对于空竭系统，随着 T_r 的增加，一个工作休假期显然会增长，且忙期开始时会积压更多的顾客，从而使得忙期长度也增长，所以忙循环 B_c^e 相应增长。然而，非空竭系统的忙循环 B_c^n 对 T_r 的敏感度并不高，甚至，当 T_r 接近于 N 时，忙循环长度变得更短。实际上，这是由有限系统容量 N 造成的，当 T_r 接近于 N 时，会有更少的顾客接受服务，从而忙期减短。

对于忙循环关于 μ_v 的敏感度来说，我们观察到忙循环始终随 μ_v 递增。然而，图 4-40 表明，无论对于空竭系统还是非空竭系统，忙循环始终随 μ_b 递减。这也说明系统负载越低，平均忙循环越短。而且，由于实行阈值策略，对于空竭系统和非空竭系统，图 4-40 表明，当 λ 较小时，一个工作休假和整个的忙循环都随 λ 递减，而当 λ 足够大后，工作休假减少的部分不足以抵消忙期增加的部分，因而忙循环开始随 λ 递增。

4.5.3.2　系统费用

接下来，在一个固定时间区间 T 内，比较空竭系统和非空竭系统的系统费用。在空竭系统中，显然，较高的阈值 T_r 能降低服务员状态转换的频率，从而减少系统费用。因此，图 4-41 表明系统费用随 T_r 递减，尤其是当 c_s 远大于 c_r 时更是如此。而且，无论 c_s 与 c_r 的大小关系如何，当 T_r 超过某个域值后，系统费用的递减率将明显降低。

另一方面，对于非空竭系统来说，虽然较大的阈值 T_r 能减少系统运行费用，但服务员状态转换的频率不一定与 T_r 成反比。尤其是当 T_r 接近 N 时，由于越来越短的忙期，服务员的状态转换频率可能会越来越高。因此，当 $c_s > c_r$ 且 T_r 趋近于 N 时，从图 4-41 可以观察到系统费用有一个小的反弹。实际上，对于固定的 c_r 和 c_s 以及可变的 μ_v 和 μ_b，我们已经验证了当 T_r 趋近于 N 时，系统费用仍然有一个小的反弹，尤其是当 μ_b 和 μ_v 的差距较小时。图 4-41 还表明一般当 $c_s > c_r$ 时，$C_T^n > C_T^e$，这是由于 $B_c^e > B_c^n$，

且C_T^n的减速比C_T^e要快，这说明系统运行费用在系统总费用中应占更大比例。

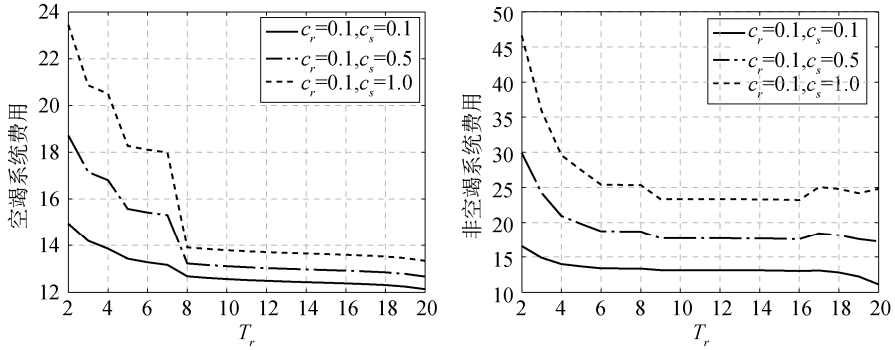

图 4-41　空竭与非空竭排队系统在 T 时间段内的系统费用关于 T_r 的敏感度

$(\lambda = 2, \mu_b = 4, \mu_v = 1, T = 60, N = 20)$

4.5.3.3　到达率与社会福利

最后，我们分别比较空竭系统和非空竭系统的顾客均衡到达率、最优到达率以及最优社会福利。在此，我们只考虑稳定的均衡到达率。图 4-42 表明 λ_1^e 总是高于 λ_2^e。原因在于顾客相信，在非空竭系统中，他们更有可能遇到低速服务的服务员。然而，当 T_r 很小或接近于 N 时，他们的差距并不明显。而且，λ_2^e 相比 λ_1^e 来说，对 T_r 的敏感度更高。因此，在非空竭系统中，通过调节阈值来控制系统负载将更有效。

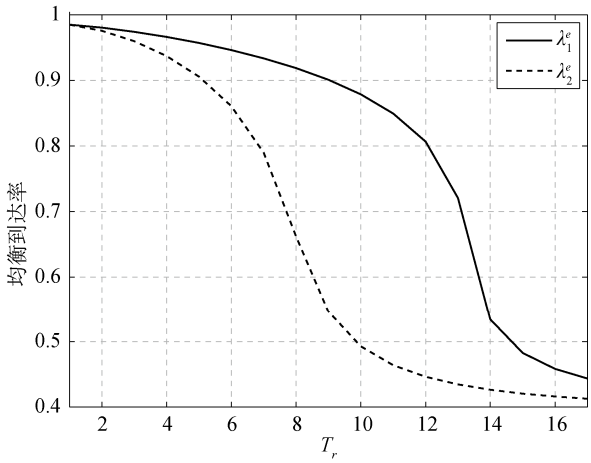

图 4-42　空竭与非空竭排队系统的均衡到达率关于 T_r 的敏感度

$(R = 10, C = 1, \mu_b = 1, \mu_v = 0.5, N = 20)$

不考虑系统费用，图 4-43 表明 $\lambda_1^* \geq \lambda_2^*, S_1(\lambda^*) \geq S_2(\lambda^*)$，且随着 T_r 的增大，二者几乎相等。而且，对于空竭系统和非空竭系统，它们的最优阈值并不相等，

即空竭系统的最优阈值为 $T_r=2$，非空竭系统的最优阈值为 $T_r=1$。这说明，尽管没考虑系统费用，至少对于空竭系统来说，采取阈值策略对增加社会福利是有益的，并能提高服务员效率。

类似于图 4-43，图 4-46 表明 $\lambda_1^*\geq\lambda_2^*$ 且 $S_1(\lambda^*)\geq S_2(\lambda^*)$。而且，$\lambda_1^*$ 和 $S_1(\lambda^*)$ 均随 N 递减，而 λ_2^* 和 $S_2(\lambda^*)$ 几乎保持不变。这说明对于社会计划者而言，空竭系统的系统负载应高于非空竭系统的系统负载，这使得空竭系统的性能指标对 N 变化的敏感度更高。

另一方面，在社会福利中考虑系统费用，图 4-44 说明，当 $T_r>2$ 时，最优到达率和最优社会福利不一定随 T_r 递减。这主要是由于系统费用一般情况下是随 T_r 递减的。而且，$S_1(\lambda^*)$ 并不总是大于 $S_2(\lambda^*)$，这是由于在非空竭系统中，系统费用随 T_r 的减速更高。因此，如果考虑系统费用，空竭服务规则对于社会计划者来说并不一定是更好的选择，即当 $T_r>11$ 时，他应选择非空竭服务规则。而且，与图 4-43 类似，空竭系统和非空竭系统的最优阈值并不一致。

与图 4-43 和图 4-44 相反，图 4-45 表明，当 μ_v 比 μ_b 低很多时，为了保证系统正常运行，随着 T_r 的增大，社会计划者始终希望更多的顾客加入系统，尤其是对于非空竭服务系统。然而，显然，这与顾客的个人意愿是冲突的。而且，不同于最优到达率，最优社会福利始终随 T_r 递减，因此 $T_r=1$ 对于空竭和非空竭系统来说都是最优的。

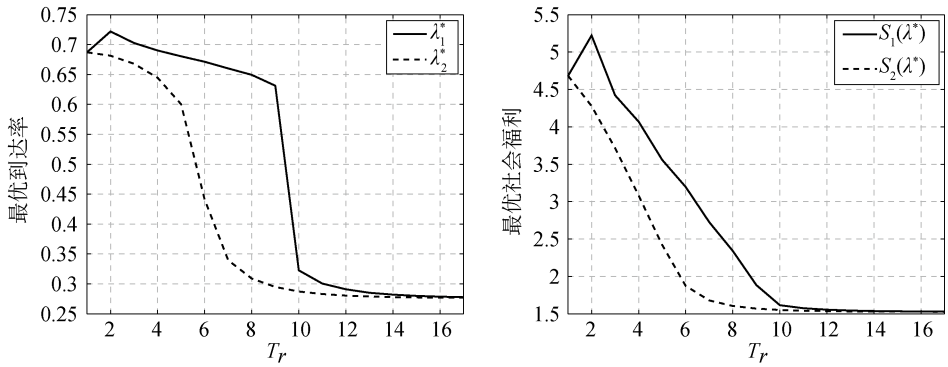

图 4-43　空竭与非空竭排队系统的最优到达率和最优社会福利（不包括系统费用）关于 T_r 的敏感度 $(R=10, C=1, \mu_b=1, \mu_v=0.5, N=20)$

和图 4-46 相比，除了空竭系统的最优到达率和最优社会福利随 N 递减之外，图 4-47 还表明在非空竭系统中，二者均具有先递增而后不变最后递减的变化趋势。主要原因就是对于非空竭系统，当 N 相比 T_r 较小时，状态转换费用随 N 递减，而当 N 相比 T_r 大很多时，系统运行费用随 N 递增。

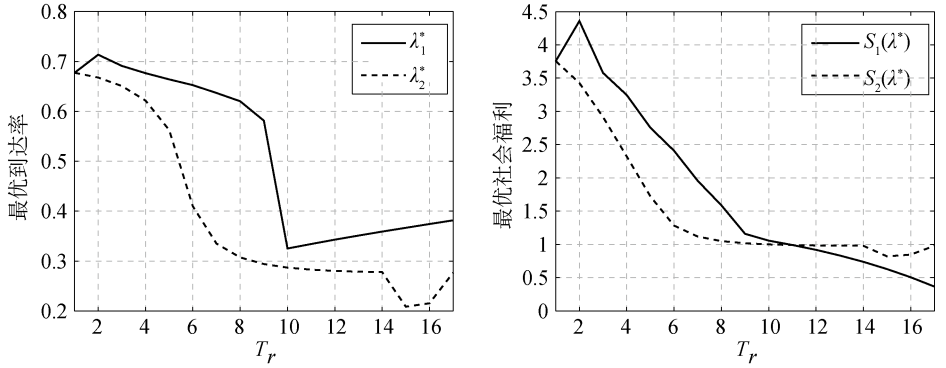

图 4-44　空竭与非空竭排队系统的最优到达率和最优社会福利（包括系统费用）关于 T_r 的敏感度（$R=10, C=1, c_r=0.1, c_s=0.5, \mu_b=1, \mu_v=0.5, N=20$）

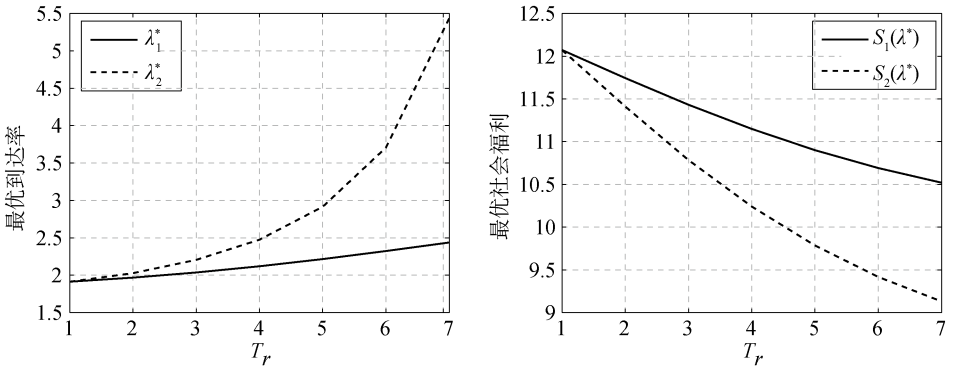

图 4-45　空竭与非空竭排队系统的最优到达率和最优社会福利（包括系统费用）关于 T_r 的敏感度（$R=10, C=1, c_r=0.1, c_s=0.5, \mu_b=2, \mu_v=0.1, N=10$）

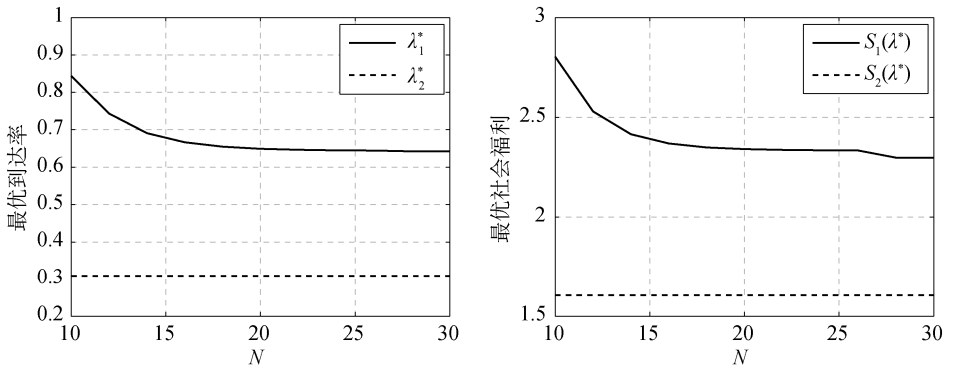

图 4-46　空竭与非空竭排队系统的最优到达率和最优社会福利（不包括系统费用）关于 N 的敏感度（$R=10, C=1, \mu_b=1, \mu_v=0.5, T_r=8$）

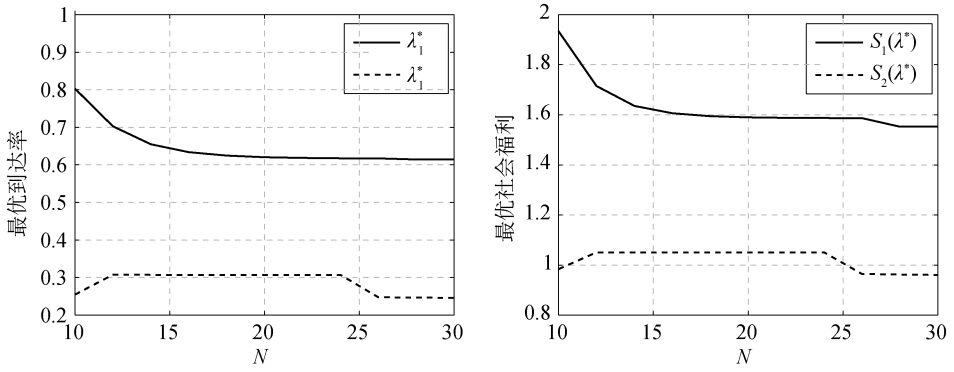

图 4-47　空竭与非空竭排队系统的最优到达率和最优社会福利（包括系统费用）关于 N 的敏感度（$R=10, C=1, c_r=0.1, c_s=0.5, \mu_b=1, \mu_v=0.5, T_r=8$）

4.6　相关文献评述

最初的关于排队系统中工作休假策略的详细描述是由 Servi 和 Finn[9]给出的。Servi 和 Finn[9]采用古典生灭过程方法研究了多重工作休假的 M/M/1 排队，得到了稳态队长和等待时间。之后，Wu 与 Takagi[72]等文献进一步扩展了 M/M/1 工作休假排队，使用嵌入 Markov 链和复函数围道积分研究了多重工作休假的 M/G/1 排队系统。接下来，应用拟生灭过程理论及方法，李继红、田乃硕等[71,73,74]分别分析了具有多重工作休假的 Markov 排队系统、M/G/1 排队系统，以及具有单重工作休假的 GI/M/1 排队系统，得到了稳态队长分布与随机分解结果。更多关于工作休假的介绍及相关文献请参考田乃硕等[10]给出的综述以及李继红[11]的博士学位论文。

从经济学的研究角度来看，针对空竭服务系统，孙微等[67,68,70]、王金亭等[75,76]和 Guha 等[77]分别得到了具有多重工作休假、单重工作休假、双阶段工作休假和双级工作休假的排队系统中顾客的均衡止步策略。针对非空竭服务系统，Dimitrakopoulos 和 Burnetas[66]考虑了具有工作休假和阈值策略的队长不可视 M/M/1 排队系统中顾客的均衡止步行为。孙微等[69]研究了具有工作休假和阈值策略的有限容量排队系统中顾客的均衡和最优止步策略，并研究了系统费用对顾客止步行为和最优社会福利的影响。

第 5 章　具有中途退出行为的排队系统

5.1　中途退出行为

在实际的随机服务系统中，顾客除了在进入队列前可能会产生止步行为之外，由于一些外部因素的影响，也有可能迫使或诱使其产生中途退出队列的行为。比如，在服务员休假期或慢速服务期（工作休假期），一些外生的退出机遇会有规律的到达系统，或者，服务员在提供服务的过程中，可能会由于自身原因或不可抗拒的外力原因而导致其失去服务能力。在这些时刻，在场顾客均将面临是否中途退出系统以及以何种方式退出系统的选择。

本章将研究由上述这些随机性外因导致的顾客退出行为对系统性能的影响，尤其是对顾客止步行为的影响，从而强调顾客不同阶段行为之间的制约关系，凸显顾客在进入复杂系统前的能动性与前瞻性。

本章内容为作者研究成果[78,79]的进一步完善和补充。

5.2　具有不可靠服务员与维修员的排队系统

服务员因自身或外力原因失效后，在场顾客面临多种选择，如全部离开（“清场”系统）、部分离开（同时独立决策或顺序独立决策）及全部等待。在服务员维修期间，维修员仅以一定的概率保证一次性修理成功，因此，单次修理过程完成后，在场顾客鉴于维修结果仍面临类似选择。本节将对其中的两类模型中顾客的止步行为进行对比。

5.2.1　模型描述

分析两类具有不可靠服务员和不可靠维修员的 Markov 单服务员排队系统，他们的区别在于分别采用清场机制（服务员因不可抗力的灾难性原因而导致失效）和非

清场机制（服务员因自身寿命终结而导致失效）。假定顾客的潜在到达率 Λ 足够大，实际到达率为 λ，服务员服务率为 μ。在服务员处于忙期或闲期时，可能会由于外力原因或自身原因导致服务员失效，失效的发生过程服从参数为 η 的 Poisson 过程。

为了便于比较，我们假定外来不可抗力的灾难性原因的到达间隔与服务员的寿命均服从同样的参数为 η 的 Poisson 分布。在每个失效时刻，对于清场系统，所有在场顾客均立即离场，而对于非清场系统，所有在场顾客均与服务员一起经历一个修理时间。在修理期间，顾客可正常到达系统。在经历一个服从参数为 γ 的指数分布的修理时间后，服务员仅以概率 $p(0<p<1)$恢复正常。一旦服务员修理失败（概率为 $q=1-p$），所有在场顾客均离场，同时下一个修理时间开始。

令$(L_i(t),I_i(t))(i=1, 2)$分别表示清场系统和非清场系统中时刻 t 的系统状态，其中 $L_i(t)$表示时刻 t 的系统队长，$I_i(t)$表示时刻 t 的服务员状态，且

$$I_i(t)=\begin{cases}0, \text{服务员处于修理期}\\ 1, \text{服务员处于正常状态}\end{cases}$$

我们考虑完全不可视排队系统，即到达顾客观测不到系统状态 $L_i(t)$与 $I_i(t)$。

对于每个进入系统的顾客，他服务完成可获得服务收益 R（如果中途退出的话将一无所获），但也需承担单位等待损耗 c。我们采用线性等待损耗函数，且假定到达过程、失效过程、服务过程和修理时间相互独立，服务顺序为先到先服务。

5.2.2　非清场排队系统

对于非清场排队系统，为了分析顾客止步行为，首先设法得到稳态队长分布的母函数。显然，过程$\{L_1(t), I_1(t)\}$为一个拟生灭过程，且状态空间为 $\Omega_1=\{(k,j)：k\geq 0, j=0, 1\}$。令$(L_1, I_1)$为过程$\{L_1(t), I_1(t)\}$的稳态极限。记稳态队长分布为

$$\boldsymbol{\pi}=(\boldsymbol{\pi}_0,\boldsymbol{\pi}_1),\boldsymbol{\pi}_j=(\pi_{0,j},\pi_{1,j},\cdots),j=0,1$$

$$\pi_{k,j}=P\{L_1=k,I_1=j\}=\lim_{t\to\infty}P\{L_1(t)=k,I_1(t)=j\},\quad (k,j)\in\Omega_1$$

由状态转移图 5-1 可知，稳态状态转移概率方程为

$$(\lambda+\gamma)\pi_{0,0}=\eta\pi_{0,1}+\gamma q\sum_{j=0}^{\infty}\pi_{j,0} \tag{5-1}$$

$$(\lambda+\gamma)\pi_{n,0}=\eta\pi_{n,1}+\lambda\pi_{n-1,0},\qquad n\geq 1 \tag{5-2}$$

$$(\lambda+\eta)\pi_{0,1}=\mu\pi_{1,1}+\gamma p\pi_{0,0} \tag{5-3}$$

$$(\lambda+\mu+\eta)\pi_{n,1}=\lambda\pi_{n-1,1}+\mu\pi_{n+1,1}+\gamma p\pi_{n,0},\quad n\geq 1 \tag{5-4}$$

定义系统稳态队长分布的母函数为$\Pi(z)$，状态 0 和状态 1 下的偏概率母函数分别为$\Pi_0(z)$和$\Pi_1(z)$。显然，$\Pi(z)=\Pi_0(z)+\Pi_1(z)$，且

$$\Pi_0(z)=\sum_{n=0}^{\infty}\pi_{n,0}z^n,\quad \Pi_1(z)=\sum_{n=0}^{\infty}\pi_{n,1}z^n$$

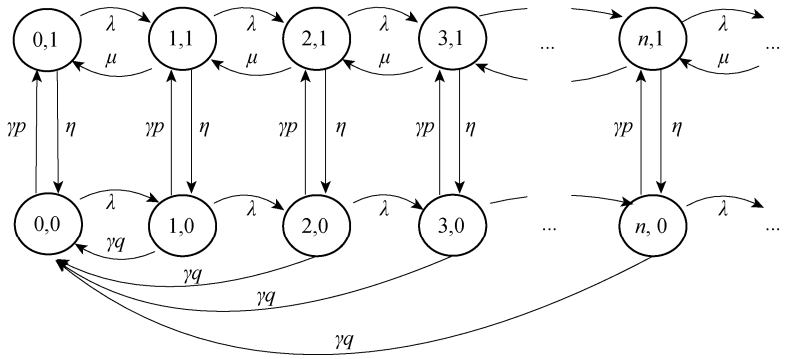

图 5-1　非清场系统的状态转移图

定理5.2.1　对于具有不可靠服务员和不可靠维修员的非清场 M/M/1 排队系统，稳态队长分布的偏概率母函数 $\Pi_0(z)$ 和 $\Pi_1(z)$ 分别为

$$\begin{aligned}\Pi_0(z)=&\frac{\eta\gamma q}{(\eta+\gamma p)(\lambda+\gamma-\lambda z)}\left(1+\gamma p\eta\left(\frac{z_0(1-z)}{(1-z_0)(\lambda+\gamma-\lambda z_0)}-\frac{z}{\lambda+\gamma-\lambda z}\right)\right.\\&\left.\times\left(\frac{\gamma p\eta z}{\lambda+\gamma-\lambda z}+\lambda z^2-(\lambda+\mu+\eta)z+\mu\right)^{-1}\right)\\\Pi_1(z)=&\frac{\gamma^2 pq\eta}{\eta+\gamma p}\left(\frac{z_0(1-z)}{(1-z_0)(\lambda+\gamma-\lambda z_0)}-\frac{z}{\lambda+\gamma-\lambda z}\right)\\&\times\left(\frac{\gamma p\eta z}{\lambda+\gamma-\lambda z}+\lambda z^2-(\lambda+\mu+\eta)z+\mu\right)^{-1}\end{aligned}\tag{5-5}$$

其中 z_0 为以下方程

$$\lambda^2z^3-\lambda(2\lambda+\mu+\eta+\gamma)z^2+((\lambda+\mu+\eta)(\lambda+\gamma)+\mu\lambda-\gamma p\eta)z-\mu(\lambda+\gamma)=0\tag{5-6}$$

的解且 $0<z_0<1$。

证明　将式（5-1）和式（5-2）关于 $n\geq 0$ 求和，得到

$$\gamma p\Pi_0(1)=\eta\Pi_1(1)\tag{5-7}$$

结合式（5-7）与平衡方程 $\Pi_0(1)+\Pi_1(1)=1$，得到

$$\Pi_0(1)=\frac{\eta}{\eta+\gamma p},\qquad \Pi_1(1)=\frac{\gamma p}{\eta+\gamma p}\tag{5-8}$$

分别用 z^0 和 z^n 与式 (5-1) 和式 (5-2) 相乘，再关于 $n\geq 0$ 求和，然后代入式 (5-8)，得到

$$\Pi_0(z)=\frac{\eta}{\lambda(1-z)+\gamma}\left(\Pi_1(z)+\frac{\gamma q}{\eta+\gamma p}\right) \tag{5-9}$$

类似地，分别用 z^0 和 z^n 与式（5-3）和式（5-4）相乘，再关于 $n\geq 0$ 求和，得到

$$\Pi_1(z)\left(\lambda z^2-z(\lambda+\mu+\eta)+\frac{\gamma p\eta z}{\lambda(1-z)+\gamma}+\mu\right)=\mu(1-z)\pi_{0,1}-\frac{\gamma^2 pq\eta}{\eta+\gamma p}\frac{z}{\lambda(1-z)+\gamma} \tag{5-10}$$

显然，关于 z 的三次方程

$$\lambda z^2-z(\lambda+\mu+\eta)+\frac{\gamma p\eta z}{\lambda(1-z)+\gamma}+\mu=0 \tag{5-11}$$

具有三个根。定义函数

$$f(z)=\lambda^2 z^3-\lambda(2\lambda+\mu+\eta+\gamma)z^2+((\lambda+\mu+\eta)(\lambda+\gamma)+\mu\lambda-\gamma p\eta)z-\mu(\lambda+\gamma)$$

有

$$f(0)=-\mu(\lambda+\gamma)<0,\qquad f(1)=\eta\gamma q>0$$

所以有一个根 $z_0\in(0,\ 1)$且很容易验证它的唯一性。在式（5-10）中，令 $z=z_0$（注意到：因为 $0<z_0<1$，所以 $\Pi_1(z_0)$ 收敛），得到

$$\pi_{0,1}=\frac{\gamma^2 pq\eta}{\eta+\gamma p}\frac{z_0}{\mu(1-z_0)(\lambda(1-z_0)+\gamma)} \tag{5-12}$$

首先将式（5-12）代入式（5-10），可求解出 $\Pi_1(z)$。然后，将 $\Pi_1(z)$ 代入式（5-9），可求解出 $\Pi_0(z)$。

鉴于定理 5.2.1 中的结果，非清场系统中的稳态平均队长 $E[L_1]$为

$$\begin{aligned}E[L_1]=\Pi_0'(1)+\Pi_1'(1)=&\frac{p(\gamma+\eta)}{\gamma(\eta+\gamma p)}\left(\lambda+\gamma+\frac{\gamma^2 z_0}{(1-z_0)(\lambda(1-z_0)+\gamma)}\right)\\&+\frac{p(\gamma+\eta)}{\eta(\eta+\gamma p)q}\left(\lambda-\mu-q\eta+\frac{p\eta\lambda}{\gamma}\right)+\frac{\lambda\eta}{\gamma(\eta+\gamma p)}\end{aligned} \tag{5-13}$$

因此，稳态下任意顾客的平均逗留时间为 $E[W_1]=E[L_1]/\lambda$。然而，对于一个进入系统的标记顾客，他可能面临两种结果：服务完成离开系统（情况 1）或由于维修员修理失败而退出系统（情况 2）。现在，我们设法导出它在两种情况下的条件平均逗留时间。定义该顾客在情况 i 下的逗留时间为 W_{1i}，Laplace 变换为 $W_{1i}^*(s)(i=1,2)$，首先，在情况 1 下，该顾客最终完成服务，所以有

$$\begin{aligned}W_{11}^*(s)=&\sum_{n=0}^{\infty}\pi_{n,1}\sum_{j=0}^{\infty}\int_0^{\infty}\frac{(\eta t)^j}{j!}e^{-\eta t}\frac{\mu(\mu t)^n}{n!}e^{-\mu t}\mathrm{d}t\left(\frac{\gamma p}{\gamma+s}\right)^j\left(\frac{\mu}{\mu+s}\right)^{n+1}\\&+\frac{\gamma p}{\gamma+s}\sum_{n=0}^{\infty}\pi_{n,0}\sum_{j=0}^{\infty}\int_0^{\infty}\frac{(\eta t)^j}{j!}e^{-\eta t}\frac{\mu(\mu t)^n}{n!}e^{-\mu t}\mathrm{d}t\left(\frac{\gamma p}{\gamma+s}\right)^j\left(\frac{\mu}{\mu+s}\right)^{n+1}\end{aligned}$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty}\pi_{n,1}\left(\frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}\right)^{n+1}\sum_{j=0}^{\infty}\left(\frac{\eta\gamma p}{\gamma+s}\frac{1}{\eta+\mu}\right)^{j}\frac{(j+n)!}{j!n!} \\
&\quad + \frac{\gamma p}{\gamma+s}\sum_{n=0}^{\infty}\pi_{n,0}\left(\frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}\right)^{n+1}\sum_{j=0}^{\infty}\left(\frac{\eta\gamma p}{\gamma+s}\frac{1}{\eta+\mu}\right)^{j}\frac{(j+n)!}{j!n!} \\
&= \frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}\left(\Pi_1\left(\frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}+\frac{\eta\gamma p}{\gamma+s}\frac{1}{\eta+\mu}\right)\right. \\
&\quad \left. + \frac{\gamma p}{\gamma+s}\Pi_0\left(\frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}+\frac{\eta\gamma p}{\gamma+s}\frac{1}{\eta+\mu}\right)\right)
\end{aligned}
\tag{5-14}
$$

因此，顾客最终完成服务离开系统的概率 p_{11} 为

$$
\begin{aligned}
p_{11} &= \sum_{n=0}^{\infty}\pi_{n,1}\sum_{j=0}^{\infty}\int_0^{\infty}\frac{(\eta t)^j}{j!}e^{-\eta t}p^j\frac{\mu(\mu t)^n}{n!}e^{-\mu t}\mathrm{d}t \\
&\quad + \sum_{n=0}^{\infty}\pi_{n,0}\sum_{j=0}^{\infty}\int_0^{\infty}\frac{(\eta t)^j}{j!}e^{-\eta t}p^{j+1}\frac{\mu(\mu t)^n}{n!}e^{-\mu t}\mathrm{d}t \\
&= \frac{\mu}{\eta+\mu}\left(\Pi_1\left(\frac{\eta p+\mu}{\eta+\mu}\right)+p\Pi_0\left(\frac{\eta p+\mu}{\eta+\mu}\right)\right)
\end{aligned}
\tag{5-15}
$$

在式（5-5）中，令 $z=(\eta p+\mu)/(\eta+\mu)$ 即可得到 p_{11} 的表达式。定义顾客在情况 1 下的条件逗留时间的 Laplace 变换为 $\tilde{W}_{11}^{*}(s)$，所以有 $\tilde{W}_{11}^{*}(s)=W_{11}^{*}(s)/p_{11}$，且顾客在情况 1 下的平均逗留时间 $E[W_{11}]$ 为

$$
\begin{aligned}
E[W_{11}] &= -\tilde{W}_{11}^{*\prime}(0) \\
&= \frac{1}{\mu}+\left(\frac{\gamma+\eta p}{\gamma(\eta+\mu)}\Pi_1'\left(\frac{\eta p+\mu}{\eta+\mu}\right)+\frac{p(\gamma+\eta p)}{\gamma(\eta+\mu)}\Pi_0'\left(\frac{\eta p+\mu}{\eta+\mu}\right)+\frac{p}{\gamma}\Pi_0\left(\frac{\eta p+\mu}{\eta+\mu}\right)\right) \\
&\quad \times\left(\Pi_1\left(\frac{\eta p+\mu}{\eta+\mu}\right)+p\Pi_0\left(\frac{\eta p+\mu}{\eta+\mu}\right)\right)^{-1}
\end{aligned}
\tag{5-16}
$$

将式（5-5）关于 z 求导，令 $z=(\eta p+\mu)/(\eta+\mu)$，并将其结果与式（5-15）一并代入式（5-16），得到 $E[W_{11}]$。定义该顾客因服务员修理失败而离开系统的概率为 p_{10}，他在情况 2 下的平均逗留时间为 $E[W_{10}]$。显然，$p_{10}=1-p_{11}$，且

$$
E[W_1]=P_{11}E[W_{11}]+p_{10}E[W_{10}] \tag{5-17}
$$

所以根据式（5-13）、式（5-15）和式（5-16），可得到 $E[W_{10}]$的表达式。该顾客的平均剩余效用 $U_R^1(\lambda)$ 为

$$
U_R^1(\lambda)=p_{11}(R-cE[W_{11}])+p_{10}(-cE[W_{10}])=p_{11}R-cE[W_1] \tag{5-18}
$$

定义顾客正的均衡到达率为 λ_1^e。求解 $U_R^1(\lambda)=0$，可得到 λ_1^e。平均单位社会福利为

$$
SW_1(\lambda)=\lambda(p_{11}R-cE[W_1]) \tag{5-19}
$$

求解优化问题 $\max SW_1(\lambda)$，可得到顾客的最优到达率 λ_1^*。

有趣的是，图 5-2 表明 λ_1^e 和 λ_1^* 不一定随 p 递增，这似乎有悖常理。其实，这种现象的主要原因在于顾客的平均逗留时间随 p 增长很快，尤其是情况 2 下顾客的平均逗留时间更是如此，这就导致一些顾客放弃进入队列而止步离开，从而，λ_1^e 随 p 递减。为了最大化社会福利，只要 p 不是很小，社会计划者也希望 λ_1^* 随 p 递减。图 5-3 表明 λ_1^e 和 λ_1^* 均随 η 递减，且随着服务员失效发生频率的增加，所有顾客都将选择止步。另一方面，图 5-4 则表明 λ_1^e 和 λ_1^* 均随 γ 递增，但当它们超过某个域值后，通过提高修理速度来吸引更多顾客收效甚微。而且，显然，在图 5-2 至图 5-4 中，$\lambda_1^e \geq \lambda_1^*$ 均成立。

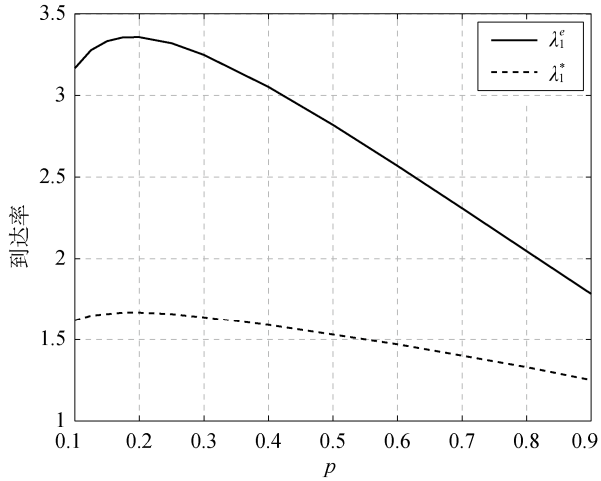

图 5-2　非清场系统中顾客的均衡和最优到达率关于 p 的敏感度（$R=10, c=1, \mu=2, \gamma=1, \eta=0.2$）

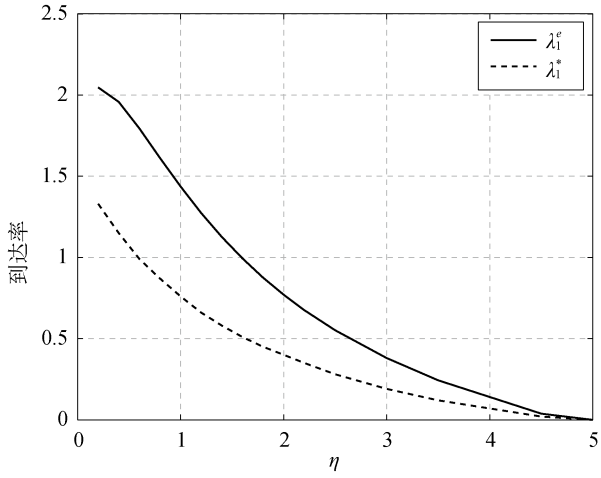

图 5-3　非清场系统中顾客的均衡和最优到达率关于 η 的敏感度（$R=10, c=1, \mu=2, \gamma=1, p=0.8$）

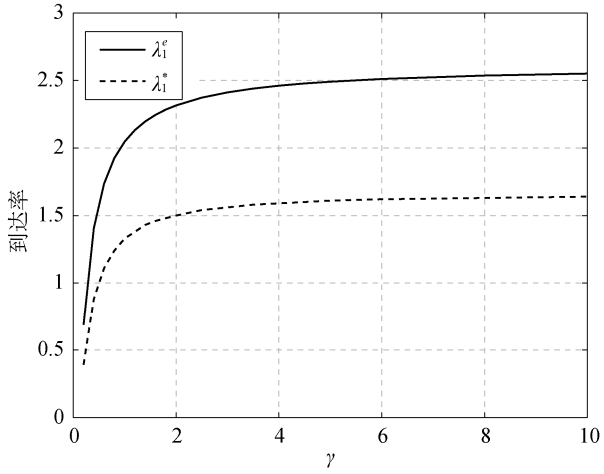

图 5-4　非清场系统中顾客的均衡和最优到达率关于 γ 的敏感度（$R=10, c=1, \mu=2, \eta=0.2, p=0.8$）

除了 p 相对较小的情况下，图 5-5 表明 $E[W_{10}^e]$ 几乎均大于 $E[W_{11}^e]$。事实上，只要 p 较大（如 $p=0.8$），顾客就有希望等待较短的时间而服务完成离开，而不是被迫清场。由于到达率随 η 锐减，图 5-6 表明 $E[W_{10}^e]$ 和 $E[W_{11}^e]$ 也随 η 递减。而且，图 5-7 表明，由于当修理率超过某个域值之后 λ_1^e 增长很慢，因此 $E[W_{10}^e]$ 和 $E[W_{11}^e]$ 几乎不再受 γ 的影响。

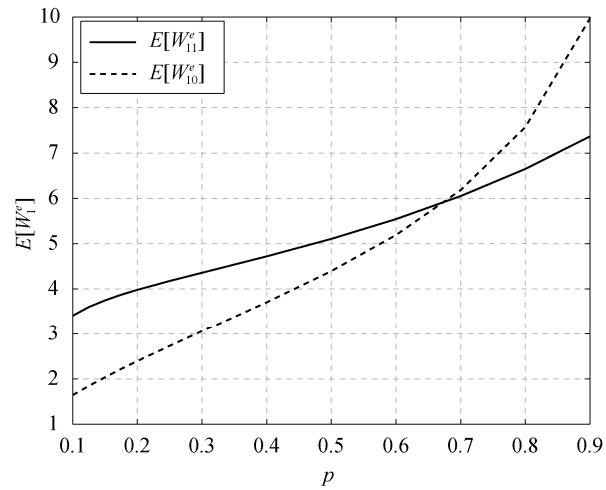

图 5-5　非清场系统中顾客的均衡逗留时间关于 p 的敏感度（$\mu=2, \gamma=1, \eta=0.2$）

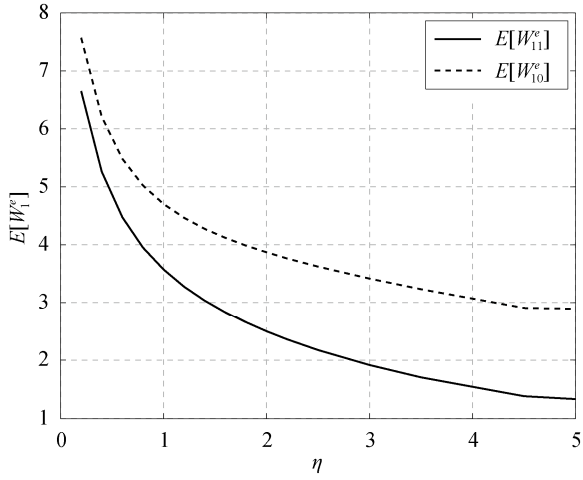

图 5-6　非清场系统中顾客的均衡逗留时间关于 η 的敏感度（$\mu=2,\gamma=1,p=0.8$）

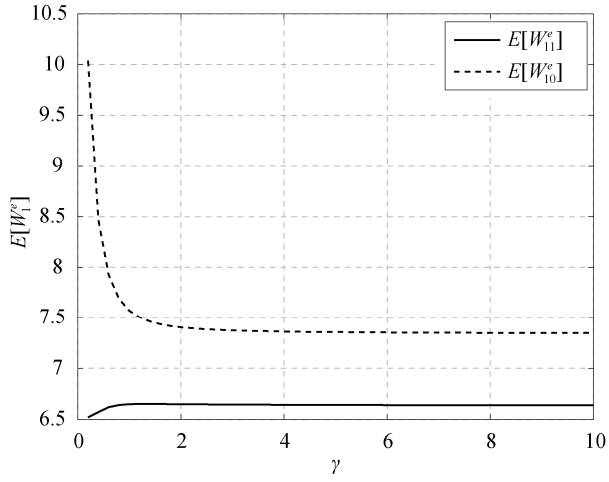

图 5-7　非清场系统中顾客的均衡逗留时间关于 γ 的敏感度（$\mu=2,\ \eta=0.2,p=0.8$）

5.2.3　清场排队系统

对于清场排队系统，过程$\{L_2(t),I_2(t)\}$仍为一个拟生灭过程，且状态空间仍然是Ω_1。类似地，首先需要得到稳态队长分布的母函数。根据状态转移图 5-8 所示，可得到状态转移概率方程为

$$(\lambda+\gamma)\pi_{0,0}=\eta\sum_{j=0}^{\infty}\pi_{j,1}+\gamma q\sum_{j=0}^{\infty}\pi_{j,0} \tag{5-20}$$

$$(\lambda+\gamma)\pi_{n,0}=\lambda\pi_{n-1,0},\qquad n\geq 1 \tag{5-21}$$

$$(\lambda+\eta)\pi_{0,1}=\mu\pi_{1,1}+\gamma p\pi_{0,0} \tag{5-22}$$

$$(\lambda+\mu+\eta)\pi_{n,1}=\lambda\pi_{n-1,1}+\mu\pi_{n+1,1}+\gamma p\pi_{n,0},\qquad n\geq 1 \tag{5-23}$$

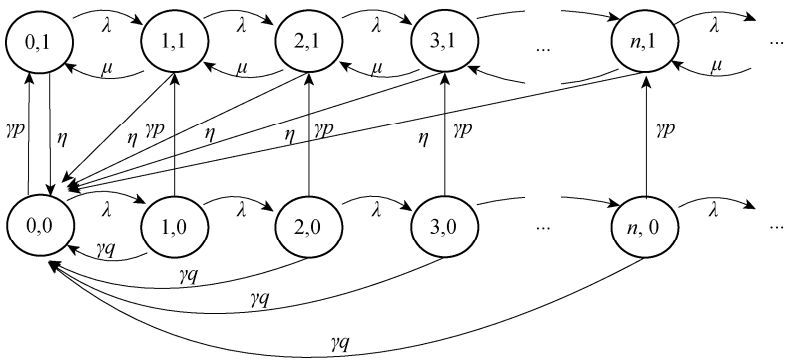

图 5-8　清场排队系统的状态转移图

定理 5.2.2　对于具有不可靠服务员和不可靠维修员的 M/M/1 清场排队系统，稳态队长分布的偏概率母函数 $\Pi_0(z)$ 和 $\Pi_1(z)$ 为

$$\begin{aligned}\Pi_0(z)&=\frac{\eta\gamma}{(\eta+\gamma p)(\lambda+\gamma-\lambda z)}\\ \Pi_1(z)&=\frac{\gamma^2 p\eta}{(\eta+\gamma p)(\lambda z^2-(\lambda+\mu+\eta)z+\mu)}\left(\frac{z_1(1-z)}{(1-z_1)(\lambda+\gamma-\lambda z_1)}-\frac{z}{\lambda+\gamma-\lambda z}\right)\end{aligned} \tag{5-24}$$

其中

$$z_1=\frac{(\lambda+\mu+\eta)-\sqrt{(\lambda+\mu+\eta)^2-4\lambda\mu}}{2\lambda} \tag{5-25}$$

证明　将式（5-20）和式（5-21）关于 $n\geq 0$ 求和，得到

$$\Pi_0(1)=\frac{\eta}{\eta+\gamma p},\qquad \Pi_1(1)=\frac{\gamma p}{\eta+\gamma p} \tag{5-26}$$

分别用 z^0 和 z^n 与式(5-20)和式(5-21)相乘, 再关于 $n\geq 0$ 求和, 然后代入式(5-26), 得到

$$\Pi_0(z)=\frac{\eta\gamma}{(\eta+\gamma p)(\lambda(1-z)+\gamma)} \tag{5-27}$$

类似地，分别用 z^0 和 z^n 与式（5-22）和式（5-23）相乘，再关于 $n\geq 0$ 求和，得到

$$\Pi_1(z)(\lambda z^2-z(\lambda+\mu+\eta)+\mu)=\mu(1-z)\pi_{0,1}-\frac{\gamma^2 p\eta}{\eta+\gamma p}\frac{z}{\lambda(1-z)+\gamma} \tag{5-28}$$

显然，关于 z 的二次方程

$$\lambda z^2-z(\lambda+\mu+\eta)+\mu=0 \tag{5-29}$$

有两个根,

$$z_1=\frac{(\lambda+\mu+\eta)-\sqrt{(\lambda+\mu+\eta)^2-4\lambda\mu}}{2\lambda}$$

为其中一个，且 $z_1\in(0,1)$。在式（5-28）中，令 $z=z_1$，得到

$$\pi_{0,1}=\frac{\gamma^2 p\eta}{\eta+\gamma p}\frac{z_1}{\mu(1-z_1)(\lambda(1-z_1)+\gamma)} \tag{5-30}$$

将式（5-30）代入式（5-28），可得到 $\Pi_1(z)$ 的表达式。

鉴于定理 5.2.2 中的结果，清场系统中的稳态平均队长 $E[L_2]$为

$$E[L_2]=\Pi_0'(1)+\Pi_1'(1)=\frac{1}{\eta+\gamma p}\left(\frac{\gamma^2 pz_1}{(1-z_1)(\lambda(1-z_1)+\gamma)}-\frac{\gamma p}{\eta}(\mu-\lambda)\right)+\frac{\lambda}{\gamma} \tag{5-31}$$

因此，稳态下任意顾客的平均逗留时间为 $E[W_2]=E[L_2]/\lambda$。类似于非清场系统, 对于一个进入系统的标记顾客, 也可能面临同样的两种结果 :服务完成离开系统(情况 1）或者由于服务员失效或维修员修理失败而退出系统（情况 2）。定义该顾客在情况 i 下的逗留时间为 W_{2i}，Laplace 变换为 $W_{2i}^*(s)(i=1,2)$。在情况 1 下，我们有

$$\begin{aligned}W_{21}^*(s)&=\sum_{n=0}^{\infty}\pi_{n,1}\int_0^{\infty}\frac{\mu(\mu t)^n}{n!}e^{-(\eta+\mu)t}\mathrm{d}t\left(\frac{\mu}{\mu+s}\right)^{n+1}\\&\quad+\frac{\gamma p}{\gamma+s}\sum_{n=0}^{\infty}\pi_{n,0}\int_0^{\infty}\frac{\mu(\mu t)^n}{n!}e^{-(\eta+\mu)t}\mathrm{d}t\left(\frac{\mu}{\mu+s}\right)^{n+1}\\&=\frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}\left(\Pi_1\left(\frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}\right)+\frac{\gamma p}{\gamma+s}\Pi_0\left(\frac{\mu}{\mu+s}\frac{\mu}{\eta+\mu}\right)\right)\end{aligned} \tag{5-32}$$

因此，顾客最终完成服务离开系统的概率 p_{21} 为

$$\begin{aligned}p_{21}&=\sum_{n=0}^{\infty}\pi_{n,1}\int_0^{\infty}\frac{\mu(\mu t)^n}{n!}e^{-(\eta+\mu)t}\mathrm{d}t+p\sum_{n=0}^{\infty}\pi_{n,0}\int_0^{\infty}\frac{\mu(\mu t)^n}{n!}e^{-(\eta+\mu)t}\mathrm{d}t\\&=\frac{\mu}{\eta+\mu}\left(\Pi_1\left(\frac{\mu}{\eta+\mu}\right)+p\Pi_0\left(\frac{\mu}{\eta+\mu}\right)\right)\end{aligned} \tag{5-33}$$

在式（5-24）中，令 $z=\mu/(\eta+\mu)$ 即可得到 p_{21} 的表达式。定义顾客在情况 1 下的条件逗留时间的 Laplace 变换为 $\tilde{W}_{21}^*(s)$，所以有 $\tilde{W}_{21}^*(s)=W_{21}^*(s)/p_{21}$，且顾客在情况 1 下的平均逗留时间 $E[W_{21}]$为

$$\begin{aligned}E[W_{21}]&=-\tilde{W}_{21}^{*\prime}(0)\\&=\frac{1}{\mu}+\left(\frac{1}{\eta+\mu}\left(\Pi_1'\left(\frac{\mu}{\eta+\mu}\right)+p\Pi_0'\left(\frac{\mu}{\eta+\mu}\right)\right)+\frac{p}{\gamma}\Pi_0\left(\frac{\mu}{\eta+\mu}\right)\right)\\&\quad\times\left(\Pi_1\left(\frac{\mu}{\eta+\mu}\right)+p\Pi_0\left(\frac{\mu}{\eta+\mu}\right)\right)^{-1}\end{aligned} \tag{5-34}$$

将式（5-24）关于 z 求导，令 $z=\mu/(\eta+\mu)$，并将其结果与式（5-33）一并代入式（5-34），得到 $E[W_{21}]$。定义该顾客因服务员失效或因维修员修理失败而离开系统的概率为 p_{20}，其在情况 2 下的平均逗留时间为 $E[W_{20}]$。显然，p_{20}=1-p_{21}，且

$$E[W_2]=p_{21}E[W_{21}]+p_{20}E[W_{20}] \tag{5-35}$$

所以根据式（5-31）、式（5-34）和式（5-35），可得到 $E[W_{20}]$的表达式。该顾客的平均剩余效用 $U_R^2(\lambda)$ 为

$$U_R^2(\lambda)=p_{21}R-cE[W_2] \tag{5-36}$$

定义顾客正的均衡到达率为 λ_2^e。求解 $U_R^2(\lambda)=0$，可得到 λ_2^e。单位社会福利为

$$SW_2(\lambda)=\lambda(p_{21}R-cE[W_2]) \tag{5-37}$$

求解最优化问题 $\max SW_2(\lambda)$，可得到顾客的最优到达率 λ_2^*。

图 5-9 表明 λ_2^e 和 λ_2^* 均随 p 递增但增速逐渐放缓。因此，由于采用清场机制，即使 p 值较大，对于顾客甚至社会计划者来说也并不具有很大的吸引力。相反，图 5-10 表明，λ_2^e 和 λ_2^* 显然会受到 η 的影响，且当 η 较小时，它们还可能随 η 递增。主要原因在于当 η 较小时，$E[W_{20}^e]$和 $E[W_{21}^e]$均随 η 递减。与图 5-4 类似，图 5-11 也说明当 γ 超过某个域值后，λ_2^e 和 λ_2^* 很难再有增长的空间，且 $\lambda_2^e\geq\lambda_2^*$。在图 5-12 至图 5-14 中，$E[W_{20}^e]\geq E[W_{21}^e]$始终成立，且 $E[W_{21}^e]$随 η 的增大有微小的反弹。事实上，即使 $E[W_{21}^e]$足够小，如果服务员失效过于频繁，顾客也不可能最终接受服务。

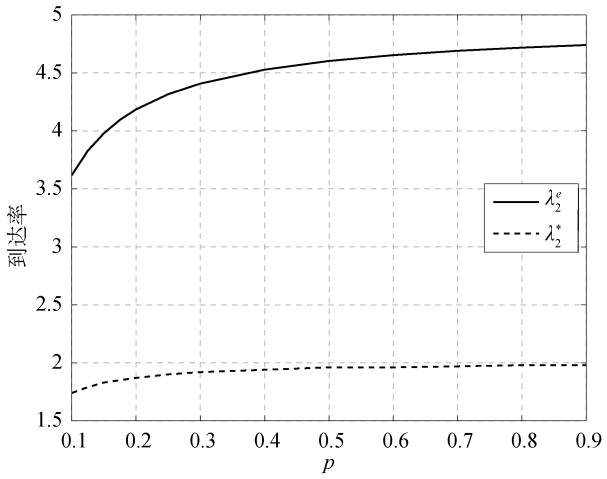

图 5-9　清场系统中顾客的均衡和最优到达率关于 p 的敏感度（$R=10, c=1, \mu=2, \gamma=1, \eta=0.2$）

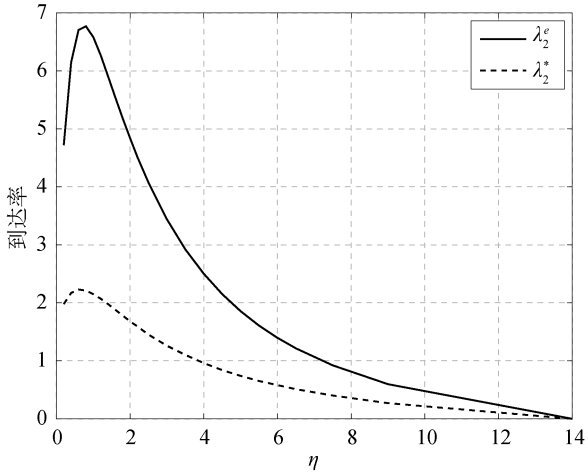

图 5-10　清场系统中顾客的均衡和最优到达率关于 η 的敏感度（$R = 10, c = 1, \mu = 2, \gamma = 1, p = 0.8$）

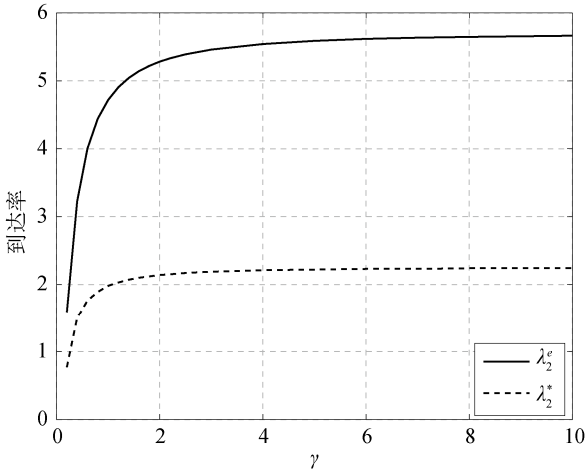

图 5-11　清场系统中顾客的均衡和最优到达率关于 γ 的敏感度（$R = 10, c = 1, \mu = 2, \eta = 0.2, p = 0.8$）

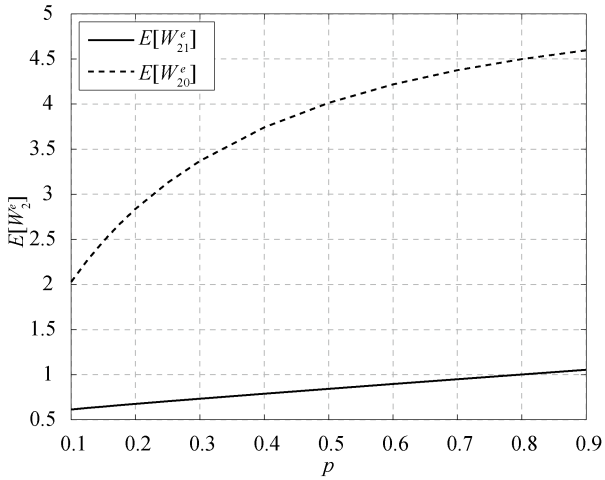

图 5-12　清场系统中顾客的均衡逗留时间关于 p 的敏感度（$\mu = 2, \gamma = 1, \eta = 0.2$）

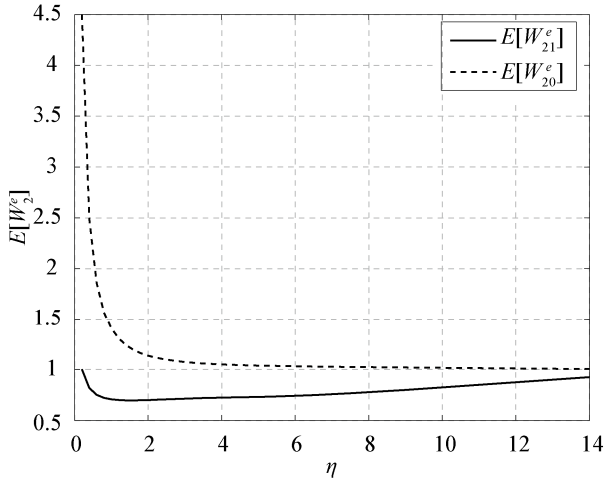

图 5-13　清场系统中顾客的均衡逗留时间关于 η 的敏感度（$\mu = 2, \gamma = 1, p = 0.8$）

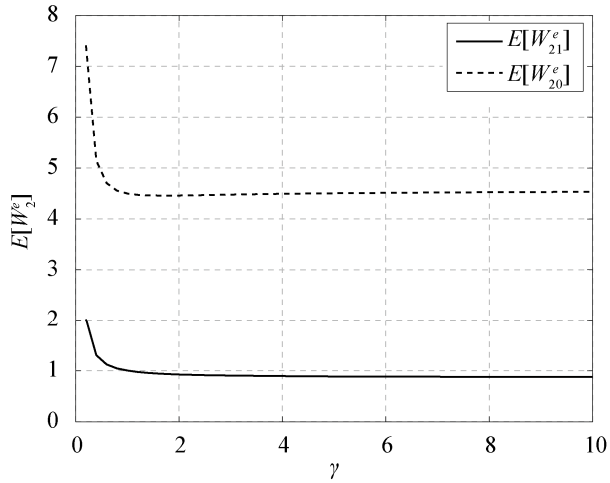

图 5-14　清场系统中顾客的均衡逗留时间关于 γ 的敏感度（$\mu=2,\eta=0.2,p=0.8$）

5.2.4　数值比较

针对以上清场系统和非清场系统中顾客的到达率和逗留时间的数值分析，表 5-1 总结且列举出了主要的对比结果。

表 5-1　清场系统和非清场系统中顾客的到达率和逗留时间的比较结果

<table>
<tr><td colspan="2">非清场系统</td><td colspan="2">清场系统</td></tr>
<tr><td>λ_1</td><td>$\lambda_1^e \geq \lambda_1^*$
如果 p 不是太小，λ_1^e 与 λ_1^* 随 p 递减
λ_1^e 与 λ_1^* 随 η 递减</td><td>$\lambda_2^e \geq \lambda_2^*$
λ_2^e 与 λ_2^* 随 p 递增
如果 η 较小，λ_2^e 与 λ_2^* 随 η 递增</td><td>λ_2</td></tr>
<tr><td colspan="4">λ_i^e 与 λ_i^* 均随 $\gamma(i=1,2)$ 递增
$\lambda_1^e \leq \lambda_2^e$ 且 $\lambda_1^* \leq \lambda_2^*$</td></tr>
<tr><td>$E[W_1^e]$</td><td>如果 p 不是太小，$E[W_{11}^e] \leq E[W_{10}^e]$
$E[W_{11}^e]$ 随 η 递减
$E[W_{11}^e]$ 随 γ 递增</td><td>$E[W_{21}^e] \leq E[W_{20}^e]$
如果 η 不是太小，$E[W_{21}^e]$ 随 η 递增
$E[W_{21}^e]$ 随 γ 递减</td><td>$E[W_2^e]$</td></tr>
<tr><td colspan="4">$E[W_{i1}^e]$ 与 $E[W_{i0}^e]$ 均随 p 递增，且 $E[W_{i0}^e]$ 增速更快$(i=1,2)$
$E[W_{11}^e] \geq E[W_{21}^e]$ 且 $E[W_{10}^e] \geq E[W_{20}^e]$</td></tr>
</table>

从表 5-1 可知，清场系统相对于非清场系统而言，顾客的均衡行为更加偏离社会最优行为。而且，在非清场系统中，无论顾客最终是否接受了服务，其在均衡状

态下的平均逗留时间相对更长，这使得顾客和社会计划者都倾向于采用清场机制。

除了比较顾客的到达率和逗留时间之外，我们继续比较一个顾客最终能够接受服务的概率。图 5-15 至图 5-17 表明 $p_{11}^e \geq p_{21}^e$ 始终成立。因此，显然，清场机制加大了顾客中途被清场的风险。而且，有趣的是，高修理速率对于顾客来说没有任何益处。也就是说，修理速率越高，顾客最终接受服务的概率越低。然而，图 5-18 至图 5-20 表明 $SW_1^* \leq SW_2^*$ 始终成立。因此，虽然在非清场系统中顾客最终被服务的概率更大，但是他们更喜欢加入清场系统，且社会计划者的观点也是如此。

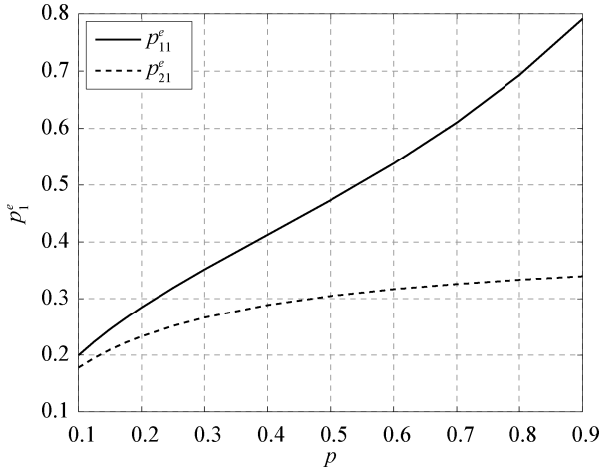

图 5-15　清场系统与非清场系统中顾客的均衡服务概率关于 p 的敏感度（$\mu = 2, \eta = 0.2, \gamma = 1$）

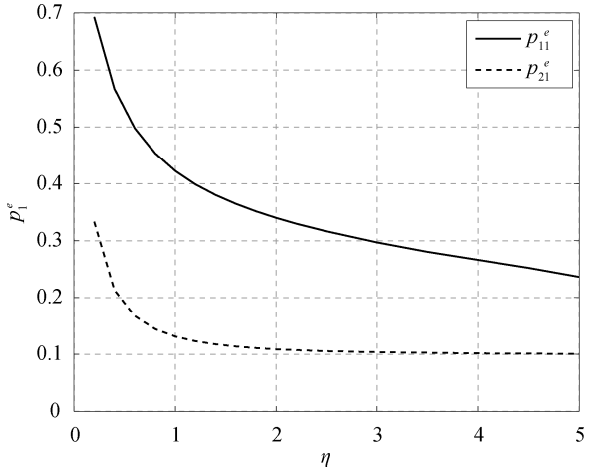

图 5-16　清场系统与非清场系统中顾客的均衡服务概率关于 η 的敏感度（$\mu = 2, \gamma = 1, p = 0.8$）

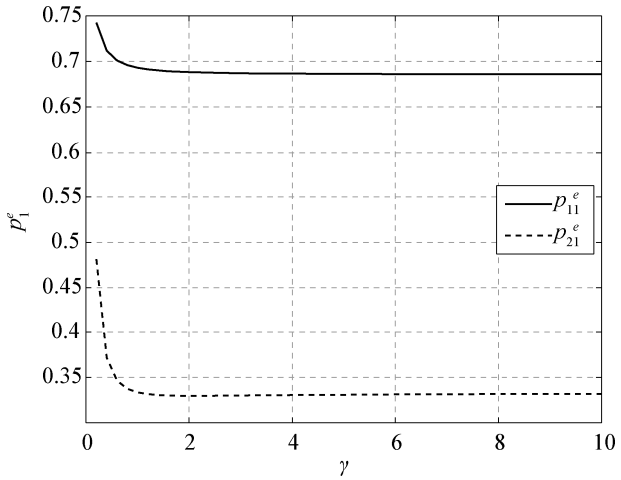

图 5-17　清场系统与非清场系统中顾客的均衡服务概率关于 γ 的敏感度（$\mu = 2, \eta = 0.2, p = 0.8$）

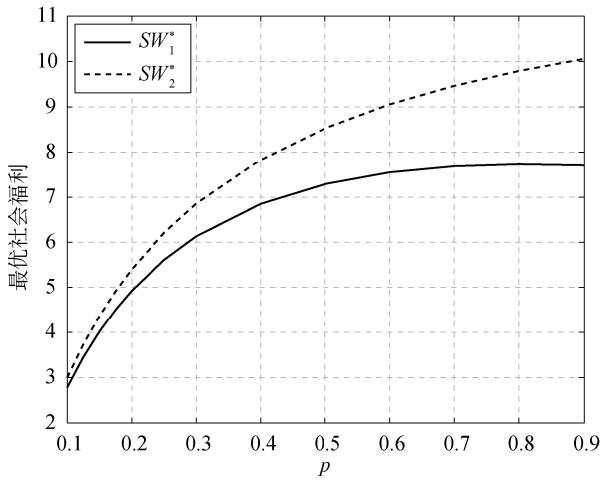

图 5-18　清场系统与非清场系统中的最优社会福利关于 p 的敏感度（$\mu = 2, \eta = 0.2, \gamma = 1$）

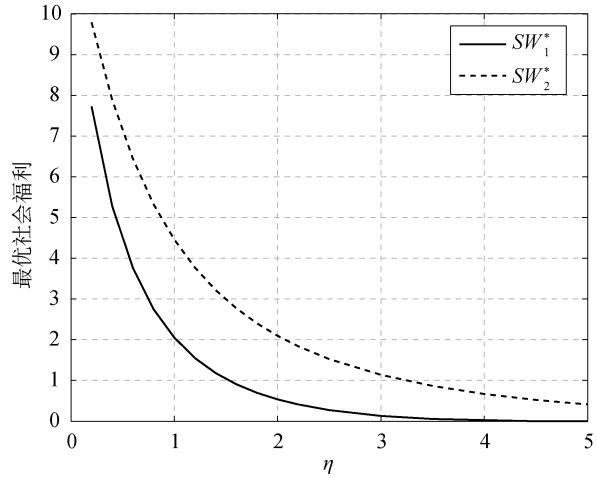

图 5-19　清场系统与非清场系统中的最优社会福利关于 η 的敏感度（$\mu = 2, \gamma = 1, p = 0.8$）

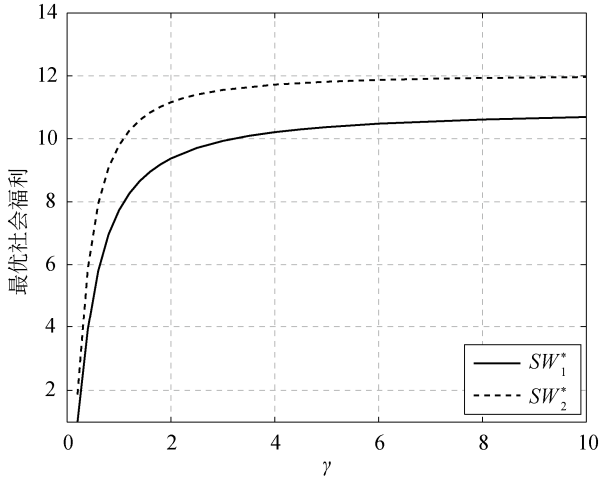

图 5-20　清场系统与非清场系统中的最优社会福利关于 γ 的敏感度（$\mu = 2, \eta = 0.2, p = 0.8$）

5.3　具有顺序退出行为与休假策略的排队系统

在服务员可靠的前提条件下，虽然服务速率控制是调节系统运作的有效方法，但在服务员休假期或服务慢速期，由于一些外部退出机遇的不断出现以及对未来等待损耗的不可预见性，顾客难免也会趁机产生中途退出系统的行为，尽管顾客已知退出收益（可能为零）必将显著低于进入系统前最初的服务收益。

5.3.1　模型描述

分析具有顺序退出行为和单重休假策略的 Markov 单服务员排队系统。假定潜在顾客的到达率 Λ 足够大，实际到达率为 λ，服务员服务率为 μ。一旦系统空竭，服务员开始一个服从参数为 γ 的指数分布的休假时间。休假结束时，如果系统中仍无顾客，服务员处于闲期并等待顾客到达。在休假期间，退出机遇的到达过程服从参数为 ξ 的 Poisson 过程。在退出机遇到达的瞬间，顾客顺次决定是否中途退出，且决策顺序与他们在队列中的位置相一致，即排在前面的顾客先做决策。每位顾客以概率 p 退出系统，以概率 q 继续排队，且 $p+q=1$。一旦有顾客决定继续排队或所有顾客均选择退出，退出过程结束。因此，在退出机遇到达时，退出顾客数服从参数为 p 的几何分布。

令$(L(t), I(t))$表示时刻 t 的系统状态，其中 $L(t)$表示系统队长，$I(t)$表示服务员状态，且

$$I(t)=\begin{cases}0, & \text{服务员处于休假期}\\ 1, & \text{服务员处于忙期或闲期}\end{cases}$$

我们考虑四种系统信息水平，即完全/几乎可视系统与几乎/完全不可视系统。

对于每个进入系统的顾客，服务完成可获得服务收益 R，中途退出可获得一部分补偿收益 $r(0\leq r<R)$，但也需承担单位等待损耗 c。我们采用线性等待损耗函数，且假定顾客到达过程，退出机遇到达过程，服务过程和休假时间相互独立，服务顺序为先到先服务。

5.3.2　完全可视排队系统

对于完全可视排队系统，我们可以很容易地得到顾客在忙期的均衡阈值止步策略。标记一个顾客，定义在他观测到系统状态$(n, 1)$后的平均逗留时间为 $E[W_1^{fo}(n)]$，

服务完成后的平均剩余效用为$U_1^{fo}(n)$，均衡止步阈值为n_1^e。因此，显然，

$$U_1^{fo}(n)=R-cE[W_1^{fo}(n)]=R-\frac{c(n+1)}{\mu} \tag{5-38}$$

所以

$$n_1^e=\left\lfloor\frac{R\mu}{c}\right\rfloor-1 \tag{5-39}$$

否则，如果顾客观测到系统状态$(n, 0)$，可能面临两种可能的结果：完成服务离开系统或在休假期中途退出。定义该顾客最终完成服务的概率为p_1^{fo}，中途退出的概率为p_0^{fo}，因此，得到

$$p_1^{fo}=\int_0^\infty e^{-\xi t}\gamma e^{-\gamma t}\mathrm{d}t+\int_0^\infty(1-e^{-\xi t})\sum_{k=0}^{n}p^k q\gamma e^{-\gamma t}\mathrm{d}t=1-\frac{\xi p^{n+1}}{\gamma+\xi} \tag{5-40}$$

$$p_0^{fo}=\int_0^\infty(1-e^{-\xi t})p^{n+1}\gamma e^{-\gamma t}\mathrm{d}t=\frac{\xi p^{n+1}}{\gamma+\xi} \tag{5-41}$$

定义该顾客逗留时间的 Laplace 变换为$W_0^{*fo}(s)$，且

$$\begin{aligned}W_0^{*fo}(s)=&\int_0^\infty e^{-\xi t}\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\left(\frac{\mu}{\mu+s}\right)^{n+1}\\&+\int_0^\infty(1-e^{-\xi t})\sum_{k=n+1}^{\infty}p^k q\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\\&+\int_0^\infty(1-e^{-\xi t})\sum_{k=0}^{\infty}p^k q\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\left(\frac{\mu}{\mu+s}\right)^{n+1-k}\\=&\frac{\gamma}{\gamma+s}\left(\frac{\gamma}{\gamma+\xi}\left(\frac{\mu}{\mu+s}\right)^{n+1}+\frac{\xi q}{\gamma+\xi}\left(\frac{\mu}{\mu+s}\right)^{n+1}\right.\\&\left.\times\frac{\mu^{n+1}-(p(\mu+s))^{n+1}}{\mu^{n+1}-\mu^n p(\mu+s)}+\frac{\xi p^{n+1}}{\gamma+\xi}\right)\end{aligned} \tag{5-42}$$

则该顾客的平均逗留时间$E[W_0^{fo}(n)]$为

$$E[W_0^{fo}(n)]=-W_0^{'*fo}(0)=\frac{n+1}{\mu}+\frac{1}{\gamma}-\frac{\xi p(1-p^{n+1})}{\mu q(\gamma+\xi)} \tag{5-43}$$

顾客的平均剩余效用$U_0^{fo}(n)$为

$$\begin{aligned}U_0^{fo}(n)&=p_1^{fo}R+p_0^{fo}r-cE[W_0^{fo}(n)]\\&=\left(1-\frac{\xi p^{n+1}}{\gamma+\xi}\right)R+\frac{\xi r p^{n+1}}{\gamma+\xi}-c\left(\frac{n+1}{\mu}+\frac{1}{\gamma}-\frac{\xi p(1-p^{n+1})}{\mu q(\gamma+\xi)}\right)\end{aligned} \tag{5-44}$$

定义顾客在休假期的均衡止步阈值为n_0^e。求解$U_0^{fo}(n)=0$，可得到n_0^e。至于n_0^e和n_1^e的大小关系，由于中途退出现象的存在，$n_0^e<n_1^e$不一定成立，二者的关系依赖于参数γ, ξ和p的取值。

图 5-21 和图 5-22 表明 $n_0^e < n_1^e$，但在图 5-23 和图 5-24 中，该关系式并不是始终成立的。而且，当 $p > 0.5$ 时，n_0^e 波动较大，且当 p 接近于 1 时，毕竟 $0 \le r < R$，所以 n_0^e 可能会降低。图 5-23 表明 n_0^e 也可能随 γ 降低，且当 γ 足够大时，$n_0^e = n_1^e$。也就是说，当休假时间很短时，服务员休假与否对顾客来说无异。

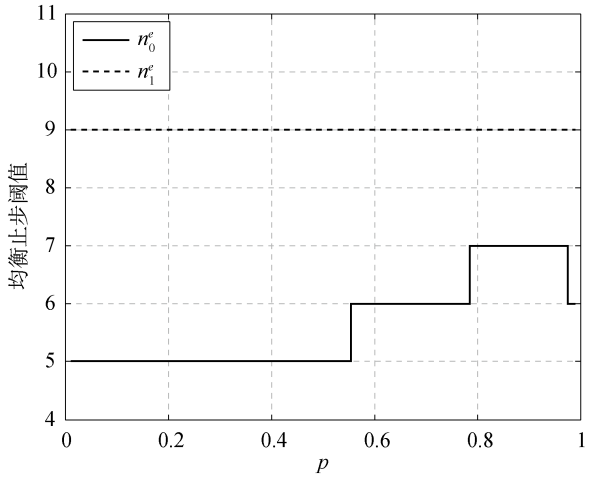

图 5-21　均衡止步阈值关于 p 的敏感度
（$R = 5, r = 2, c = 1, \mu = 2, \gamma = 0.5, \xi = 4$）

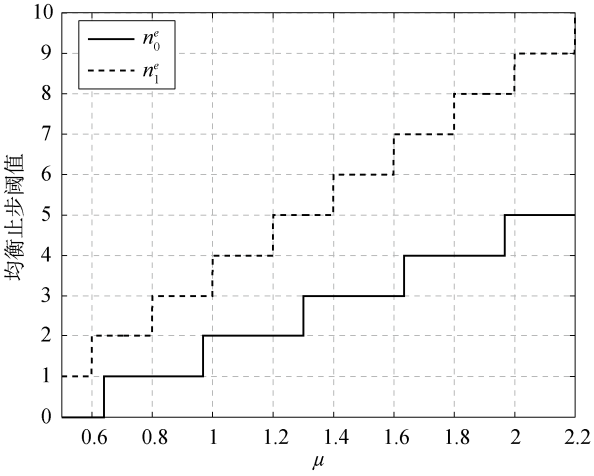

图 5-22　均衡止步阈值关于 μ 的敏感度
（$R = 5, r = 2, c = 1, \gamma = 0.5, \xi = 4, p = 0.1$）

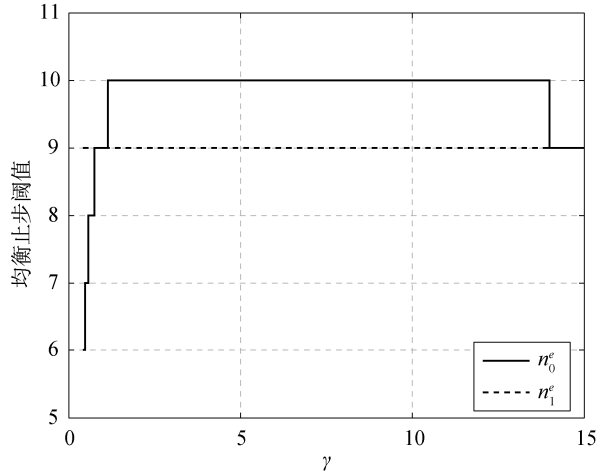

图 5-23　均衡止步阈值关于 γ 的敏感度

（$R=5, r=2, c=1, \mu=2, \xi=8, p=0.8$）

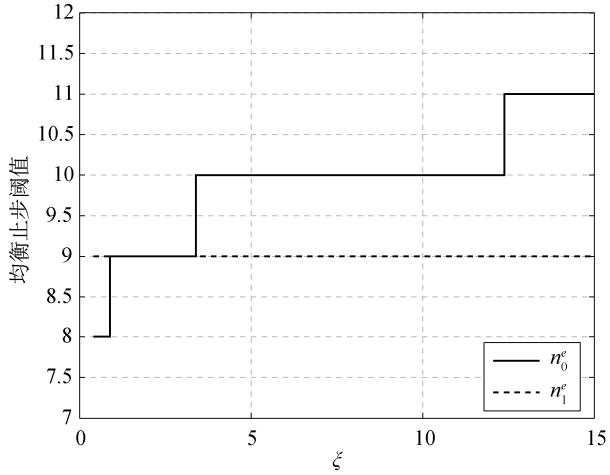

图 5-24　均衡止步阈值关于 ξ 的敏感度（$R=5, r=2, c=1, \mu=2, \gamma=3, p=0.8$）

5.3.3　几乎可视排队系统

对于几乎可视排队系统，首先设法导出稳态队长分布。在此，我们只对队长分布的推导思路加以说明。记顾客的均衡止步阈值为 n^e，根据状态转移图（如图 5-25 所示），得到均衡状态下系统状态转移概率方程为

$$(\lambda+\gamma+\xi)\pi_{0,0}=\mu\pi_{1,1}+\xi\sum_{j=0}^{n^e+1}p^j\pi_{j,0} \tag{5-45}$$

$$(\lambda+\gamma+\xi)\pi_{n,0}=\lambda\pi_{n-1,0}+\xi\sum_{j=n}^{n^e+1}p^{j-n}q\pi_{j,0},\quad 1\le n\le n^e \tag{5-46}$$

$$(\gamma+\xi p)\pi_{n^e+1,0}=\lambda\pi_{n^e,0} \tag{5-47}$$

$$\lambda\pi_{0,1}=\gamma\pi_{0,0} \tag{5-48}$$

$$(\lambda+\mu)\pi_{n,1}=\lambda\pi_{n-1,1}+\mu\pi_{n+1,1}+\gamma\pi_{n,0},\quad 1\le n\le n^e \tag{5-49}$$

$$\mu\pi_{n^e+1,1}=\lambda\pi_{n^e,1}+\gamma\pi_{n^e+1,0} \tag{5-50}$$

根据式(5-47)和递推公式(5-46)，我们可以倒序地得到概率$\{\pi_{n,0},0\le n\le n^e\}$。然后根据式（5-45）和式（5-48），可分别得到$\pi_{1,1}$和$\pi_{0,1}$的表达式。最后，根据式（5-49）、式（5-50）和$\pi_{1,1},\pi_{0,1}$，可以正序地得到概率$\{\pi_{n,1},2\le n\le n^e+1\}$。显然，所有概率均与$\pi_{n^e+1,0}$相关。对于$\pi_{n^e+1,0}$，可通过求解以下平衡方程

$$\sum_{n=0}^{n^e+1}(\pi_{n,0}+\pi_{n,1})=1 \tag{5-51}$$

得到。

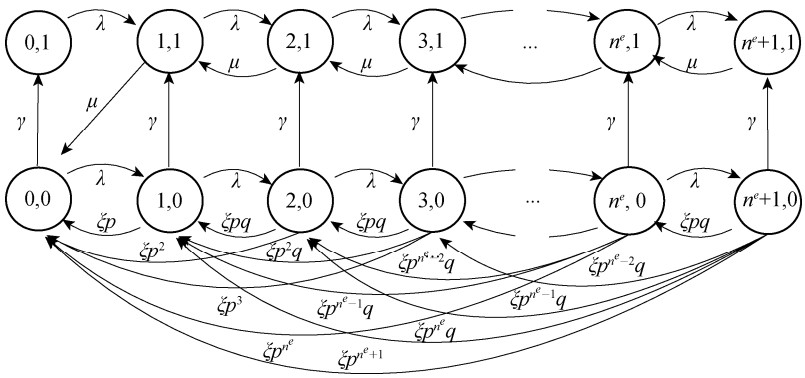

图 5-25　几乎可视排队系统的状态转移图

然后，分析顾客的均衡阈值止步策略。标记一顾客，该顾客观测到队长 n 且加入队列，定义顾客服务完成或退出系统后的平均剩余效用为$U^{ao}(n)$。根据式（5-38）和式（5-44），得到

$$\begin{aligned}U^{ao}(n)&=\frac{\pi_{n,1}}{\pi_{n,1}+\pi_{n,0}}U_1^{fo}(n)+\frac{\pi_{n,0}}{\pi_{n,1}+\pi_{n,0}}U_0^{fo}(n)\\&=\frac{\pi_{n,1}}{\pi_{n,1}+\pi_{n,0}}\left(R-\frac{c(n+1)}{\mu}\right)+\frac{\pi_{n,0}}{\pi_{n,1}+\pi_{n,0}}\left(\left(1-\frac{\xi p^{n+1}}{\gamma+\xi}\right)R+\frac{\xi rp^{n+1}}{\gamma+\xi}\right.\\&\quad\left.-c\left(\frac{n+1}{\mu}+\frac{1}{\gamma}-\frac{\xi p(1-p^{n+1})}{\mu q(\gamma+\xi)}\right)\right)\end{aligned} \tag{5-52}$$

求解$U^{ao}(n)=0$，得到n^e。

图 5-26 和图 5-27 表明n^e随μ或R递增，且递增率几乎恒定，即每当μ增加 0.2 个单位（或R增加 0.5 个单位），n^e增加一个单位。然而，我们数值证实了n^e几乎不随其他参数变动，如ξ, p或γ。因此，这说明发生在休假期间的顾客退出行为对顾客的止步策略几乎没有影响。

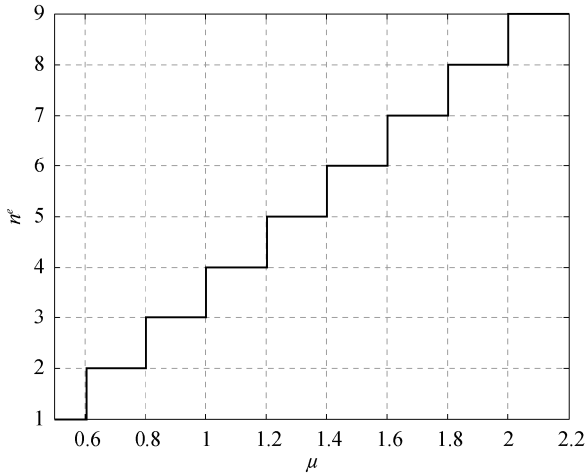

图 5-26　均衡止步阈值关于μ的敏感度

（$R=5, r=2, c=1, \lambda=3, \gamma=0.5, \xi=4, p=0.1$）

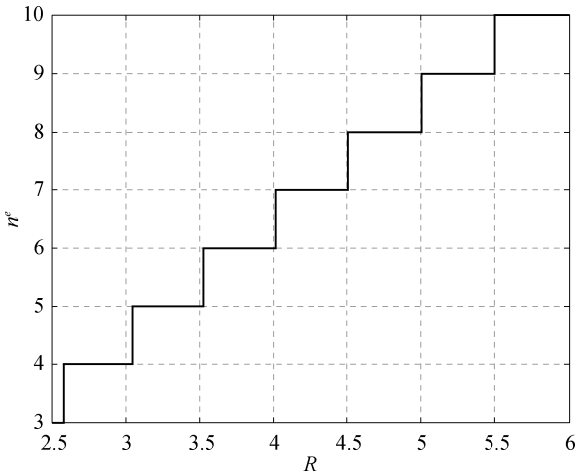

图 5-27　均衡止步阈值关于R的敏感度

（$r=2, c=1, \lambda=3, \mu=2, \gamma=0.5, \xi=4, p=0.1$）

5.3.4　几乎不可视排队系统

类似于几乎可视排队系统，对于几乎不可视排队系统，首先仍然是导出稳态队长分布。由于到达顾客可观测到服务员状态，所以顾客在状态 0 和状态 1 分别具有不同的实际到达率，分别记为 λ_0 和 λ_1。显然，过程$\{L(t),\ I(t)\}$为一个拟生灭过程，且状态空间为$\Omega=\{(k,j):k\geq 0,j=0,1\}$。令$(L,\ I)$为过程$\{L(t),\ I(t)\}$的稳态极限。记稳态队长分布为

$$\pi_{k,j}=P\{L=k,I=j\}=\lim_{t\to\infty}P\{L(t)=k,I(t)=j\},\quad (k,j)\in\Omega$$

根据状态转移图（如图 5-28 所示），可得到状态转移概率方程为

$$(\lambda_0+\gamma+\xi)\pi_{0,0}=\mu\pi_{1,1}+\xi\sum_{j=0}^{\infty}p^j\pi_{j,0} \tag{5-53}$$

$$(\lambda_0+\gamma+\xi)\pi_{n,0}=\lambda_0\pi_{n-1,0}+\xi\sum_{j=n}^{\infty}p^{j-n}q\pi_{j,0},\qquad n\geq 1 \tag{5-54}$$

$$\lambda_1\pi_{0,1}=\gamma\pi_{0,0}; \tag{5-55}$$

$$(\lambda_1+\mu)\pi_{n,1}=\lambda_1\pi_{n-1,1}+\mu\pi_{n+1,1}+\gamma\pi_{n,0},\qquad n\geq 1 \tag{5-56}$$

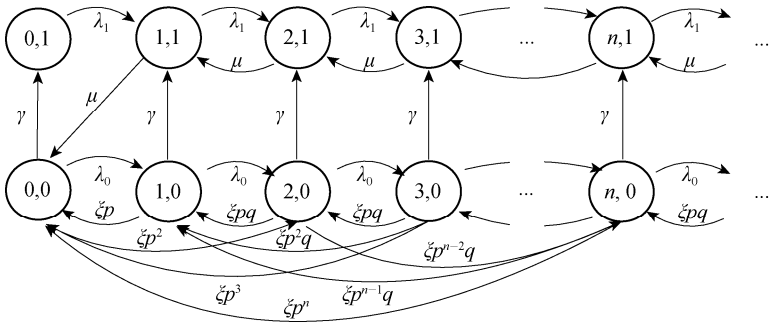

图 5-28　几乎不可视排队系统的状态转移图

定理 5.3.1　如果 $\rho=\lambda_1/\mu<1$，则稳态队长分布为

$$\begin{cases}\pi_{n,0}=\pi_{0,0}x_1^n, & n\geq 0\\ \pi_{n,1}=\left(\left(\dfrac{\gamma}{\lambda_1}-\dfrac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\rho^n+\dfrac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}x_1^n\right)\pi_{0,0}, & n\geq 0\end{cases} \tag{5-57}$$

其中

$$x_1=\frac{(\lambda_0(1+p)+\gamma+\xi p)-\sqrt{(\lambda_0(1+p)+\gamma+\xi p)^2-4p\lambda_0(\lambda_0+\gamma+\xi)}}{2p(\lambda_0+\gamma+\xi)} \tag{5-58}$$

且

$$\pi_{0,0}=\left(\frac{1}{1-x_1}+\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\frac{1}{1-\rho}+\frac{\gamma x_1}{(1-x_1)^2(\mu x_1-\lambda_1)}\right)^{-1} \tag{5-59}$$

证明 首先，重写式（5-54），得到

$$\xi q\sum_{j=n+1}^{\infty}p^{j-n}\pi_{j,0}-(\lambda_0+\gamma+\xi p)\pi_{n,0}+\lambda_0\pi_{n-1,0}=0,\qquad n\geq 1 \tag{5-60}$$

显然，它的特征方程是一个齐次线性差分方程

$$\xi q\sum_{j=n+1}^{\infty}p^{j-n}x^{j-n+1}-(\lambda_0+\gamma+\xi p)x+\lambda_0=0 \tag{5-61}$$

简化式（5-61），得到一个一元二次特征方程

$$p(\xi+\lambda_0+\gamma)x^2-(\lambda_0(1+p)+\xi p+\gamma)x+\lambda_0=0 \tag{5-62}$$

求解式（5-62），得到如下两个根：

$$x_{1,2}=\frac{(\lambda_0(1+p)+\gamma+\xi p)\mp\sqrt{(\lambda_0(1+p)+\gamma+\xi p)^2-4p\lambda_0(\lambda_0+\gamma+\xi)}}{2p(\lambda_0+\gamma+\xi)} \tag{5-63}$$

其中 $0\leq x_1<1$ 且 $x_2>1$，则式（5-54）的通解为

$$\pi_{n,0}=A_1x_1^n+B_1x_2^n,\qquad n\geq 0$$

其中 A_1 和 B_1 为未知系数。因为 $x_2>1$，所以 $B_1=0$，进而 $A_1=\pi_{0,0}$，即

$$\pi_{n,0}=\pi_{0,0}x_1^n,\qquad n\geq 0 \tag{5-64}$$

类似地，式（5-56）的特征方程为一个非齐次线性差分方程

$$\mu x^2-(\lambda_1+\mu)x+\lambda_1=\gamma\pi_{n,0}=\gamma x_1^n\pi_{0,0} \tag{5-65}$$

因此，式（5-56）的通解为 $\pi_{n,1}=\pi_{n,1}^{\text{hom}}+\pi_{n,1}^{\text{spec}}(n\geq 0)$，其中 $\pi_{n,1}^{\text{hom}}=A_2+B_2\rho^n$ 且 $\pi_{n,1}^{\text{spec}}$ 为式（5-56）的一个特解。由于式（5-65）的非齐次部分为参数 x_1 的几何形式，我们考虑一个形为 $\pi_{n,1}^{\text{spec}}=Cx_1^n$ 的特解。将其代入式（5-56），得到

$$C=\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\pi_{0,0} \tag{5-66}$$

根据式（5-53）和式（5-55），得到 $A_2=0$ 且

$$B_2=\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\pi_{0,0} \tag{5-67}$$

因此，

$$\pi_{n,1}=B_2\rho^n+Cx_1^n,\qquad n\geq 0 \tag{5-68}$$

且 $\pi_{0,0}$ 可通过求解平衡方程

$$\sum_{n=0}^{\infty}(\pi_{n,0}+\pi_{n,1})=1 \tag{5-69}$$

得到。

现在，在忙期和休假期，分别定义稳态队长分布的偏概率母函数为 $\Pi_0(z)$ 和 $\Pi_1(z)(|z|\leq 1)$。鉴于定理 5.3.1 的结果，得到

$$\begin{aligned}\Pi_0(z)&=\sum_{n=0}^{\infty}\pi_{n,0}z^n=\pi_{0,0}\sum_{n=0}^{\infty}(x_1z)^n\\&=\frac{1}{1-x_1z}\left(\frac{1}{1-x_1}+\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\frac{1}{1-\rho}+\frac{\gamma x_1}{(1-x_1)^2(\mu x_1-\lambda_1)}\right)^{-1}\end{aligned}\tag{5-70}$$

$$\begin{aligned}\Pi_1(z)&=\sum_{n=0}^{\infty}\pi_{n,1}z^n\\&=\pi_{0,0}\sum_{n=0}^{\infty}\left(\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)(\rho z)^n+\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}(x_1z)^n\right)\\&=\left(\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\frac{1}{1-\rho z}+\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\frac{1}{1-x_1z}\right)\\&\quad\times\left(\frac{1}{1-x_1}+\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\frac{1}{1-\rho}+\frac{\gamma x_1}{(1-x_1)^2(\mu x_1-\lambda_1)}\right)^{-1}\end{aligned}\tag{5-71}$$

如果标记顾客在忙期进入系统，他的条件平均逗留时间 $E[W_1^{au}(\lambda_0,\lambda_1)]$ 为

$$\begin{aligned}E[W_1^{au}(\lambda_0,\lambda_1)]&=\sum_{n=0}^{\infty}\pi_{n,1}\left(\frac{n+1}{\mu}\right)=\frac{1}{\mu}(\Pi_1'(1)+\Pi_1(1))\\&=\frac{1}{\mu}\left(\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\frac{\rho}{(1-\rho)^2}+\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\frac{x_1}{(1-x_1)^2}\right)\\&\quad\times\left(\frac{1}{1-x_1}+\left(\frac{\gamma}{\lambda_1}-\frac{\gamma x_1}{(1-x_1)(\mu x_1-\lambda_1)}\right)\frac{1}{1-\rho}+\frac{\gamma x_1}{(1-x_1)^2(\mu x_1-\lambda_1)}\right)^{-1}\end{aligned}\tag{5-72}$$

另一方面，如果标记顾客在休假期进入系统，类似于完全可视情形，可能面临两种结果：完成服务离开系统或在休假期中途退出。定义顾客逗留时间的 Laplace 变换为 $W_0^{*au}(s)$，则

$$\begin{aligned}W_0^{*au}(s)&=\sum_{n=0}^{\infty}\pi_{n,0}\left(\int_0^{\infty}(1-e^{-\xi t})\sum_{k=n+1}^{\infty}p^kq\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\right.\\&\quad+\int_0^{\infty}e^{-\xi t}\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\left(\frac{\mu}{\mu+s}\right)^{n+1}\\&\quad\left.+\int_0^{\infty}(1-e^{-\xi t})\sum_{k=0}^{\infty}p^kq\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\left(\frac{\mu}{\mu+s}\right)^{n+1-k}\right)\\&=\frac{\gamma}{(\gamma+s)(\gamma+\xi)}\left(\frac{\gamma\mu}{\mu+s}\Pi_0\left(\frac{\mu}{\mu+s}\right)\right.\end{aligned}\tag{5-73}$$

$$+\frac{\xi\mu q}{\mu q-sp}\left(\frac{\mu}{\mu+s}\Pi_0\left(\frac{\mu}{\mu+s}\right)-p\Pi_0(p)\right)+\xi p\Pi_0(p)\Bigg)$$

顾客的条件平均逗留时间 $E[W_0^{au}(\lambda_0)]$ 为

$$E[W_0^{au}(\lambda_0)]=\frac{-W_0'^{*au}(0)}{\Pi_0(1)}=\frac{1}{\gamma}+\frac{1}{\mu(1-x_1)}-\frac{p\xi}{\mu q(\gamma+\xi)}\left(1-\frac{p(1-x_1)}{1-px_1}\right) \tag{5-74}$$

定义该顾客最终完成服务的概率为 p_{01}^{au}，中途退出的概率为 p_{00}^{au}。因此，得到

$$\begin{aligned}p_{01}^{au}&=\sum_{n=0}^{\infty}\frac{\pi_{n,0}}{\Pi_0(1)}\left(\int_0^{\infty}e^{-\xi t}\gamma e^{-\gamma t}\mathrm{d}t+\int_0^{\infty}(1-e^{-\xi t})\sum_{k=0}^{n}p^kq\gamma e^{-\gamma t}\mathrm{d}t\right)\\&=1-\frac{p\xi}{\gamma+\xi}\frac{1-x_1}{1-px_1}\end{aligned} \tag{5-75}$$

$$p_{00}^{au}=\sum_{n=0}^{\infty}\frac{\pi_{n,0}}{\Pi_0(1)}\left(\int_0^{\infty}(1-e^{-\xi t})\sum_{k=n+1}^{\infty}p^kq\gamma e^{-\gamma t}\mathrm{d}t\right)=\frac{p\xi}{\gamma+\xi}\frac{1-x_1}{1-px_1} \tag{5-76}$$

显然，$p_{01}^{au}=1-p_{00}^{au}$。该顾客在休假期到达的平均剩余效用 $U_0^{au}(\lambda_0)$ 为

$$U_0^{au}(\lambda_0)=p_{01}^{au}R+p_{00}^{au}r-cE[W_0^{au}(\lambda_0)] \tag{5-77}$$

该顾客在忙期或闲期到达的平均剩余效用 $U_1^{au}(\lambda_0,\lambda_1)$ 为

$$U_1^{au}(\lambda_0,\lambda_1)=R-cE[W_1^{au}(\lambda_0,\lambda_1)] \tag{5-78}$$

分别定义顾客在忙期或休假期的正的均衡到达率为 λ_{s1}^e 和 λ_{s0}^e ①。求解 $U_0^{au}(\lambda_0)=0$ 和 $U_1^{au}(\lambda_0,\lambda_1)=0$，可得到 λ_{s1}^e 和 λ_{s0}^e。单位社会福利 $SW_s^{au}(\lambda_0,\lambda_1)$ 为

$$\begin{aligned}SW_s^{au}(\lambda_0,\lambda_1)=&\lambda_0\Pi_0(1)(p_{01}^{au}R+p_{00}^{au}r-cE[W_0^{au}(\lambda_0)])\\&+\lambda_1\Pi_1(1)(R-cE[W_1^{au}(\lambda_0,\lambda)])\end{aligned} \tag{5-79}$$

求解在约束 $\lambda_1<\mu$ 下的优化问题 $\max SW_s^{au}(\lambda_0,\lambda_1)$，可得到顾客的最优到达率 λ_{s0}^* 和 λ_{s1}^*。将其代入式（5-79），可得到最优社会福利 SW_s^{*au}。

图 5-29 和图 5-30 表明 λ_{s0}^e 随 ξ 线性递增，而随 p 凸增。比较图 5-30 和图 5-21，发现当 p 接近于 1 时，在同样的条件下，λ_{s0}^e 并不减少。这说明信息量的多少的确影响顾客的均衡止步策略。然而，λ_{s1}^e 并不受 ξ 或 p 的影响（$\lambda_{s0}^e\equiv1.643$）②。也就是说，在几乎不可视情形下，尽管在上一个休假期顾客的中途退出行为会影响下一个忙期开始时的队长，但该忙期内到达的顾客在做决策时，并不将这种退出行为加以考虑。而且，$\lambda_{s0}^e>\lambda_{s1}^e$，这种现象在没有退出行为的系统中一般是不可能出现的。

① λ_{s1}^e 和 λ_{s0}^e 的下标“s”是为了将其与多重休假系统中的相应指标相区分。

② 记多重休假系统中顾客的均衡到达率分别为 λ_{m1}^e 和 λ_{m0}^e，且已验证，在与图 5-29 和图 5-30 相同的条件下，$\lambda_{m0}^e=\lambda_{s0}^e$ 且 $\lambda_{m1}^e\equiv1.634<\lambda_{s1}^e\equiv1.634$。

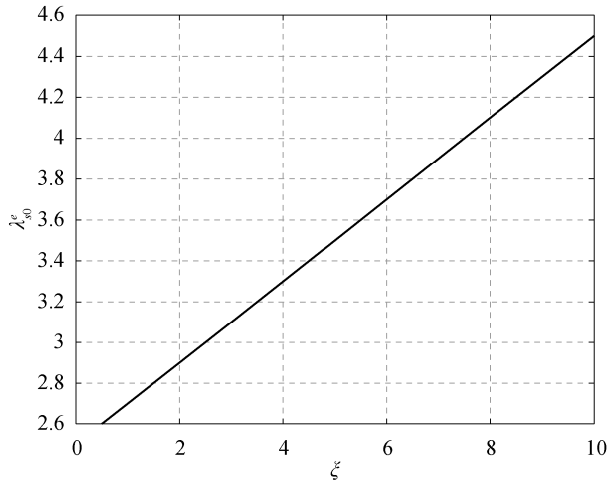

图 5-29　休假期间的均衡到达率关于 ξ 的敏感度
（$R=5, r=2, c=1, \mu=2, \gamma=0.5, p=0.2$）

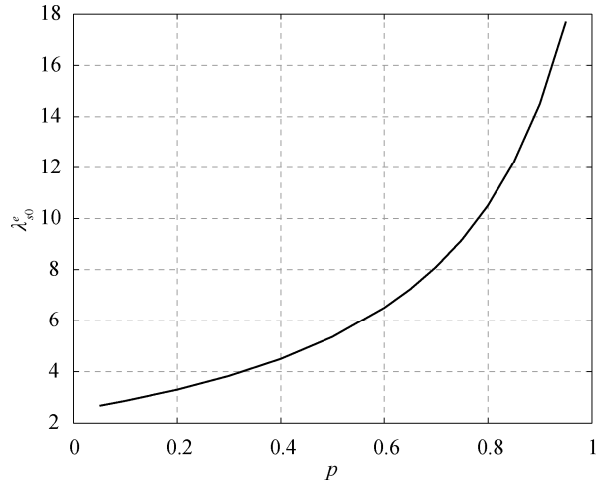

图 5-30　休假期间的均衡到达率关于 p 的敏感度
（$R=5, r=2, c=1, \mu=2, \gamma=0.5, \xi=4$）

5.3.5　完全不可视排队系统

在完全不可视排队系统中，到达顾客不能观测到系统状态，因此，只要将定理 5.3.1 中的 λ_0 和 λ_1 替换为 λ，即得到稳态队长分布，且状态转移图如图 5-31 所示。

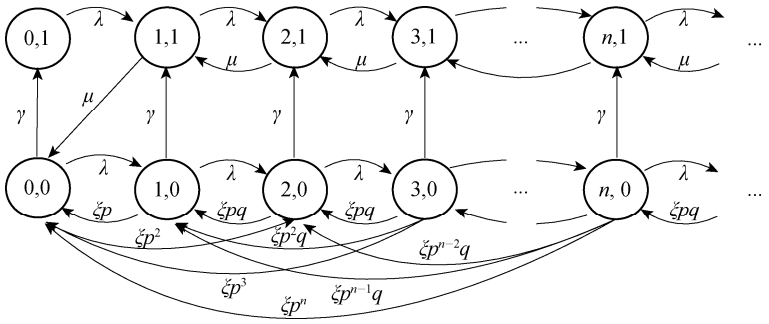

图 5-31　完全不可视排队系统的状态转移图

类似于完全可视和几乎不可视情形，标记一顾客，如果该顾客进入系统，他可能面临两种结果：完成服务离开系统或在休假期中途退出。定义该顾客最终完成服务的概率为 p_1^{fu}，中途退出的概率为 p_0^{fu}。因此，

$$\begin{aligned} p_1^{fu} &= \sum_{n=0}^{\infty}\pi_{n,1} + \sum_{n=0}^{\infty}\pi_{n,0}\left(\int_0^{\infty} e^{-\xi t}\gamma e^{-\gamma t}\mathrm{d}t + \int_0^{\infty}(1-e^{-\xi t})\sum_{k=0}^{\infty}p^k q\gamma e^{-\gamma t}\mathrm{d}t\right) \\ &= 1-\frac{p\xi}{\gamma+\xi}\frac{1}{1-px}\left(\frac{1}{1-x}+\left(\frac{\gamma}{\lambda}-\frac{\gamma x}{(1-x)(\mu x-\lambda)}\right)\frac{1}{1-\rho}+\frac{\gamma x}{(1-x)^2(\mu x-\lambda)}\right)^{-1} \end{aligned} \tag{5-80}$$

$$\begin{aligned} p_0^{fu} &= \sum_{n=0}^{\infty}\pi_{n,0}\left(\int_0^{\infty}(1-e^{-\xi t})\sum_{k=n+1}^{\infty}p^k q\gamma e^{-\gamma t}\mathrm{d}t\right) \\ &= \frac{p\xi}{\gamma+\xi}\frac{1}{1-px}\left(\frac{1}{1-x}+\left(\frac{\gamma}{\lambda}-\frac{\gamma x}{(1-x)(\mu x-\lambda)}\right)\frac{1}{1-\rho}+\frac{\gamma x}{(1-x)^2(\mu x-\lambda)}\right)^{-1} \end{aligned} \tag{5-81}$$

其中 $\rho=\lambda/\mu$ 且

$$x=\frac{(\lambda(1+p)+\gamma+\xi p)-\sqrt{(\lambda(1+p)+\gamma+\xi p)^2-4p\lambda(\lambda+\gamma+\xi)}}{2p(\lambda+\gamma+\xi)} \tag{5-82}$$

显然，$p_1^{fu}=1-p_0^{fu}$。定义顾客逗留时间的 Laplace 变换为 $W^{*fu}(s)$，则

$$\begin{aligned} W^{*fu}(s) = \sum_{n=0}^{\infty}\pi_{n,0}\Bigg(&\int_0^{\infty}(1-e^{-\xi t})\sum_{k=0}^{n}p^k q\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\left(\frac{\mu}{\mu+s}\right)^{n+1-k} \\ &+\int_0^{\infty}e^{-\xi t}\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\left(\frac{\mu}{\mu+s}\right)^{n+1} \\ &+\int_0^{\infty}(1-e^{-\xi t})\sum_{k=n+1}^{\infty}p^k q\gamma e^{-\gamma t}\mathrm{d}t\left(\frac{\gamma}{\gamma+s}\right)\Bigg)+\sum_{n=0}^{\infty}\pi_{n,1}\left(\frac{\mu}{\mu+s}\right)^{n+1} \\ =\Bigg(&\frac{\mu}{\mu+s}\left(\left(\frac{\gamma}{\lambda}-\frac{\gamma x}{(1-x)(\mu x-\lambda)}\right)\frac{\mu+s}{\mu+s-\lambda}+\frac{\gamma x}{(1-x)(\mu x-\lambda)}\frac{\mu+s}{\mu(1-x)+s}\right) \end{aligned}$$

$$+\frac{\gamma}{(\gamma+s)(\gamma+\xi)}\left(\frac{\gamma\mu}{\mu+s}\frac{\mu+s}{\mu(1-x)+s}+\frac{\xi\mu q}{\mu q-sp}\left(\frac{\mu}{\mu+s}\frac{\mu+s}{\mu(1-x)+s}-\frac{p}{1-px}\right)\right.$$

$$\left.+\frac{\xi p}{1-px}\right)\left(\frac{1}{1-x}+\left(\frac{\gamma}{\lambda}-\frac{\gamma x}{(1-x)(\mu x-\lambda)}\right)\frac{1}{1-\rho}+\frac{\gamma x}{(1-x)^2(\mu x-\lambda)}\right)^{-1}$$

该顾客的平均逗留时间 $E[W^{fu}(\lambda)]$ 为

$$\begin{aligned}E[W^{fu}(\lambda)]&=-W'^{*fu}(0)\\&=\left(\frac{1}{\mu}\left(\left(\frac{\gamma}{\lambda}-\frac{\gamma x}{(1-x)(\mu x-\lambda)}\right)\frac{\rho}{(1-\rho)^2}+\frac{\gamma x}{(1-x)(\mu x-\lambda)}\frac{x}{(1-x)^2}\right)\right.\\&\quad\left.+\left(\frac{1}{\gamma(1-x)}+\frac{1}{\mu(1-x)^2}-\frac{p\xi}{\mu q(\gamma+\xi)}\left(\frac{1}{1-x}-\frac{p}{1-px}\right)\right)\right)\\&\quad\times\left(\frac{1}{1-x}+\left(\frac{\gamma}{\lambda}-\frac{\gamma x}{(1-x)(\mu x-\lambda)}\right)\frac{1}{1-\rho}+\frac{\gamma x}{(1-x)^2(\mu x-\lambda)}\right)^{-1}\end{aligned}\tag{5-84}$$

顾客的平均剩余效用 $U^{fu}(\lambda)$ 为

$$U^{fu}(\lambda)=p_1^{fu}R+p_0^{fu}r-cE[W^{fu}(\lambda)]\tag{5-85}$$

定义顾客正的均衡到达率为 λ_s^e。求解 $U^{fu}(\lambda)=0$，可得到 λ_s^e。单位社会福利 $SW_s^{fu}(\lambda)$ 为

$$SW_s^{fu}(\lambda)=\lambda(p_1^{fu}R+p_0^{fu}r-cE[W^{fu}(\lambda)])\tag{5-86}$$

求解在约束 $\lambda<\mu$ 下的优化问题 $\max SW_s^{fu}(\lambda)$，可得到顾客的最优到达率 λ_s^*。将其代入式（5-86），可得到最优社会福利 SW_s^{*fu}。

5.3.6　数值比较

对于几乎不可视和完全不可视排队系统，我们比较单重休假和多重休假策略下顾客的均衡和最优止步行为以及最优社会福利[①]。

类似于单重休假系统中的符号定义，在几乎不可视的多重休假系统中，定义顾客在休假期和忙期的最优到达率分别为 λ_{m0}^* 和 λ_{m1}^*（λ_{m0}^e 和 λ_{m1}^e 已经在以上脚注中定义），以及最优社会福利为 SW_m^{*au}。然后，在完全不可视的多重休假系统中，定义顾客的均衡和最优到达率分别为 λ_m^e 和 λ_m^*，以及最优社会福利为 SW_m^{*fu}。

图 5-32 表明 λ_{s0}^* 和 λ_{m0}^* 均随 ξ 或 p 递增。与图 5-29 相比，我们发现最优到达率与均衡到达率相似，都具有递增的趋势，但是 λ_{s0}^* 和 λ_{m0}^* 不再相等，而是，$\lambda_{s0}^*<\lambda_{m0}^*$。另一方面，图 5-33 表明，除了 p 很大的情况，λ_{s1}^* 和 λ_{m1}^* 均随 ξ 或 p 递

① 对于完全（几乎）可视情形，顾客的均衡止步行为在两种休假策略下（几乎）无异。

减。这说明不同于顾客的均衡行为，社会计划者希望忙期到达的顾客在决策之前能够考虑上一个休假期出现的中途退出行为。而且，显然，$\lambda_{si}^{*}<\lambda_{si}^{e}$ 且 $\lambda_{mi}^{*}<\lambda_{mi}^{e}(i=0,1)$。

对于完全不可视排队系统，图 5-34 和图 5-35 表明 $\lambda_{m}^{e}<\lambda_{s}^{e}$，但是 $\lambda_{m}^{*}>\lambda_{s}^{*}$。因此，顾客和社会计划者对休假策略的选择具有不同的偏好。而且，$\lambda_{s}^{*}<\lambda_{s}^{e}$ 且 $\lambda_{m}^{*}<\lambda_{m}^{e}$ 依然成立。将几乎不可视和完全不可视排队系统中顾客的均衡止步行为进行比较，发现 $\lambda_{s1}^{e}<\lambda_{s}^{e}<\lambda_{s0}^{e}$ 且 $\lambda_{m1}^{e}<\lambda_{m}^{e}<\lambda_{m0}^{e}$。对于几乎不可视排队系统，在服务员处于休假期时，社会计划者希望顾客选择多重休假策略，而在忙期时，他希望顾客选择单重休假策略。

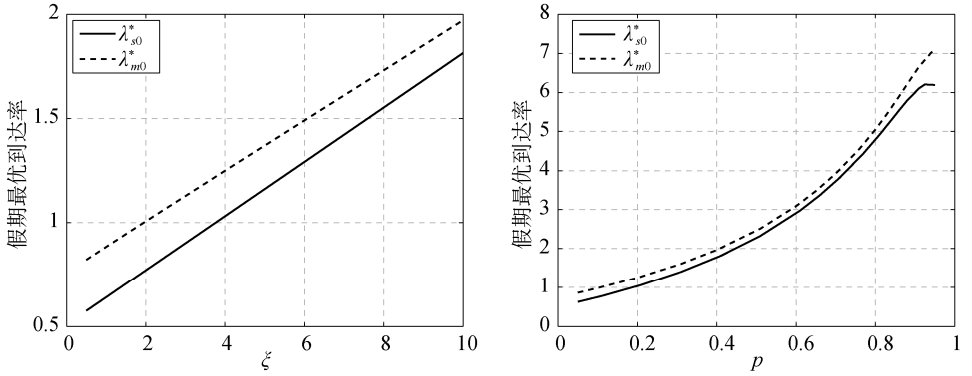

图 5-32　λ_{s0}^{*} 与 λ_{m0}^{*}：(1)$R=5$，$r=2$，$c=1$，$\mu=2$，$\gamma=0.5$，$p=0.2$；(2)$R=5$，$r=2$，$c=1$，$\mu=2$，$\gamma=0.5$，$\xi=4$

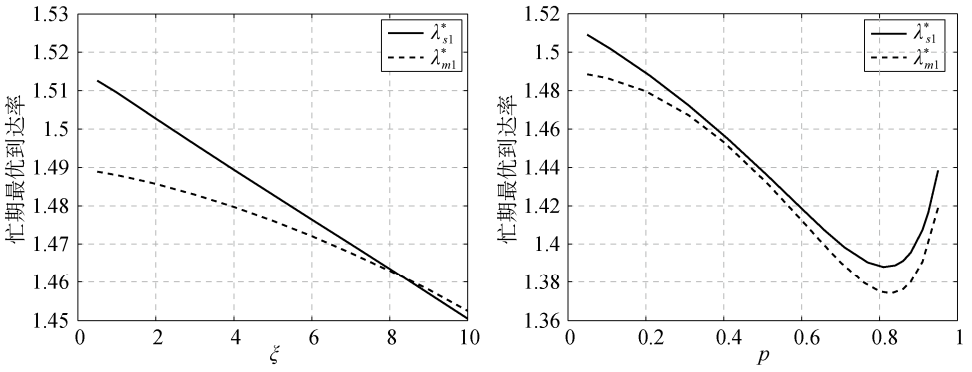

图 5-33　λ_{s1}^{*} 与 λ_{m1}^{*}：(1)$R=5$，$r=2$，$c=1$，$\mu=2$，$\gamma=0.5$，$p=0.2$；(2)$R=5$，$r=2$，$c=1$，$\mu=2$，$\gamma=0.5$，$\xi=4$

最后，比较最优社会福利。图 5-36 和图 5-37 都说明多重休假系统中的最优社会福利低于单重休假系统中的最优社会福利，尤其是在完全不可视排队系统中更是明显。而且，显然，$SW_{s}^{*fu}<SW_{s}^{*au}$ 且 $SW_{m}^{*fu}<SW_{m}^{*au}$，这是由信息量的多少决定的。

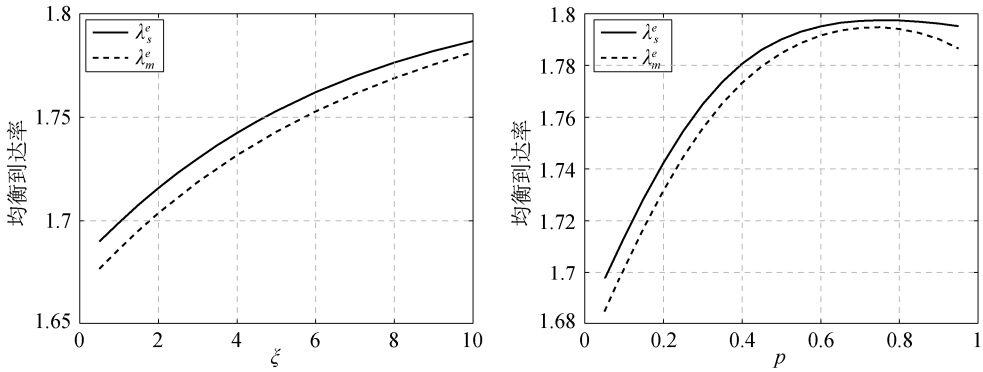

图 5-34　λ_s^e 与 λ_m^e : (1)$R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $p=0.2$;

(2)$R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $\xi=4$

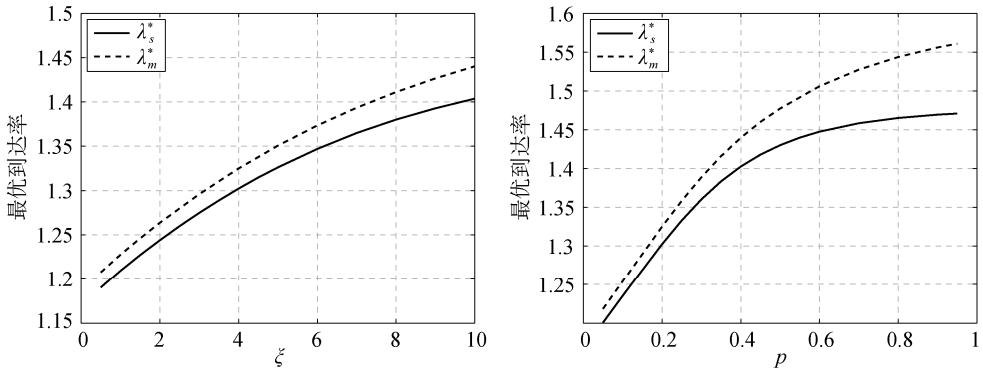

图 5-35　λ_s^* 与 λ_m^* : (1)$R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $p=0.2$;

(2) $R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $\xi=4$

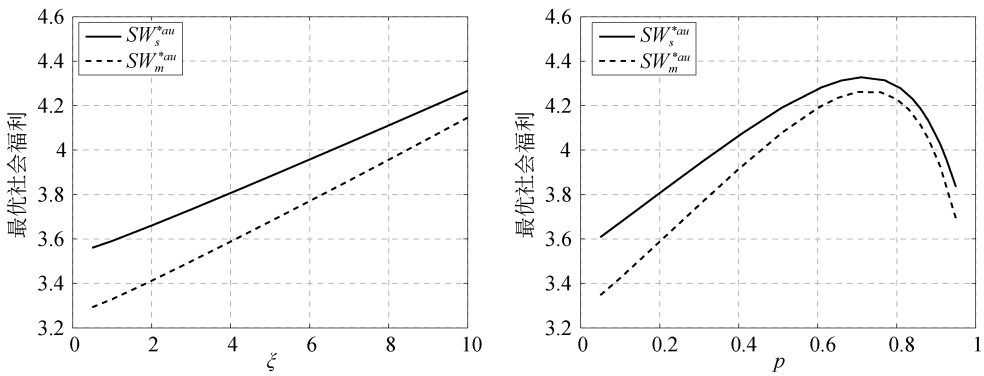

图 5-36　SW_s^{*au} 与 SW_m^{*au} : (1)$R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $p=0.2$;

(2) $R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $\xi=4$

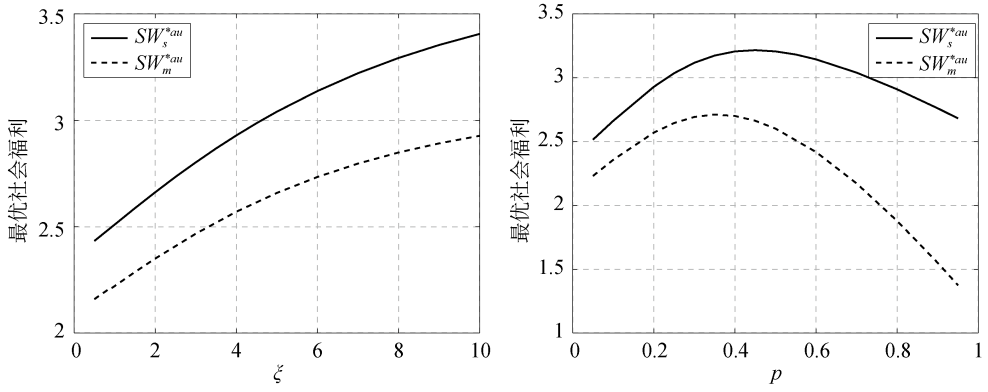

图 5-37　SW_s^{*fu} 与 SW_m^{*fu} : (1)$R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $p=0.2$;
(2)$R=5$, $r=2$, $c=1$, $\mu=2$, $\gamma=0.5$, $\xi=4$

5.4　相关文献评述

从经典随机服务系统的研究角度，假定在服务员休假期或慢速服务期由外因而导致顾客中途退出的研究工作，共分三类：独立退出行为（Independent Abandonment）、同时退出行为（Synchronized Abandonment）与顺序退出行为（Sequential Abandonment）。Altman 和 Yechiali[80]首次考虑了顾客在服务员休假期的独立退出行为，其假定每个顾客在休假期开始时开启各自的“计时器”，若倒计时完毕则退出系统，其倒计时时长为外生量，其分布由外因事先给定。更多关于独立退出行为的文献可参阅 Laxmi 和 Jyothsna[81]以及岳德权等[82]等的作品。除了独立退出行为之外，Adan 等[60]、Kapodistria[83]考虑了一些马尔科夫与非马尔科夫休假系统中顾客的同时退出行为，其假定外生的退出机遇（Abandonment Opportunity）的到达过程服从某种点过程（如 Poisson 过程），每位顾客独立且同时做出是否退出队列的决策。此时，一次性退出系统的顾客数服从二项分布。接下来，Dimou 等[84]考虑了休假系统中顾客的顺序退出行为，每位顾客独立且按一定顺序顺次做出是否退出队列的决策，一旦有顾客决定继续等待，退出过程结束。此时，一次性退出系统的顾客数服从几何分布。

上述文献均假定服务员可靠，而在现实的服务系统之中，服务员可能会因自身或外力作用而失去服务能力。Yechiali[85]首先考虑了 M/M/c 清场系统中顾客在服务员修理期的独立退出行为，得到了最终接受服务的顾客的平均逗留时间、接受服务顾客的比例、清场率、退出率等服务质量指标。然而，Economou 和

Kapodistria[86]则考虑了类似模型中顾客的同时退出行为。作为补充与扩展, Dimou 与 Economou[87]研究了 M/M/1 清场系统中顾客在服务员修理期的顺序退出行为。

相比之下，从经济学的视角来分析具有顾客退出行为的排队系统中顾客的止步行为的研究工作还是很少的。Boudali 和 Economou[88]假定在服务员修理期间禁止顾客进入系统, 得到了可视与不可视情形下顾客的均衡和最优止步策略。而后, Boudali 和 Economou[89]又去除了该假定，研究了类似的模型，得到了顾客的均衡止步策略。

从以上文献中看到，在研究顾客的中途退出行为时，都是提前为顾客的等待时间阈值设定一个外生的概率分布，或假定诱使顾客退出的外来机遇（包括服务员失效等）按某种规律出现或发生。相反，从内生量的角度考虑顾客的中途退出行为的研究成果较少。最初，Hassin 和 Haviv[90]研究了一个较简单的带有同类顾客止步和中途退出行为的不可视 Markov 单服务员排队系统，得到了顾客的均衡和社会最优止步和中途退出策略组合。而后，Haviv 和 Ritov[91]又研究了只带有同类顾客中途退出行为的不可视 Markov 多服务员排队系统。与此同时，除了同类顾客, Mandelbaum 和 Shimkin 等[92–94]考虑了更一般的带有异类顾客中途退出行为的不可视 Markov 多服务员排队系统，均得到了均衡状态下顾客的退出阈值策略。

第 6 章　具有部分服务信息的排队系统

对于刚刚到达一个服务系统的顾客来说，只知道部分服务信息的情况非常普遍，也就是对服务系统的某些指标或参数，如服务时间的具体分布，并不明确知道。因此，顾客需要根据已有的不完全服务信息，来对服务时间的分布进行估计和猜测，以此做出是否排队等候或止步离开的决定。至于顾客将采取哪种标准或原理来对此进行估计，我们在本章假设最大熵原理。

本章考虑了两个具有不完全服务信息和不同类型顾客的队长不可视排队系统，顾客类型分别为风险中立型和风险规避型。基于最大熵原理，主要分析了顾客的均衡止步策略和社会最优止步策略，以及服务员在均衡状态下对于不同类型的部分服务信息所给出的垄断价格和获得的最大利润。

本章内容为作者研究成果[95]的进一步完善和补充。

6.1　最大熵原理

在不完全信息的前提下，最大熵原理本质上是一种偏差最小的统计方法。根据最大熵原理，在给定的不完全信息下，对一个随机事件或随机变量的分布函数进行估计，所得的此分布函数即为最大熵分布。最大熵原理说明决策者应当选择留下最大剩余不确定性（即最大熵）的那个概率分布，这样就没有引入任何附加假设或偏差。因此，最大熵分布是所有符合不完全信息的概率分布中最无偏最客观的分布，它没有掺杂进任何已知信息以外的信息。

在本章考虑的排队系统中，顾客只被告知服务时间分布的部分信息。因此，顾客首先需要对服务时间分布函数的形式进行估计。根据部分服务信息和最大熵原理得到最大熵分布后，计算平均剩余效用，以指导其做出是否加入队列的决策。

6.2　模型描述

为以下分析方便起见，结合各种类型的部分服务信息，首先给出一些关于最大熵分布的预备知识[96]。

引理 6.2.1　如果给定部分服务信息是服务时间的取值范围$[s_1, s_2]$，则服务时间的最大熵分布是区间$[s_1, s_2]$上的均匀分布。

引理 6.2.2　如果给定部分服务信息是服务时间的均值$\bar{s}$，则服务时间的最大熵分布是参数为$1/\bar{s}$的指数分布。

引理 6.2.3　如果给定部分服务信息是服务时间的取值范围$[s_1, s_2]$和均值$\bar{s}$，则服务时间的最大熵分布是参数为λ_1的截尾指数分布，也就是最大熵分布密度函数为

$$f(x)=\frac{\lambda_1 e^{-\lambda_1 x}}{e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}}, s_1 \le x \le s_2 \tag{6-1}$$

其中，λ_1满足方程

$$e^{-\lambda_1(s_1-s_2)}=\frac{\lambda_1(\bar{s}-s_2)-1}{\lambda_1(\bar{s}-s_1)-1} \tag{6-2}$$

引理 6.2.4　如果给定部分服务信息是服务时间的均值$\bar{s}$和方差σ_s^2，则:

（1）若服务时间的变差系数$\sigma_s/\bar{s}$满足$\sigma_s/\bar{s}<1$，则服务时间的最大熵分布是截尾正态分布，其密度函数为

$$f(x)=\frac{1}{\Phi\left(\frac{\mu}{\sigma}\right)}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{6-3}$$

其中，μ和σ是以下方程的解：

$$\begin{cases}\mu+\sigma\dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}=\bar{s}\\ \sigma^2\left(1-\dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\left(1-\dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\right)\right)=\sigma_s^2\end{cases} \tag{6-4}$$

（2）若服务时间的变差系数$\sigma_s/\bar{s}$满足$\sigma_s/\bar{s}=1$，则服务时间的最大熵分布是参数为$1/\bar{s}$ (或$1/\sigma_s$)的指数分布；

（3）若服务时间的变差系数 $\sigma_s / \overline{s}$ 满足 $\sigma_s / \overline{s} > 1$，则服务时间的最大熵分布不存在。

假设潜在顾客的到达过程服从参数为 Λ 的泊松分布，且顾客类型相同。对于刚刚到达的顾客，有两种纯策略可供选择：加入队列等待服务或立即止步离开。在队长不可视的排队系统中，顾客采取进入概率为 $q(0<q<1)$的混合止步策略。因此，实际到达率 λ 为 $\lambda=\Lambda q$，且均衡状态下有 $\lambda_e=\Lambda q_e$。

基于经典 M/G/1 排队的关于稳态逗留时间的结果[7]，有其一阶矩 $E[W]$为

$$E[W]=\overline{s}_1+\frac{\lambda\overline{s}_2}{2(1-\rho)} \tag{6-5}$$

和二阶矩 $E[W^2]$为

$$E[W^2]=\overline{s}_2+\frac{\lambda^2\overline{s}_2^{\,2}}{2(1-\rho)^2}+\frac{\lambda\overline{s}_1\overline{s}_2}{1-\rho}+\frac{\lambda\overline{s}_3}{3(1-\rho)} \tag{6-6}$$

其中 $\rho=\lambda\overline{s}_1, \overline{s}_1, \overline{s}_2, \overline{s}_3$ 分别是服务时间的前三阶矩。

假设顾客不允许中途退出。服务完成所得的效用为 R，单位等待费用为 c，等待费用函数为 $c(W)$。风险中立顾客采用线性费用函数，即 $c(W)=cE[W]$，而风险规避顾客采用二次费用函数，即 $c(W)=cE[W^2]$。因此，风险中立顾客的剩余效用函数为 $U_R=R-cE[W]$，风险规避顾客的剩余效用函数为 $U_R=R-cE[W^2]$。

6.3 风险中立顾客

首先考虑具有风险中立顾客和不完全服务信息的队长不可视排队系统。基于最大熵原理，首先考虑在各种部分服务信息下顾客的均衡止步行为，其部分服务信息包括：取值范围、均值、取值范围和均值、均值和方差。这里用符号 R-信息、M-信息、MR-信息和 MV-信息分别表示各类信息，并用上标 R、M、MR、MV 分别对其进行标记。然后给定几种常见的服务时间分布函数：均匀分布、指数分布、截尾指数分布，考虑社会最优止步行为并把两者结果进行对比。最后讨论均衡状态下，服务员在各种部分信息下的最优价格和最大利润，并进行比较。

6.3.1 纳什均衡

6.3.1.1 R-信息

如果到达顾客只被告知服务时间的取值范围$[s_1, s_2]$，$0 < s_1 < s_2$，根据

引理 6.2.1，顾客估计服务时间为区间$[s_1, s_2]$上的均匀分布。因此，根据式（6-5），顾客的平均逗留时间 $E[W(q)]$为

$$E\left[W\left(q\right)\right]=\frac{s_1+s_2}{2}+\frac{\Lambda q\left(s_1{}^2+s_1s_2+s_2{}^2\right)}{3\left(2-\Lambda q\left(s_1+s_2\right)\right)} \tag{6-7}$$

平均剩余效用为 $U_R(q) = R-cE[W(q)]$。因为在均衡状态下，队长不可视排队系统中顾客的平均剩余效用为零，所以解方程 $U_R(q) = 0$ 得到顾客的均衡进入概率 q_e 为

$$q_e^R=\frac{6\left(2R-c\left(s_1+s_2\right)\right)}{\Lambda\left(6R\left(s_1+s_2\right)-c\left(s_1{}^2+4s_1s_2+s_2{}^2\right)\right)} \tag{6-8}$$

6.3.1.2　M-信息

如果到达顾客只被告知服务时间的均值$\overline{s}$，根据引理 6.2.2，顾客估计服务时间是参数为 $1/\overline{s}$ 的指数分布。因此，顾客的平均逗留时间为

$$E[W(q)=\overline{s}+\frac{\Lambda q\overline{s}^2}{1-\Lambda q\overline{s}} \tag{6-9}$$

平均剩余效用为$U_R(q)=R-cE[W(q)]$。解$U_R(q)=0$，得到顾客的均衡进入概率为

$$q_e^M=\frac{R-c\overline{s}}{R\Lambda\overline{s}} \tag{6-10}$$

6.3.1.3　MR-信息

如果到达顾客既被告知服务时间的取值范围$[s_1, s_2]$，又被告知均值$\overline{s}$，根据引理 6.2.3，顾客估计服务时间是参数为 λ_1 的截尾指数分布，其中 λ_1 满足式（6-2）。因此，顾客的平均逗留时间为

$$E[W(q)]=\overline{s}+\frac{\Lambda q}{2\left(1-\Lambda q\overline{s}\right)}\left(\frac{2}{\lambda_1{}^2}+\frac{s_1^2e^{-\lambda_1s_1}-s_2^2e^{-\lambda_1s_2}}{e^{-\lambda_1s_1}-e^{-\lambda_1s_2}}+\frac{2\left(s_1e^{-\lambda_1s_1}-s_2e^{-\lambda_1s_2}\right)}{\lambda_1\left(e^{-\lambda_1s_1}-e^{-\lambda_1s_2}\right)}\right) \tag{6-11}$$

平均剩余效用为$U_R\left(q\right)=R-cE[W\left(q\right)]$。解$U_R\left(q\right)=0$，得到$q_e^{MR}$的分子为

$$q_e^N=2\lambda_1{}^2\left(e^{-\lambda_1s_1}-e^{-\lambda_1s_2}\right)\left(R-c\overline{s}\right) \tag{6-12}$$

分母为

$$\begin{aligned}q_e^D=\Lambda\Big(&2\left(e^{-\lambda_1s_1}-e^{-\lambda_1s_2}\right)\left(\left(R-c\overline{s}\right)\lambda_1{}^2s+c\right)+c\lambda_1{}^2\left(s_1^2e^{-\lambda_1s_1}-s_2^2e^{-\lambda_1s_2}\right)\\&+2c\lambda_1\left(s_1e^{-\lambda_1s_1}-s_2e^{-\lambda_1s_2}\right)\Big).\end{aligned} \tag{6-13}$$

6.3.1.4　MV-信息

如果到达顾客既被告知服务时间的均值为$\overline{s}$，又被告知方差为σ_s^2，则顾客的

平均逗留时间为

$$E[W(q)]=\overline{s}+\frac{\Lambda q\left(\overline{s}^2+\sigma_s^2\right)}{2\left(1-\Lambda q\overline{s}\right)} \tag{6-14}$$

平均剩余效用为$U_R(q)=R-cE[W(q)]$。解$U_R(q)=0$，得到顾客的均衡进入概率为

$$q_e^{MV}=\frac{2\left(R-c\overline{s}\right)}{\Lambda\left(2R\overline{s}+c\left(\sigma_s^2-\overline{s}^2\right)\right)} \tag{6-15}$$

因此，关于服务时间的均值和方差的部分信息对风险中立顾客做出决定来说已经足够，顾客无须再对具体的服务时间分布进行估计。

6.3.2 社会最优

对于具有均值$\overline{s}$和方差σ_s^2的任意服务时间分布，下面来分析不同情况下顾客的社会最优混合止步策略。首先假设$R\geq c\overline{s}$，单位社会福利$SW(q)$为

$$SW(q)=\Lambda q\left(R-c\left(\overline{s}+\frac{\Lambda q\left(\overline{s}^2+\sigma_s^2\right)}{2\left(1-\Lambda q\overline{s}\right)}\right)\right) \tag{6-16}$$

则社会最优目标函数为

$$\max_{0\leq q\leq 1}\Lambda q\left(R-c\left(\overline{s}+\frac{\Lambda q(\overline{s}^2+\sigma_s^2}{2\left(1-\Lambda q\overline{s}\right)}\right)\right) \tag{6-17}$$

基于一阶最优条件$SW'(q)=0$，得到社会最优进入概率q^*为

$$q^*=\frac{2R\overline{s}+c\left(\sigma_s^2-\overline{s}^2\right)-\sqrt{c\left(\sigma_s^2+\overline{s}^2\right)\left(2R\overline{s}+c\left(\sigma_s^2-\overline{s}^2\right)\right)}}{\overline{s}\Lambda\left(2R\overline{s}+c\left(\sigma_S^2-\overline{s}^2\right)\right)} \tag{6-18}$$

根据假设$R\geq c\overline{s}$，q^*非负。因此，如果$q^*<1$，$\lambda^*=\Lambda q$；否则，$\lambda^*=\Lambda$，$q^*=1$。

接下来，将考虑几种常见的服务时间分布函数，导出顾客的社会最优混合止步策略，并与顾客在各种部分信息下的均衡止步策略进行比较。

6.3.2.1 均匀分布

如果实际的服务时间分布函数为区间$[s_1, s_2]$上的均匀分布，则假设$R\geq cs_2$。单位社会福利函数为

$$SW(q)=\Lambda q\left(R-c\left(\frac{s_1+s_2}{2}+\frac{\Lambda q\left(s_1^2+s_1s_2+s_2^2\right)}{3\left(2-\Lambda q\left(s_1+s_2\right)\right)}\right)\right) \tag{6-19}$$

则社会最优目标函数为

$$\max_{0\le q\le 1}\quad \Lambda q\left(R-c\left(\frac{s_1+s_2}{2}+\frac{\Lambda q\left(s_1^2+s_1s_2+s_2^2\right)}{3\left(2-\Lambda q\left(s_1+s_2\right)\right)}\right)\right) \tag{6-20}$$

基于一阶最优条件 $SW'(q)=0$，得到社会最优进入概率 q^* 的分子为

$$\begin{aligned} q_N^* = {} & 2\left(6R\left(s_1+s_2\right)-c\left(s_1^2+4s_1s_2+s_2^2\right)\right) \\ & -2\sqrt{2c\left(s_1^2+s_1s_2+s_2^2\right)\left(6R\left(s_1+s_2\right)-c\left(s_1^2+4s_1s_2+s_2^2\right)\right)} \end{aligned} \tag{6-21}$$

分母为

$$q_D^* = \left(s_1+s_2\right)\left(6R\left(s_1+s_2\right)-c\left(s_1^2+4s_1s_2+s_2^2\right)\right)\Lambda \tag{6-22}$$

根据假设 $R>c\left(s_1+s_2\right)/2$，$q^*$ 非负。因此，如果 $q^*<1$，则 $\lambda^*=\Lambda q^*$；否则，$\lambda^*=\Lambda$，$q^*=1$。

接下来，比较三种部分服务信息下的均衡和社会最优止步策略。如果顾客只被告知服务时间的取值范围$[s_1, s_2]$，则最大熵分布与实际分布吻合，也就是顾客的估计是准确的，这样均衡进入概率 q_e 由式（6-8）给出：如果顾客被告知服务时间的均值 $\bar{s}$ 和方差 σ_s^2，即 $\bar{s}=\left(s_1+s_2\right)/2$，$\sigma_s^2=\left(s_2-s_1\right)^2/12$，则顾客的决策独立于服务时间的真实分布，所以均衡进入概率与只告知服务时间取值范围时相等。

如果顾客只被告知服务时间的均值 $\bar{s}$，即 $\bar{s}=(s_1+s_2)/2$，则最大熵分布是参数为 $1/\bar{s}$ 的指数分布，与实际服务时间分布不符。此时，均衡进入概率为

$$q_e=\frac{R-c\bar{s}}{R\Lambda\bar{s}}=\frac{2R-c\left(s_1+s_2\right)}{R\Lambda\left(s_1+s_2\right)} \tag{6-23}$$

如果顾客被告知服务时间的取值范围$[s_1, s_2]$和均值 $\bar{s}=\left(s_1+s_2\right)/2$，则最大熵分布是区间$[s_1, s_2]$上参数为 λ_1 的截尾指数分布，其中 λ_1 满足

$$e^{-\lambda_1\left(s_1-s_2\right)}=\frac{\lambda_1\left(s_1-s_2\right)-2}{\lambda_1\left(s_2-s_1\right)-2}$$

解得 $\lambda_1=0$。根据洛必达法则（L'Hospital Rule），把 $\lambda_1=0$ 代入式（6-12）和式（6-13）中，得到均衡进入概率

$$q_e=\frac{6\left(R-c\bar{s}\right)}{\Lambda\left(6\bar{s}\left(R-c\bar{s}\right)+c\left(s_1^2+s_1s_2+s_2^2\right)\right.}=\frac{6\left(2R-c\left(s_1+s_2\right)\right)}{\Lambda\left(6R\left(s_1+s_2\right)-c\left(s_1^2+4s_1s_2+s_2^2\right)\right)} \tag{6-24}$$

与只被告知服务时间取值范围$[s_1, s_2]$时的均衡进入概率相等。另一方面，通过把 $\lambda_1=0$ 代入式（6-1），发现服务时间的密度函数变为

$$f(x)=\frac{\lambda_1e^{-\lambda_1x}}{e^{-\lambda_1s_1}-e^{-\lambda_1s_2}}=\frac{e^{-\lambda_1x}-\lambda_1xe^{-\lambda_1x}}{s_2e^{-\lambda_1s_2}-s_1e^{-\lambda_1s_1}}=\frac{1}{s_2-s_1},s_1\le x\le s_2 \tag{6-25}$$

说明区间$[s_1, s_2]$上的参数为$\lambda_1 = 0$的截尾指数分布实际上是均匀分布。这个结果说明当实际服务时间分布为均匀分布时，顾客被告知服务时间均值和取值范围与顾客只被告知取值范围效果相同，服务时间的均值这个附加信息对顾客的决策不起作用。

当实际的服务时间分布为均匀分布时，图 6-1 比较了顾客的均衡进入概率q_e和社会最优进入概率q^*。观察到所有的q_e都大于q^*。与式（6-25）一致，部分信息取值范围、取值范围和均值、均值和方差下的均衡进入概率都是相等的，且大于部分信息为均值下的均衡进入概率。

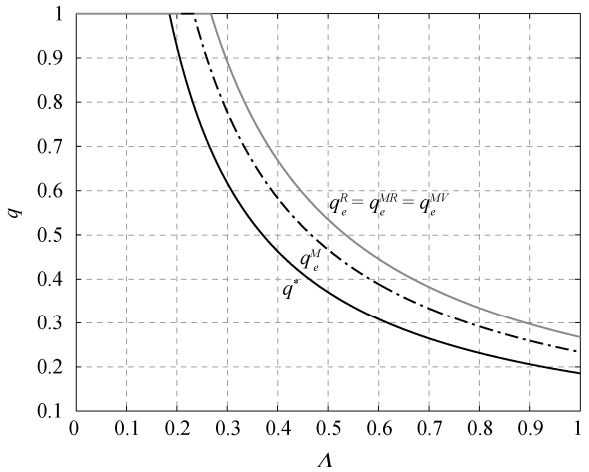

图 6-1　服务时间为均匀分布时的q_e和q^*

6.3.2.2　指数分布

如果实际的服务时间分布函数是参数为$1/\bar{s}$的指数分布，则假设$R \geq c\bar{s}$。单位社会福利函数为

$$SW(q) = \Lambda q\left(R - c\left(\bar{s} + \frac{\Lambda q\bar{s}^2}{1-\Lambda q\bar{s}}\right)\right) \tag{6-26}$$

则社会最优目标函数为

$$\max_{0\leq q\leq 1} \quad \Lambda q\left(R - c\left(\bar{s} + \frac{\Lambda q\bar{s}^2}{1-\Lambda q\bar{s}}\right)\right) \tag{6-27}$$

基于一阶最优条件$SW'(q) = 0$，得到社会最优进入概率q^*为

$$q^* = \frac{R - \sqrt{Rcs}}{R\Lambda s} \tag{6-28}$$

基于假设$R \geq c\bar{s}$，q^*非负。因此，如果$q^* < 1$，$\lambda^* = \Lambda q^*$；否则，$\lambda^* = \Lambda$，$q^* = 1$。

如果顾客只被告知服务时间的均值 $\bar{s}$，则最大熵分布与实际分布吻合。这样均衡进入概率 q_e 由式（6-10）给出；如果顾客被告知服务时间的均值 $\bar{s}$ 和方差 $\sigma_s^2=\bar{s}^2$，则顾客的决策独立于服务时间的真实分布，所以均衡进入概率与只告知服务时间均值时相等。

当实际的服务时间分布为指数分布时，图 6-2 比较了顾客的均衡进入概率 q_e 和社会最优进入概率 q^*。显然，部分信息均值、均值和方差下的均衡进入概率大于社会最优进入概率。

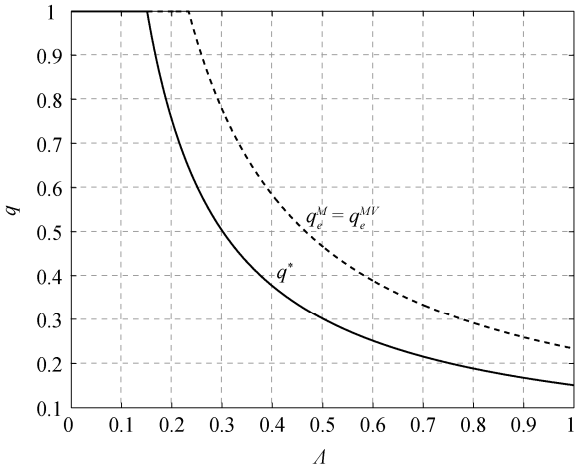

图 6-2　服务时间为指数分布时的 q_e 和 q^*

6.3.2.3　截尾指数分布

如果实际的服务时间分布函数是区间 $[s_1,s_2]$ 上参数为 λ_1 的截尾指数分布，则假设 $R\geq c\overline{s_1}$。单位社会福利函数为

$$SW(q)=\Lambda q\left(R-c\left(\overline{s_1}+\frac{\Lambda q\overline{s_2}}{2\left(1-\Lambda q\overline{s_1}\right)}\right)\right) \tag{6-29}$$

其中

$$\begin{cases}\overline{s_1}=\dfrac{1}{\lambda_1}+\dfrac{s_1e^{-\lambda_1 s_1}-s_2e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}}\\ \overline{s_2}=\dfrac{2}{\lambda_1^{\,2}}+\dfrac{s_1^2e^{-\lambda_1 s_1}-s_2^2e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}}+\dfrac{2\left(s_1e^{-\lambda_1 s_1}-s_2e^{-\lambda_1 s_2}\right)}{\lambda_1\left(e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}\right)}\end{cases} \tag{6-30}$$

则社会最优目标函数为

$$\max_{0\le q\le 1} \quad \Lambda q\left(R-c\left(\overline{s}_1+\frac{\Lambda q\overline{s}_2}{2\left(1-\Lambda q\overline{s}_1\right)}\right)\right) \tag{6-31}$$

基于一阶最优条件 $SW'(q)=0$，得到社会最优进入概率 q^* 为

$$q^*=\frac{2\overline{s}_1\left(R-c\overline{s}_1\right)+c\overline{s}_2-\sqrt{c\overline{s}_2\left(2\overline{s}_1\left(R-c\overline{s}_1\right)+c\overline{s}_2\right)}}{\overline{s}_1\Lambda\left(2\overline{s}_1\left(R-c\overline{s}_1\right)+c\overline{s}_2\right)} \tag{6-32}$$

其中 $\overline{s}_1$ 和 $\overline{s}_2$ 由式(6-30)给出。基于假设 $R\geq c\overline{s}_1$，q^* 非负。因此，如果 $q^*<1$，$\lambda^*=\Lambda q^*$；否则，$\lambda^*=\Lambda$，$q^*=1$。

如果顾客被告知服务时间的取值范围为$[s_1, s_2]$，则最大熵分布为均匀分布，与实际分布不符。因此，q_e 由式（6-8）给出。如果顾客只被告知服务时间的均值 $\overline{s}_1$，则最大熵分布为指数分布，与实际分布也不符。因此，q_e 为

$$q_e=\frac{R\lambda_1\left(e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}\right)-c\left(\left(e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}\right)+\lambda_1\left(s_1e^{-\lambda_1 s_1}-s_2e^{-\lambda_1 s_2}\right)\right)}{R\Lambda\left(\left(e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}\right)+\lambda_1\left(s_1e^{-\lambda_1 s_1}-s_2e^{-\lambda_1 s_2}\right)\right)} \tag{6-33}$$

如果顾客被告知服务时间的取值范围$[s_1,s_2]$和均值 $\overline{s}_1$，则最大熵分布与实际服务时间分布吻合，q_e 由式（6-12）和式（6-13）给出。如果顾客被告知服务时间的均值 $\overline{s}_1$ 和方差 $\sigma_s^2=\overline{s}_2-\overline{s}_1^2$，其中 $\overline{s}_1$、$\overline{s}_2$ 由式（6-30）给出，则顾客的决策独立于服务时间的真实分布，所以均衡进入概率与被告知服务时间均值和取值范围时相等。

当实际的服务时间分布为截尾指数分布时，图 6-3 和图 6-4 比较了顾客的均衡进入概率 q_e 和社会最优进入概率 q^*。当 $\lambda_1=1$ 时，图 6-3 中的均衡进入概率存在关系 $q_e^{MR}=q_e^{MV}>q_e^M>q_e^R$。然而，当 $\lambda_1=0.2$ 时，图 6-4 中的均衡进入概率存在关系 $q_e^{MR}=q_e^{MV}>q_e^R>q_e^M$。而且，在图 6-3 中，当顾客只被告知服务时间范围时，均衡进入概率甚至小于社会最优进入概率。

6.3.3 利润最大

现在考虑均衡状态下服务员在各种类型的部分服务信息下的利润和定价问题。假设服务价格为 p，则顾客在某种部分服务信息下的均衡进入概率满足

$$R-p=cE[W(q_e)] \tag{6-34}$$

基于式（6-5），解方程得

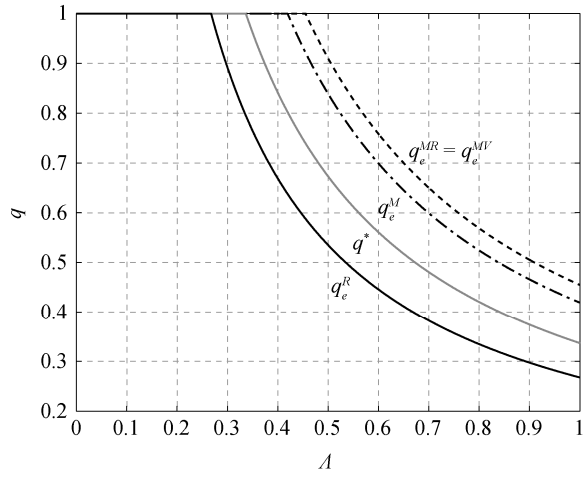

图 6-3　服务时间分布是参数为 λ_1=1 的截尾指数分布时的 q_e 和 q^*

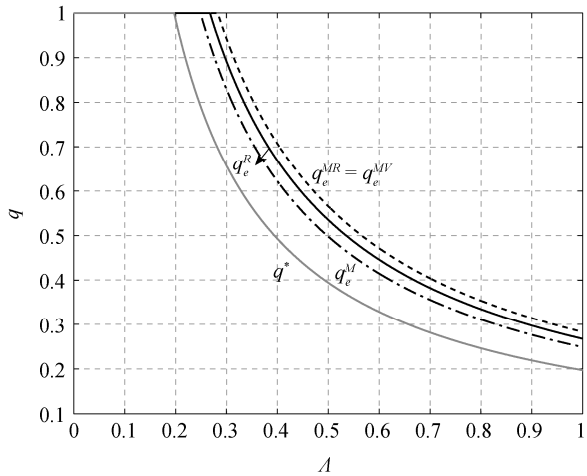

图 6-4　服务时间分布是参数为 λ_1 = 0.1 的截尾指数分布时的 q_e 和 q^*

$$q_e(p)=\frac{2\left((R-p)-c\overline{s}\right)}{\Lambda\left(2\overline{s}\left(R-p\right)+c\left(\sigma_s^2-\overline{s}^2\right)\right)} \tag{6-35}$$

因此，均衡状态下服务员的利润 $\Pi(p)$为

$$\prod(p)=\Lambda q_e(p)p=\frac{2p\left((R-p)-c\overline{s}\right)}{2\overline{s}\left(R-p\right)+c\left(\sigma_s^2-\overline{s}^2\right)} \tag{6-36}$$

解方程 $\prod'(p)=0$，得到垄断价格为

$$p^*=\frac{2R\overline{s}+c\left(\sigma_s^2-\overline{s}^2\right)-\sqrt{\left(\overline{s}^2+\sigma_s^2\right)\left(2R\overline{s}c+c^2\left(\sigma_s^2-\overline{s}^2\right)\right.}}{2\overline{s}} \tag{6-37}$$

相应的最大利润为 $\prod^* = \Lambda q_e\left(p^*\right)p^*$。根据最大熵原理，对于任意的服务时间分布函数，给定不同的部分服务信息，顾客将对服务时间的均值 $\bar{s}$ 和方差 σ_s^2 给出不同的估计结果。因此显然，部分服务信息的种类直接关系到服务员的利润和价格。

假设实际服务时间分布分别为均匀分布，指数分布和截尾指数分布，在各种类型的部分服务信息下，对均衡状态下服务员的利润和价格进行一下数值比较。

当实际的服务时间分布为均匀分布时，图 6-5 对部分信息取值范围（取值范围和均值、均值和方差）或均值下的服务员利润进行了比较。对应于 $q_e^R = q_e^{MR} = q_e^{MV}$，部分信息取值范围(取值范围和均值、均值和方差)下的利润存在关系 $\prod^R = \prod^{MR} = \prod^{MV}$。而且，对于任意给定的价格，关系 $\prod^M \le \prod^R = \prod^{MR} = \prod^{MV}$ 始终保持。如果部分信息为均值，则垄断价格为 $P^{M^*} = 5.5279$，最大利润为 $\prod^{M^*} = 1.5279$。如果部分信息为取值范围（取值范围和均值、均值和方差），则垄断价格为 $P^{R^*} = P^{MR^*} = P^{MV^*} = 5.9464$，最大利润为 $\Pi^{R^*} = \Pi^{MR^*} = 1.9464$。这说明顾客需根据不同的部分服务信息来设定不同的垄断价格以使利润最大化。而且，如果有可能，告知顾客服务时间的取值范围（取值范围和均值、均值和方差）要优于只告知其均值。

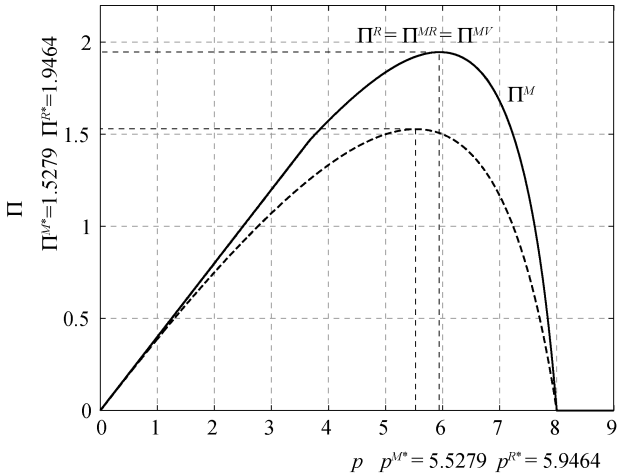

图 6-5　服务时间为均匀分布时的服务员利润

当实际的服务时间分布为指数分布时，图 6-6 展示了部分信息均值（均值和方差）下的服务员利润。如果部分信息为均值（均值和方差），则垄断价格为 $P^{M^*} = P^{MV^*} = 6.8377$，最大利润为 $\Pi^{M^*} = \Pi^{MV^*} = 4.6754$。

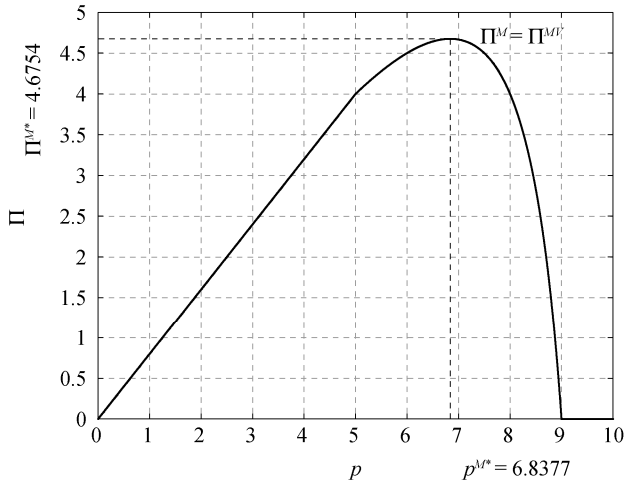

图 6-6　服务时间为指数分布时的服务员利润

当实际的服务时间分布为截尾指数分布时，图 6-7 和图 6-8 对部分信息取值范围、均值或取值范围和均值（均值和方差）下的服务员利润进行了比较。图 6-7 中，$\lambda_1=1$，而图 5-8 中，$\lambda_1=0.5$。也就是说，取值范围下的最大利润和均指下的最大利润之间的大小关系由截尾指数分布的参数 λ_1 决定。从图 6-7 观察到：如果部分信息为取值范围，则垄断价格为 P^{R^*}=5.9464，最大利润为 Π^{R^*}=1.9464；如果部分信息为均值，则垄断价格为 P^{M^*}=5.8927，最大利润为 Π^{M^*}= 2.0584；如果部分信息为取值范围和均值（均值和方差），则垄断价格为 $P^{MR^*}=P^{MV^*}$=6.3146，最大利润为 $\Pi^{MR^*}=\Pi^{MV^*}$=2.5586。然而在图 6-8 中，如果部分信息为均值，则 P^{M^*}=5.7151，Π^{M^*}=1.7789；如果部分信息为取值范围和均值（均值和方差），则 $P^{MR^*}=P^{MV^*}$=6.1337，$\Pi^{MR^*}=\Pi^{MV^*}$=2.235。因此，顾客被告知的部分服务信息越多越丰富，服务员的利润就会越大。

6.4　风险规避顾客

除了考虑风险中立顾客，下面继续考虑具有风险规避顾客的队长不可视排队系统。显然，由定义可知，无论在何种部分信息下，风险规避顾客的均衡进入概率不会大于风险中立顾客的均衡进入概率。平行于对具有风险中立顾客的排队系统的研究，考虑在各种部分服务信息下顾客的均衡止步行为，其部分服务信息包括：取值范围、均值、取值范围和均值、均值和方差、前三阶原点矩。这里用符

号 3M-信息表示前三阶原点矩信息，并用上标 3*M* 对其进行标记。然后给定几种常见的服务时间分布函数：均匀分布、指数分布、截尾指数分布、截尾正态分布，考虑社会最优止步行为并把两者结果进行对比。最后讨论均衡状态下，服务员在各种部分信息下的最优价格和最大利润，并进行比较。

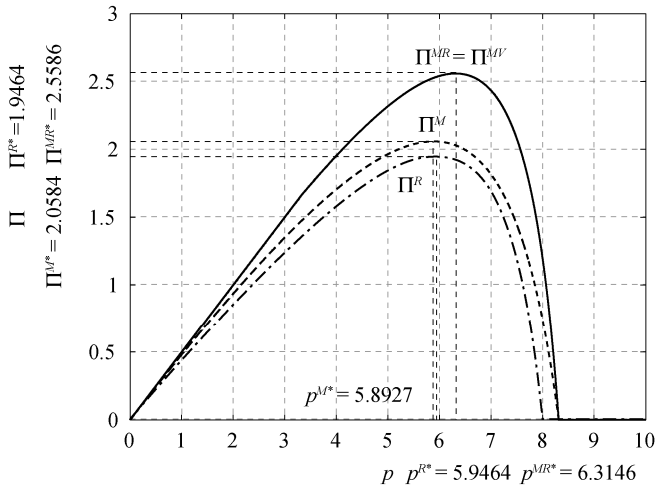

图 6-7　服务时间是参数为 $\lambda_1 = 1$ 的截尾指数分布时的服务员利润

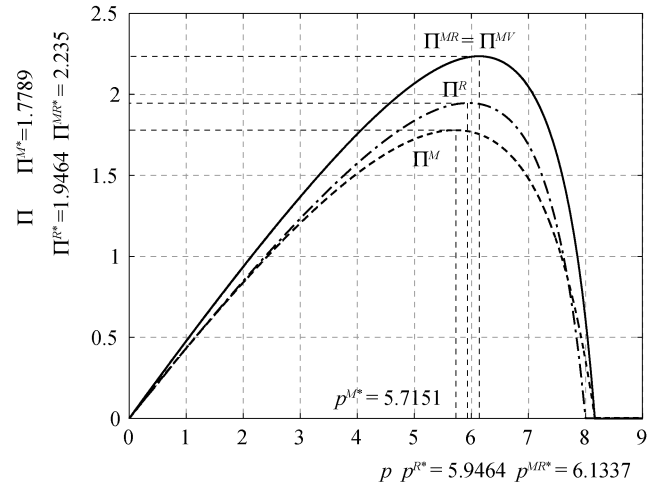

图 6-8　服务时间分布是参数为 $\lambda_1 = 0.5$ 的截尾指数分布时的服务员利润

根据引理 6.2.4，为保证最大熵分布的存在性，假设变差系数 $\sigma_s / \bar{s}$ 满足 $\sigma_s / \bar{s} < 1$，特别地，如果实际服务时间分布函数为指数分布，则变差系数满足 $\sigma_s / \bar{s} = 1$。

6.4.1　纳什均衡

6.4.1.1　R-信息

如果到达顾客只被告知服务时间的取值范围为$[s_1,s_2]$，$0<s_1<s_2$，则顾客逗留时间的二阶矩$E[W^2(q)]$为

$$E[W^2(q)]=\frac{2(\Lambda q)^2\left(s_1^2+s_1s_2+s_2^2\right)^2}{9\left(2-\Lambda q(s_1+s_2)\right)^2}+\frac{\Lambda q(s_1+s_2)\left(3s_1^2+2s_1s_2+3s_2^2\right)}{6\left(2-\Lambda q(s_1+s_2)\right)}+\frac{s_1^2+s_1s_2+s_2^2}{3} \tag{6-38}$$

平均剩余效用为$U_R(q)=R-cE[W^2(q)]$。解方程$U_R(q)=0$，得到顾客的均衡进入概率q_e^R的分子为

$$\begin{aligned}q_e^N=&3(s_1+s_2)\left(c(s_1+s_2)^2-12R\right)\\&+\sqrt{3c\left(96R\left((s_1+s_2)^2\left(s_1^2+s_2^2\right)+s_1^2s_2^2\right)-c(s_1+s_2)^6-4c\left(s_1^2+s_2^2\right)^3\right)}\end{aligned} \tag{6-39}$$

分母为

$$q_e^D=\left((s_1+s_2)^2\left(-18R+c\left(s_1^2+s_2^2\right)\right)+4cs_1^2s_2^2\right)\Lambda \tag{6-40}$$

6.4.1.2　M-信息

如果到达顾客只被告知服务时间的均值为$\bar{s}$，则顾客逗留时间的二阶矩$E[W^2(q)]$为

$$E[W^2(q)]=2\bar{s}^2+\frac{2\Lambda q\bar{s}^3(2-\Lambda q\bar{s})}{(1-\Lambda q\bar{s})^2} \tag{6-41}$$

平均剩余效用为$U_R(q)=R-cE[W^2(q)]$。解方程$U_R(q)=0$，得到顾客的均衡进入概率q_e^M为

$$q_e^M=\frac{R-\bar{s}\sqrt{2Rc}}{R\Lambda\bar{s}} \tag{6-42}$$

6.4.1.3　MR-信息

如果到达顾客被告知服务时间的取值范围为$[s_1,s_2]$，均值为$\bar{s}$，则顾客逗留时间的二阶矩$E[W^2(q)]$为

$$E[W^2(q)]=\bar{s}_2+\frac{(\Lambda q)^2\bar{s}_2^2}{2(1-\Lambda q\bar{s})^2}+\frac{\Lambda q\bar{s}\bar{s}_2}{1-\Lambda q\bar{s}}+\frac{\Lambda q\bar{s}_3}{3(1-\Lambda q\bar{s})} \tag{6-43}$$

其中

$$\begin{cases}\overline{s}_2=\dfrac{2}{\lambda_1^2}+\dfrac{s_1^2e^{-\lambda_1 s_1}-s_2^2e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}}+\dfrac{2(s_1e^{-\lambda_1 s_1}-s_2e^{-\lambda_1 s_2})}{\lambda_1(e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2})}\\ \overline{s}_3=\dfrac{6}{\lambda_1^3}+\dfrac{s_1^3e^{-\lambda_1 s_1}-s_2^3e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2}}+\dfrac{3(s_1^2e^{-\lambda_1 s_1}-s_2^2e^{-\lambda_1 s_2})}{\lambda_1(e^{-\lambda_1 s_1}-e^{-\lambda_1 s_2})}\end{cases} \tag{6-44}$$

且 λ_1 满足式（6-2）。平均剩余效用为 $U_R(q)=R-cE[W^2(q)]$。解方程 $U_R(q)=0$，得到顾客均衡进入概率 q_e^{MR} 为

$$q_e^{MR}=\frac{6R\overline{s}+c\overline{s}_3-3c\overline{s}\overline{s}_2-\sqrt{c^2(\overline{s}_3+3\overline{s}\overline{s}_2)^2+18c\overline{s}_2^2(R-c\overline{s}_2)}}{(-3c\overline{s}_2^2+6R\overline{s}^2+2c\overline{s}_3\overline{s})\Lambda} \tag{6-45}$$

其中 $\overline{s}_2$、$\overline{s}_3$ 由式（6-2）和式（6-44）给出。

6.4.1.4 MV-信息

如果到达顾客被告知服务时间的均值为 $\overline{s}$，方差为 σ_s^2，则根据引理 6.2.4，如果变差系数满足 $\sigma_s/\overline{s}<1$，最大熵分布为截尾正态分布。顾客逗留时间的二阶矩 $E[W^2(q)]$ 为

$$E[W^2(q)]=(\overline{s}^2+\sigma_s^2)+\frac{(\Lambda q)^2(\overline{s}^2+\sigma_s^2)^2}{2(1-\Lambda q\overline{s})^2}+\frac{\Lambda q(3\overline{s}(\overline{s}^2+\sigma_s^2)+\overline{s}_3)}{3(1-\Lambda q\overline{s})} \tag{6-46}$$

其中

$$\overline{s}_3=\int_0^\infty x^3\frac{1}{\Phi\left(\dfrac{\mu}{\sigma}\right)}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \tag{6-47}$$

且 μ 和 σ 是方程（6-4）的解。平均剩余效用为 $U_R(q)=R-cE[W^2(q)]$。解方程 $U_R(q)=0$，得到顾客均衡进入概率 q_e^{MV} 的分子为

$$\begin{aligned}q_e^N=&-(6\overline{s}R+c\overline{s}_3)+3c\overline{s}(\overline{s}^2+\sigma_s^2)\\&+\sqrt{-c(\overline{s}^2+\sigma_s^2)^2(-18R+9c(\overline{s}^2+2\sigma_s^2))-6c\overline{s}_3\overline{s}(\overline{s}^2+\sigma_s^2)-c\overline{s}_3^2}\end{aligned} \tag{6-48}$$

分母为

$$q_e^D=(-6R\overline{s}^2+3c\overline{s}^4+6c\sigma_s^2\overline{s}^2+3c(\sigma_s^2)^2-2c\overline{s}_3\overline{s})\Lambda \tag{6-49}$$

其中 $\overline{s}_3$ 由式（6-4）和式（6-47）给出。另一方面，如果变差系数满足 $\sigma_s/\overline{s}=1$，顾客估计服务时间分布服从参数为 $1/\overline{s}$（或 $1/\sigma_s$）的指数分布，q_e 由式（6-42）给出。

6.4.1.5 3M-信息

如果到达顾客被告知服务时间的前三阶原点矩为 $\{\overline{s}_1,\overline{s}_2,\overline{s}_3\}$，则顾客逗留时间的二阶矩 $E[W^2(q)]$ 为

$$E[W^2(q)]=\frac{(\Lambda q)^2\overline{s}_2^2}{2(1-\Lambda q\overline{s}_1)^2}+\frac{\Lambda q(3\overline{s}_1\overline{s}_2+\overline{s}_3)}{3(1-\Lambda q\overline{s}_1)}+\overline{s}_2 \tag{6-50}$$

平均剩余效用为$U_R(q)=R-cE[W^2(q)]$。解方程$U_R(q)=0$，得到顾客均衡进入概率q_e^{3M} 为

$$q_e^{3M}=\frac{6R\overline{s}_1-3c\overline{s}_1\overline{s}_2+c\overline{s}_3-\sqrt{c\left(3c\overline{s}_1\overline{s}_2\left(3\overline{s}_1\overline{s}_2+2\overline{s}_3\right)+18\overline{s}_2^2\left(R-c\overline{s}_2\right)+c\overline{s}_3^2\right)}}{\left(6R\overline{s}_1^2+2c\overline{s}_3\overline{s}_1-3c\overline{s}_2^2\right)\Lambda} \tag{6-51}$$

因此，关于服务时间的前三阶矩的部分信息对风险规避顾客做决策来说已经足够，顾客无须再对具体的服务时间分布进行估计。

6.4.2　社会最优

一般情况下，对于前三阶原点矩为$\{\overline{s}_1,\overline{s}_2,\overline{s}_3\}$的任意服务时间分布函数，根据式（6-6），单位社会福利函数为

$$SW(q)=\Lambda q\left(R-c\left(\frac{(\Lambda q)^2\overline{s}_2^2}{2(1-\Lambda q\overline{s}_1)^2}+\frac{\Lambda q(3\overline{s}_1\overline{s}_2+\overline{s}_3)}{3(1-\Lambda q\overline{s}_1)}+\overline{s}_2\right)\right) \tag{6-52}$$

类似于风险中立顾客的讨论过程，在此也考虑几种服务时间分布函数，导出风险规避顾客的社会最优混合止步策略。

6.4.2.1　均匀分布

如果实际的服务时间分布函数为区间$[s_1,s_2]$上的均匀分布，则单位社会福利函数为

$$\begin{aligned}SW(q)=\Lambda q\Bigg(R-c\Bigg(&\frac{2(\Lambda q)^2\left(s_1^2+s_1s_2+s_2^2\right)^2}{9\left(2-\Lambda q\left(s_1+s_2\right)\right)^2}\\&+\frac{\Lambda q\left(s_1+s_2\right)\left(3s_1^2+2s_1s_2+3s_2^2\right)}{6\left(2-\Lambda q\left(s_1+s_2\right)\right)}+\frac{s_1^2+s_1s_2+s_2^2}{3}\Bigg)\Bigg)\end{aligned} \tag{6-53}$$

则社会最优目标函数为$\max\{SW(q),0\le q\le 1\}$，基于一阶最优条件$SW'(q)=0$，可得到社会最优进入概率$q^*$。鉴于$q^*$表达式的复杂性，此后将省略其书写。

如果顾客只被告知服务时间的范围$[s_1,s_2]$，则最大熵分布与实际分布吻合，q_e由式（6-39）和式（6-40）给出。如果顾客被告知服务时间的前三阶原点矩$\overline{s}_1,\overline{s}_2,\overline{s}_3$，即

$$\overline{s}_1=\frac{s_1+s_2}{2},\overline{s}_2=\frac{s_1^2+s_1s_2+s_2^2}{3},\overline{s}_3=\frac{s_2^4-s_1^4}{4(s_2-s_1)} \tag{6-54}$$

则顾客的决策独立于服务时间的真实分布，所以均衡进入概率与只告知服务时间取值范围时相等。

如果顾客只被告知服务时间的均值 $\bar{s}$，即 $\bar{s}=(s_1+s_2)/2$，则最大熵分布与实际服务时间分布不符。此时，均衡进入概率为

$$q_e=\frac{R-s\sqrt{2Rc}}{R\Lambda s}=\frac{2R-(s_1+s_2)\sqrt{2Rc}}{R\Lambda(s_1+s_2)} \tag{6-55}$$

如果顾客被告知服务时间的取值范围 $[s_1,s_2]$ 和均值 $\bar{s}=(s_1+s_2)/2$，则最大熵分布是区间 $[s_1,s_2]$ 上参数为 $\lambda_1=0$ 的截尾指数分布。根据式（6-25），均衡进入概率与部分服务信息为取值范围时的一致，且由式（6-39）和式（6-40）给出。

如果顾客被告知服务时间的均值 $\bar{s}$ 和方差 σ_s^2，则最大熵分布是均值为 $\bar{s}=(s_1+s_2)/2$，方差为 $\sigma_s^2=(s_2-s_1)^2/12$ 的截尾正态分布。因此，均衡进入概率分子为

$$\begin{aligned}q_e^N=&\left(-18R+2c(s_1+s_2)^2\right)(s_1+s_2)+c\left(s_1^3+s_2^3-6\bar{s}_3\right)\\&+\left(36c\left((2R-cs_1s_2)(s_1+s_2)^2\left(s_1^2+s_2^2\right)+c\bar{s}_3\left(\bar{s}_3+(s_1+s_2)\left(s_1^2+s_1s_2+s_2^2\right)\right)\right)\right.\\&\left.+6cs_1^2s_2^2\left(12R-cs_1s_2\right)\right)^{\frac{1}{2}}\end{aligned}$$

分母为

$$q_e^D=\left((s_1+s_2)^2\left(-9R+2c\left(s_1^2+s_2^2\right)\right)-6c\bar{s}_3(s_1+s_2)+2cs_1^2s_2^2\right)\Lambda$$

其中

$$\bar{s}_3=\int_0^{\infty}x^3\frac{1}{\Phi\left(\frac{\mu}{\sigma}\right)}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x$$

且 μ 和 σ 是方程组

$$\begin{cases}\mu+\sigma\dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}=\dfrac{s_1+s_2}{2}\\[2ex]\sigma^2\left(1-\dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\left(1-\dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\right)\right)=\dfrac{(s_2-s_1)^2}{12}\end{cases} \tag{6-56}$$

的解。

当实际的服务时间分布为均匀分布时，图 6-9 比较了风险规避顾客的均衡进入概率 q_e 和社会最优进入概率 q^*。观察到部分信息取值范围、取值范围和均值、前三阶矩下的均衡进入概率都是相等的，且大于部分信息为均值或均值和方差下的均衡进入概率，甚至部分信息为均值下的均衡进入概率要小于社会最优进入概率。

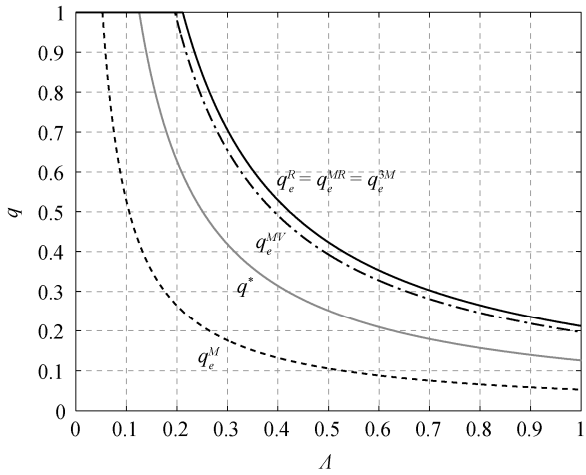

图 6-9　服务时间为均匀分布时的 q_e 和 q^*

6.4.2.2　指数分布

如果实际的服务时间分布函数是参数为 $1/\bar{s}$ 的指数分布，则单位社会福利函数为

$$SW(q)=\Lambda q\left(R-c\left(2\bar{s}^2+\frac{2\Lambda q\bar{s}^3(2-\Lambda q\bar{s})}{(1-\Lambda q\bar{s})^2}\right)\right) \tag{6-57}$$

则社会最优目标函数为 $\max\{SW(q),0\leq q\leq 1\}$。基于一阶最优条件 $SW'(q)=0$，可得到社会最优进入概率 q^*。

如果顾客被告知服务时间的均值 $\bar{s}$，则最大熵分布与实际分布吻合，这样均衡进入概率 q_e 由式（6-42）给出。如果顾客被告知服务时间的前三阶原点矩 $\bar{s}_1,\bar{s}_2,\bar{s}_3$，也就是 $\bar{s}_1=\bar{s},\bar{s}_2=2\bar{s}^2,\bar{s}_3=6\bar{s}^3$，则顾客的决策独立于服务时间的真实分布，所以均衡进入概率与只告知服务时间均值时相等。

如果顾客被告知服务时间的均值 $\bar{s}$ 和方差 σ_s^2，则 $\sigma_s^2=\bar{s}^2$，即 $\sigma_s/\bar{s}=1$。根据引理 6.2.4，最大熵分布是参数为 $1/\bar{s}$ 的指数分布，也与实际分布相吻合。换句话说，这三种类型部分信息下的均衡进入概率相同。

当实际的服务时间分布为指数分布时，图 6-10 比较了风险规避顾客的均衡

进入概率 q_e 和社会最优进入概率 q^*。显然，部分信息：均值、均值和方差、前三阶矩下的均衡进入概率不小于社会最优进入概率。

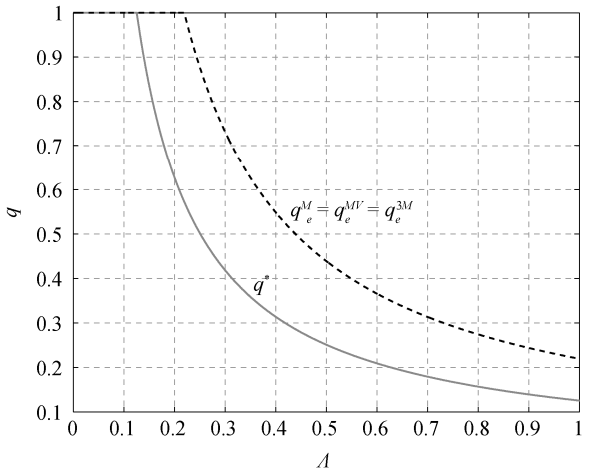

图 6-10　服务时间为指数分布时的 q_e 和 q^*

6.4.2.3　截尾指数分布

如果实际的服务时间分布函数是区间 $[s_1, s_2]$ 上参数为 λ_1 的截尾指数分布，则单位社会福利函数为

$$SW(q) = \Lambda q\left(R - c\left(\overline{s}_2 + \frac{(\Lambda q)^2 \overline{s}_2^2}{2(1-\Lambda q\overline{s}_1)^2} + \frac{\Lambda q\overline{s}_1\overline{s}_2}{1-\Lambda q\overline{s}_1} + \frac{\Lambda q\overline{s}_3}{3(1-\Lambda q\overline{s}_1)}\right)\right) \quad (6\text{-}58)$$

其中 $\overline{s}_1, \overline{s}_2, \overline{s}_3$ 由式(6-30)和式(6-44)给出。社会最优目标函数为 $\max\{SW(q), 0 \le q \le 1\}$。基于一阶最优条件 $SW'(q) = 0$，可得到社会最优进入概率 q^*。

如果顾客被告知服务时间的取值范围 $[s_1, s_2]$，则最大熵分布与实际分布不符。因此，均衡进入概率 q_e 由式（6-39）和式（6-40）给出。如果顾客被告知服务时间的均值 $\overline{s}_1$，则

$$q_e = \frac{R\lambda_1\left(e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}\right) - \sqrt{2Rc}\left(\left(e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}\right) + \lambda_1\left(s_1 e^{-\lambda_1 s_1} - s_2 e^{-\lambda_1 s_2}\right)\right)}{R\Lambda\left(\left(e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}\right) + \lambda_1\left(s_1 e^{-\lambda_1 s_1} - s_2 e^{-\lambda_1 s_2}\right)\right)} \quad (6\text{-}59)$$

如果顾客被告知服务时间的取值范围 $[s_1, s_2]$ 和均值 $\overline{s}_1$，则最大熵分布与实际分布吻合，这样均衡进入概率 q_e 由式（6-45）给出。如果顾客被告知服务时间的前三阶原点矩 $\overline{s}_1, \overline{s}_2, \overline{s}_3$，它们分别由式（6-30）和式（6-44）给出，则顾客的决策独立于服务时间的真实分布，所以均衡进入概率与告知服务时间均值和取值范围时相等。

如果顾客被告知服务时间的均值 $\bar{s}$ 和方差 σ_s^2，则最大熵分布是均值为

$$\bar{s} = \bar{s}_1 = \frac{1}{\lambda_1} + \frac{s_1 e^{-\lambda_1 s_1} - s_2 e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}} \tag{6-60}$$

方差为

$$\begin{aligned}\sigma_s^2 &= \frac{2}{\lambda_1^2} + \frac{s_1^2 e^{-\lambda_1 s_1} - s_2^2 e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}} + \frac{2(s_1 e^{-\lambda_1 s_1} - s_2 e^{-\lambda_1 s_2})}{\lambda_1 (e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2})} - \left(\frac{1}{\lambda_1} + \frac{s_1 e^{-\lambda_1 s_1} - s_2 e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}} \right)^2 \\ &= \frac{(e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2})^2 - \lambda_1^2 e^{-\lambda_1 (s_1 + s_2)} (s_2 - s_1)^2}{\lambda_1^2 (e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2})^2}\end{aligned} \tag{6-61}$$

的截尾正态分布。因此，均衡进入概率的分子为

$$\begin{aligned}q_e^N = &-\left(6\bar{s}R + c\bar{s}_3\right) + 3c\bar{s}\left(\bar{s}^2 + \sigma_s^2\right) \\ &+ \sqrt{-c\left(\left(\bar{s}^2 + \sigma_s^2\right)^2 \left(-18R + 9c\left(\bar{s}^2 + 2\sigma_s^2\right)\right) - 6c\bar{s}_3 s\left(\bar{s}^2 + \sigma_s^2\right) - c\bar{s}_3^2\right)}\end{aligned}$$

分母为

$$q_e^D = (-6R\bar{s}^2 + 3c\bar{s}^4 + 6c\sigma_s^2\bar{s}^2 + 3c(\sigma_s^2)^2 - 2c\bar{s}_3\bar{s})\Lambda$$

其中，$\bar{s}$ 和 σ_s^2 由式（6-60）和式（6-61）给出，

$$\bar{s}_3 = \int_0^\infty x^3 \frac{1}{\Phi\left(\frac{\mu}{\sigma}\right)} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x$$

μ 和 σ 是方程组

$$\begin{cases} \mu + \sigma \dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} = \dfrac{1}{\lambda_1} + \dfrac{s_1 e^{-\lambda_1 s_1} - s_2 e^{-\lambda_1 s_2}}{e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}} \\ \sigma^2 \left(1 - \dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \left(1 - \dfrac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\right)\right) = \dfrac{\left(e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}\right)^2 - \lambda_1^2 e^{-\lambda_1 (s_1 + s_2)} (s_2 - s_1)^2}{\lambda_1^2 \left(e^{-\lambda_1 s_1} - e^{-\lambda_1 s_2}\right)^2} \end{cases} \tag{6-62}$$

的解。

当实际的服务时间分布为截尾指数分布时，图 6-11 和图 6-12 比较了风险规避顾客的均衡进入概率 q_e 和社会最优进入概率 q^*。当 $\lambda_1 = 1.5$ 时，图 6-11 中的均衡进入概率存在关系 $q_e^{MR} = q_e^{3M} > q_e^{MV} > q_e^M > q_e^R$。然而，当 $\lambda_1 = 0.8$ 时，图 6-12 中的均衡进入概率存在关系 $q_e^{MR} = q_e^{3M} > q_e^{MV} > q_e^R > q_e^M$。而且，当顾客被告知服务时间范围或均值时，均衡进入概率小于社会最优进入概率。

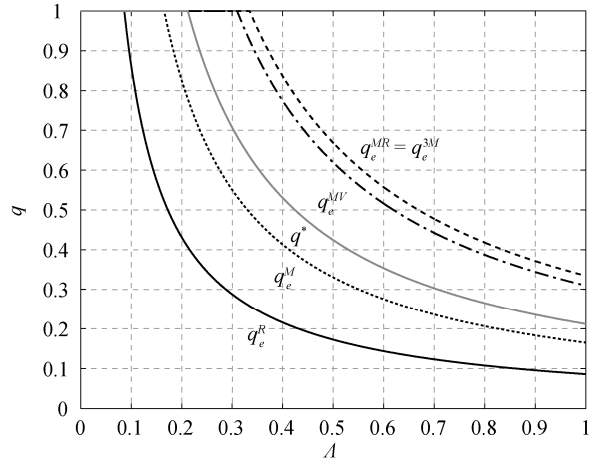

图 6-11　服务时间是参数为 $\lambda_1 = 1.5$ 的截尾指数分布时的 q_e 和 q^*

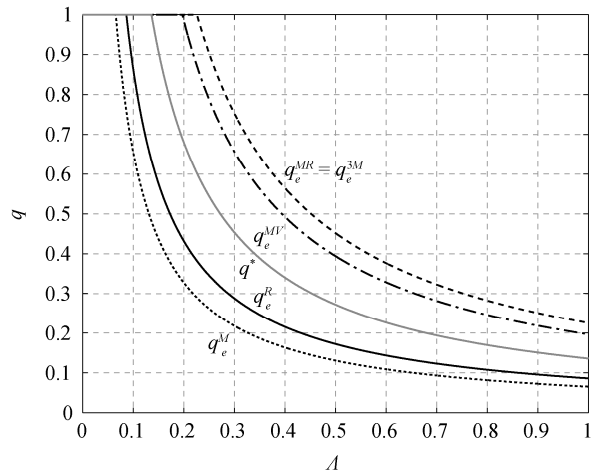

图 6-12　服务时间是参数为 $\lambda_1 = 0.8$ 的截尾指数分布时的 q_e 和 q^*

注释 6.4.1　对于均匀分布和截尾指数分布来说，如果部分服务信息为取值范围和均值，则均值总是不会大于取值范围的中值，所以顾客不会因为一个过大的服务时间均值而对队列望而却步。因此，本章通过数值算例得到的规律是合理的，即部分信息为均值和取值范围下的均衡进入概率不会小于部分信息为取值范围下的均衡进入概率。

6.4.2.4　截尾正态分布

如果实际的服务时间分布函数是区间（0,+∞）上参数为 μ 和 σ 的截尾正态分布，则单位社会福利函数为

$$SW(q)=\Lambda q\left(R-c\left(\left(\overline{s}^2+\sigma_s^2\right)+\frac{\left(\Lambda q\right)^2\left(\overline{s}^2+\sigma_s^2\right)^2}{2\left(1-\Lambda q\overline{s}\right)^2}+\frac{\Lambda q\left(3\overline{s}\left(\overline{s}^2+\sigma_s^2\right)+\overline{s}_3\right)}{3\left(1-\Lambda q\overline{s}\right)}\right)\right) \tag{6-63}$$

其中

$$\begin{cases}\overline{s}_3=\int_0^\infty x^3\dfrac{1}{\Phi\left(\dfrac{\mu}{\sigma}\right)}\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x\\ \overline{s}=\mu+\sigma\dfrac{\phi\left(\dfrac{\mu}{\sigma}\right)}{\Phi\left(\dfrac{\mu}{\sigma}\right)}\\ \sigma_s^2=\sigma^2\left(1-\dfrac{\phi\left(\dfrac{\mu}{\sigma}\right)}{\Phi\left(\dfrac{\mu}{\sigma}\right)}\left(1-\dfrac{\phi\left(\dfrac{\mu}{\sigma}\right)}{\Phi\left(\dfrac{\mu}{\sigma}\right)}\right)\right)\end{cases} \tag{6-64}$$

社会最优目标函数为 $\max\{SW(q),0\leq q\leq 1\}$。基于一阶最优条件 $SW'(q)=0$，可得到社会最优进入概率 q^*。

如果顾客被告知服务时间的均值 $\overline{s}$，即 $\overline{s}=\mu+\sigma\phi(\mu/\sigma)/\Phi(\mu/\sigma)$，则最大熵分布与实际服务时间分布不符。因此，

$$q_e=\frac{R-\overline{s}\sqrt{2Rc}}{R\Lambda\overline{s}}=\frac{R\Phi\left(\dfrac{\mu}{\sigma}\right)-\sqrt{2Rc}\left(\mu\Phi\left(\dfrac{\mu}{\sigma}\right)+\sigma\phi\left(\dfrac{\mu}{\sigma}\right)\right)}{R\Lambda\left(\mu\Phi\left(\dfrac{\mu}{\sigma}\right)+\sigma\phi\left(\dfrac{\mu}{\sigma}\right)\right)} \tag{6-65}$$

如果顾客被告知服务时间的均值 $\overline{s}$ 和方差 σ_s^2，则最大熵分布与实际服务时间分布相吻合，且 q_e 由式（6-48）和式（6-49）给出。如果顾客被告知服务时间的前三阶原点矩 $\overline{s}_1=\overline{s}$、$\overline{s}_2=\overline{s}^2+\sigma_s^2$ 和 $\overline{s}_3$，其中 $\overline{s}$、σ_s^2、$\overline{s}_3$ 由式（6-64）给出，则顾客的决策独立于服务时间的真实分布，所以均衡进入概率与告知服务时间均值和方差时相等。

当实际的服务时间分布为截尾正态分布时，图 6-13 比较了风险规避顾客的均衡进入概率 q_e 和社会最优进入概率 q^*。观察到部分信息为均值和方差下的均衡进入概率大于最优进入概率，而部分信息为均值下的均衡进入概率则小于最优进入概率。

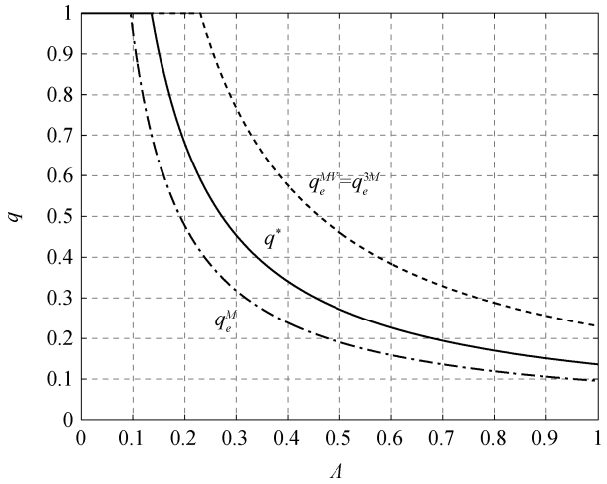

图 6-13　服务时间为截尾正态分布时的 q_e 和 q^*

6.4.3　利润最大

平行于对具有风险中立顾客的排队系统的分析过程，现在考虑在均衡状态下，具有风险规避顾客的排队系统中服务员在各种类型的部分服务信息下的利润和定价问题。假设服务价格为 p，则顾客在某种部分服务信息下的均衡进入概率满足

$$R-p=cE[W^2(q_e)] \tag{6-66}$$

基于式（6-6），解得

$$q_e(p)=\frac{6s_1(R-p)+c(s_3-3s_1s_2)-\sqrt{c^2\left((s_3+3s_1s_2)^2-18s_2^3\right)+18cs_2^2(R-p)}}{\Lambda\left(6s_1^2(R-p)-3cs_2^2+2cs_3s_1\right)} \tag{6-67}$$

因此，服务员的利润函数 $\Pi(p)$的分子为

$$\begin{aligned}\left(\prod(p)\right)^N=&p\left(6s_1(R-p)+c(s_3-3s_1s_2)\right)\\&-\sqrt{c^2\left((s_3+3s_1s_2)^2-18s_2^3\right)+18cs_2^2(R-p)}\Big)\end{aligned} \tag{6-68}$$

分母为

$$\left(\prod(p)\right)^D=6s_1^2(R-p)-3cs_2^2+2cs_3s_1 \tag{6-69}$$

令 $\prod'(p)=0$，可以得到垄断价格的显式形式，且最大利润为 $\prod^*=\Lambda q_e(p^*)p^*$。对于任意给定的实际服务时间分布，在各种部分服务时间信息下，风险规避顾客根

据最大熵原理将对服务时间的前三阶矩给予不同的估计。

假定实际服务时间分布分别为均匀分布、指数分布、截尾指数分布、截尾正态分布。部分服务信息分别为取值范围、均值、取值范围和均值、均值和方差、前三阶矩，图 6-14 至图 6-18 分别比较了不同情况下服务员的利润和价格。这里只对图 6-17 和图 6-18 加以说明。

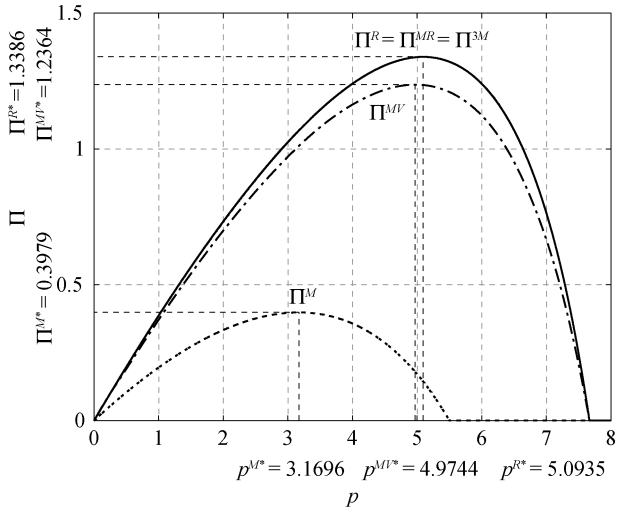

图 6-14　服务时间为均匀分布时的服务员利润

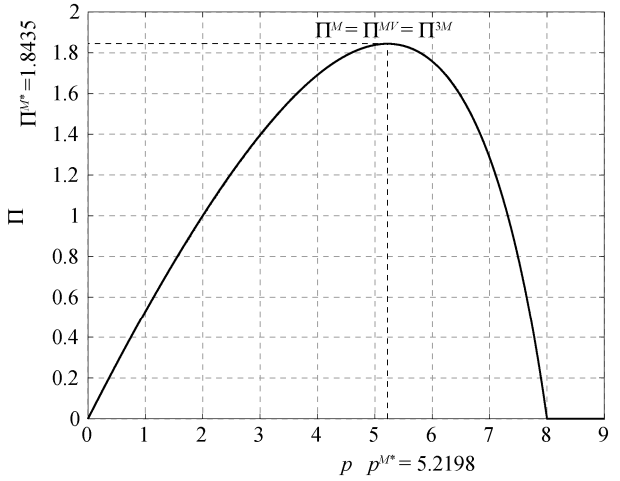

图 6-15　服务时间为指数分布时的服务员利润

从图 6-17 观察到如果部分信息为取值范围，则垄断价格为 $p^{R^*}=3.391$，最大利润为 $\prod^{R^*}=0.4257$；如果部分信息为均值，则 $P^{M^*}=3.8795$，$\prod^{M^*}=0.7028$；如果

部分信息为均值和方差，则 $p^{MV^*}=5.4465$，$\prod^{MV^*}=1.7024$；如果部分信息为取值范围和均值或前三阶矩，则 $p^{MR^*}=5.6163$，$\prod^{MR^*}=1.8676$。而在图 6-18 中，如果部分信息为均值，则 $p^{M^*}=2.9114$，$\prod^{M^*}=0.3176$；如果部分信息为均值和方差，则 $p^{MV^*}=4.651$，$\prod^{MV^*}=1.0118$；如果部分信息为取值范围和均值或前三阶矩，则 $p^{MR^*}=4.852$，$\prod^{MR^*}=1.173$。因此，在此得到相似的结论，即顾客被告知的部分服务信息越多越丰富，服务员的利润就会越大。

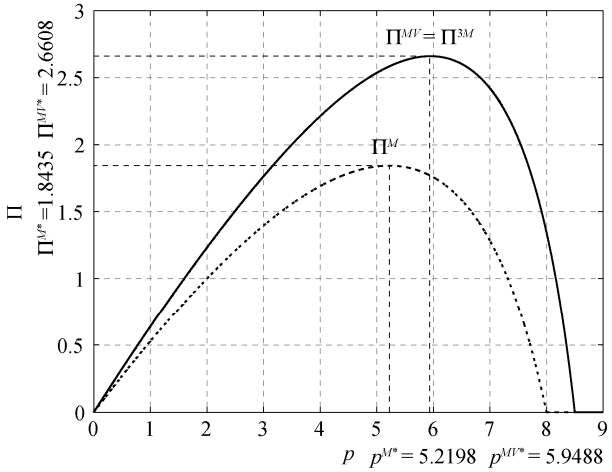

图 6-16　服务时间为截尾正态分布时的服务员利润

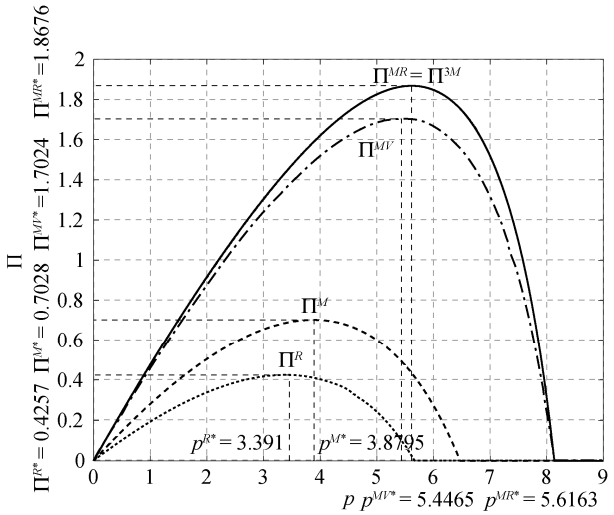

图 6-17　服务时间是参数为 $\lambda_1=3$ 的截尾指数分布时的服务员利润

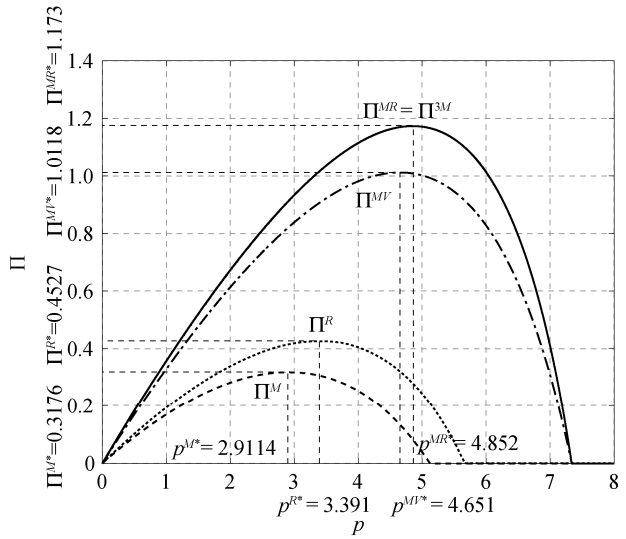

图 6-18　服务时间是参数为 $\lambda_1 = 1.5$ 的截尾指数分布时的服务员利润

6.5　相关文献评述

熵最初是热力学中的一个非常重要的概念。1948 年，电气工程师 Shannon[97] 创立了信息论，首次将信息量与熵联系起来，并用信息熵定量地刻画了随机事件的不确定性。可以看到，Shannon 的工作使熵闯入了距其出生地非常遥远的概率论与数理统计、通信、计算机领域。1957 年，Jaynes[98]首次提出了最大熵原理。对于初学者，还可以参考 Kapur 和 Kesavan[96]给出的相关文献，其中对最大熵原理给出了详尽的介绍。

除了根据最大熵原理对服务时间分布进行估计之外[95]，还可以根据最小最大后悔值准则（Minimax Regret Criterion）来进行估计。最小最大后悔值准则是管理决策中不确定型决策的一种方法。后悔值是指当某种自然状态出现时，决策者由于从若干方案中选优时没有采取能获得最大收益的方案，而采取了其他方案，以致在收益上产生的某种损失。最小最大后悔值准则一般应用于库存控制和收益管理等决策问题之中。Perakis 和 Roels[99]研究了一个具有部分需求分布信息的报童模型，并根据最小最大后悔值准则导出了报童的订货量。

然而，将最小最大后悔值准则应用于排队系统中顾客的行为分析的研究工作还很少。Veeraraghavan 和 Debo[100]考虑了具有两个服务员的排队系统，服务员服

务率未知，但顾客在选择之前被告知有关两服务员的正面和负面两种信息，根据最小最大后悔值准则，顾客会在具有较短队列长度的服务员前进行排队。孙微等[56]研究了具有多重休假策略和五类决策标准（包括最小最大后悔值准则）的排队系统中顾客的均衡止步策略。

第 7 章　具有随机服务参数的排队系统

7.1　服务参数的不确定性

在随机服务系统中，除了大家熟知的服务员服务率可变之外，和顾客最终剩余服务效用密切相关的服务参数也往往是可变的。比如，随着服务环境的恶化，顾客的单位等待损耗会增加或服务收益会减少等。重要的是，这些系统参数一般表现为随机变动，当其中某个参数变化时，服务员应采取何种定价策略以保证利润最大是其最关心的问题。再者，顾客对这种随机变动的规律（即参数的随机分布类型）已知与否会对服务员利润产生何种影响，利润对于参数不确定性的灵敏度如何，也都是系统管理者关心且急需解决的问题。

本章内容为作者研究成果[101]的进一步完善和补充。

7.2　模型描述

考虑一些队长不可视的 Markov 排队系统，潜在顾客均为风险中立顾客且到达率足够大。假设服务员服务率为 μ，单位等待损耗为 C，服务价格为 T，服务收益为 R[①]。在此，我们考虑三种类型的不确定性，即假定 μ、C、R 分别为离散型随机变量。记符号 q 表示某类服务参数，分布律为 $P=\{q=q_k\}=p_k\ (1\le k\le n)$

① 对于确定的服务参数 R、C、μ 和 T，Hassin 与 Haviv[36]已经给出了有关服务员定价和最大利润的结果。如果 $R-T<C/\mu$，则均衡到达率是零。否则，均衡到达率满足 $R-T=C/(\mu-\lambda)$ 或 $\lambda=\mu-C/(R-T)$。假设 R、C 和 μ 给定，T 是一个决策变量，则服务员利润率为 $\Pi=\lambda T$。如果 $R<C/\mu$，则服务员利润为零。否则，令 $T=R-\sqrt{RC/\mu}$，可使 Π 达到最优。因此，最大利润率为

$$\Pi^*=\begin{cases}\left(\sqrt{R\mu}-\sqrt{C}\right)^2 & R>\dfrac{C}{\mu}\\ 0 & R\le\dfrac{C}{\mu}\end{cases}$$

且 $\sum_{k=1}^{n} p_k = 1$。定义 q 的均值和方差分别为 $\bar{q} = \sum_{k=1}^{n} p_k q_k$ 和 $\mathrm{Var}(q) = \sum_{k=1}^{n} p_k (q_k - \bar{q})^2$。

对于上述每类服务参数，我们分析 5 种情况，建立 5 个子模型。首先，第一种情况，该可变服务参数的分布信息对顾客是封闭的，服务员实行单一定价策略。在其他 4 种情况中，该可变服务参数的分布信息对顾客是公开的。然而，服务员定价决策的制定受到许多因素的影响，如采用多层次定价的费用问题及相关的定价制度约束等。因此，我们将其可能的定价决策进行分化。第二种情况，服务员对于可变服务参数的 n 个实现值只能实行单一定价。第三种情况，服务员对于每个实现值都分别采用不同的定价，即对较高服务率、较低单位损耗或较高服务收益征收较高的费用。称上述两种定价策略为纯策略。第四种情况，服务员对前 $k(1\leq k\leq n\text{-}2)$个实现值采取不同的定价，对后 $n\text{-}k$ 个实现值采取单一定价。相反，第五种情况，服务员对后 $n\text{-}k(2\leq k\leq n\text{-}1)$个实现值采取不同的定价，对前 k 个实现值采取单一定价。称上述两种策略为混合策略。事实上，我们已经全面考虑了所有可能的定价策略。

现定义一些符号，特别是在各种不确定性和定价策略条件下服务员的利润率：

- λ_k　$\mu(C,R)$的第 $k(1\leq k\leq n)$个实现值 $\mu_k(C_k, R_k)$下顾客的均衡到达率；
- Π_{μ}^{un}，Π_{C}^{un}，Π_{R}^{un}　参数 $\mu(C,R)$的分布信息封闭条件下服务员的利润率；
- $\Pi_{n\mu}^{in}$，Π_{nC}^{in}，Π_{nR}^{in}　参数 $\mu(C,R)$的分布信息公开且对参数的 n 个实现值实行 n 重定价条件下服务员的利润率；
- $\Pi_{1\mu}^{in}$，Π_{1C}^{in}，Π_{1R}^{in}　参数 $\mu(C,R)$的分布信息公开且对参数的 n 个实现值实行单一定价条件下服务员的利润率；
- $\Pi_{(k+1)\mu}^{in}$，$\Pi_{(k+1)C}^{in}$，$\Pi_{(k+1)R}^{in}$　参数 $\mu(C,R)$的分布信息公开且对参数的 n 个实现值实行 $k+1$ $(1\leq k\leq n\text{-}2)$重定价条件下服务员的利润率；
- $\Pi_{(n-k+1)\mu}^{in}$，$\Pi_{(n-k+1)C}^{in}$，$\Pi_{(n-k+1)R}^{in}$　参数 $\mu(C,R)$的分布信息公开且对参数的 n 个实现值实行 $n\text{-}k+1(2\leq k\leq n\text{-}1)$重定价条件下服务员的利润率。

7.3　可变服务速率

假定服务率的分布律为 $P=\{\mu=\mu_k\}=p_k\,(1\leq k\leq n)$，$\sum_{k=1}^{n} p_k = 1$ 且 $\mu_1 > \mu_2 > \cdots > \mu_n$。

7.3.1　服务速率信息封闭

假定服务率的分布信息对顾客是封闭的，定义服务员的利润最大化价格为 T_{μ}^{un}，顾客的均衡到达率为 λ_{μ}^{un}。因为 $\lambda_{\mu}^{un} \geq 0$ 且 $\prod_{\mu}^{un} \geq 0$，不失一般性，

$$R - T_{\mu}^{un} \geq C\sum_{k=1}^{n}\frac{p_k}{\mu_k} \tag{7-1}$$

引理 7.3.1　服务员的利润最大化价格 T_{μ}^{un} 和顾客的均衡到达率 λ_{μ}^{un} 是以下方程组

$$\begin{cases} \dfrac{R}{C} - \lambda\sum\limits_{k=1}^{n}\dfrac{p_k}{(\mu_k-\lambda)^2} = \sum\limits_{k=1}^{n}\dfrac{p_k}{\mu_k-\lambda} \\ T = \lambda C\sum\limits_{k=1}^{n}\dfrac{p_k}{(\mu_k-\lambda)^2} \end{cases} \tag{7-2}$$

的解。

证明　均衡状态下，有以下关系式

$$\frac{R-T}{C} = \sum_{k=1}^{n}\frac{p_k}{\mu_k-\lambda} \tag{7-3}$$

其中，λ 可以看作是 T 的隐函数。因此，得到 λ 关于 T 的一阶导数

$$\lambda_T' = -\frac{1}{C\sum\limits_{k=1}^{n}\dfrac{p_k}{(\mu_k-\lambda)^2}}$$

由于利润 $\prod = \lambda T$，一阶最优条件可表示为 $\lambda + \lambda_T' T = 0$，即

$$\lambda - \frac{T}{C\sum\limits_{k=1}^{n}\dfrac{p_k}{(\mu_k-\lambda)^2}} = 0 \tag{7-4}$$

求解方程（7-3）与方程（7-4），得到 λ_{μ}^{un} 和 T_{μ}^{un}，然后再计算出 $\prod_{\mu}^{un} = \lambda_{\mu}^{un}\ T_{\mu}^{un}$。

7.3.2　服务速率信息公开

7.3.2.1　服务速率信息公开——n 重价格

假定服务率的分布信息对顾客是公开的，且服务员可根据服务率的 n 个实现值实行 n 重定价。如果 $R > C/\mu_n$，不论服务率为何实现值，一定有顾客进入系统，因此利润率以概率 $p_k(1\leq k\leq n)$由 n 部分构成；如果 $C/\mu_{j-1} < R \leq C/\mu_j (2\leq j\leq n)$，只有 $\mu > \mu_j$ 时才能有顾客进入，因此利润率以概率 $p_k(1\leq k\leq j-1)$ 由 $j-1$ 部分构成；然而，如果 $R\leq C/\mu_1$，将没有顾客进入系统。基于 Hassin 与 Haviv[36]给出的上述

相关结果，得到

$$\prod_{n\mu}^{in}=\begin{cases}\sum_{k=1}^{n}p_k\left(\sqrt{R\mu_k}-\sqrt{C}\right)^2, 若\dfrac{R}{C}>\dfrac{1}{\mu_n}\\ \sum_{k=1}^{j-1}p_k\left(\sqrt{R\mu_k}-\sqrt{C}\right)^2, 若\dfrac{1}{\mu_{j-1}}<\dfrac{R}{C}\le\dfrac{1}{\mu_j}(2\le j\le n)\\ 0, \qquad 若\dfrac{R}{C}\le\dfrac{1}{\mu_1}\end{cases} \tag{7-5}$$

7.3.2.2 服务速率信息公开 —— 单一价格

假定服务率的分布信息对顾客是公开的，但服务员只能实行单一定价。首先，假定价格 T 不太高，即 $T<R-C/\mu_n$，因此，对于 $\mu=\mu_k(1\le k\le n)$，有 $\lambda_k>0$。给定 $\mu=\mu_k$，顾客到达率为 $\lambda_k=\mu_k-C/(R-T)$，而给定 T，顾客的平均到达率为

$$\lambda=\sum_{k=1}^{n}p_k\lambda_k=\sum_{k=1}^{n}p_k\left(\mu_k-\frac{C}{R-T}\right)=\sum_{k=1}^{n}p_k\mu_k-\frac{C}{R-T}=\bar{\mu}-\frac{C}{R-T} \tag{7-6}$$

其中 $\bar{\mu}=\sum_{k=1}^{n}p_k\mu_k$。因为平均利润率为 $\Pi_\mu(T)=\lambda T$，因此由一阶最优条件 $\prod'_\mu(T)=0$ 可得到利润最大化价格 T_μ^1，即

$$T_\mu^1=R-\sqrt{\frac{RC}{\bar{\mu}}}\ ① \tag{7-7}$$

到达率为 $\lambda_\mu^1=\bar{\mu}-\sqrt{C\bar{\mu}/R}$。因此，平均利润率 $\prod_\mu^1$ 为

$$\prod_\mu^1=\left(\sqrt{R\bar{\mu}}-\sqrt{C}\right)^2 \tag{7-8}$$

另一方面，服务员也可能会选择一个相对较高的价格，使得当 $\mu=\mu_k$ 或 $\mu=\mu_j$ $(2\le j,k\le n, j>k)$ 时，顾客不会进入系统，即 $R-C/\mu_k\le T<R-C/\mu_{k-1}$。在此情况下，$\lambda_1,\lambda_2,\cdots,\lambda_{k-1}>0$ 且 $\lambda_k,\lambda_{k+1},\cdots,\lambda_n=0$，因此，当 $\mu=\mu_{k-1}$ 时，服务员将选择利润最大化价格 $T_{k\mu}^2=R-\sqrt{RC/\mu_{k-1}}$，相应的利润率 $\prod_{k\mu}^2$ 为

$$\prod_{k\mu}^2=\left(R-\sqrt{\frac{RC}{\mu_{k-1}}}\right)\sum_{j=1}^{k-1}p_j\left(\mu_j-\sqrt{\frac{C\mu_{k-1}}{R}}\right),\ 2\le k\le n \tag{7-9}$$

服务员将选择 $\prod_{k\mu}^2$ 与 $\prod_\mu^1$ 中的较大者。因此，实际的利润率为

① 由于假定 $T<R-C/\mu_n$，另一个解 $T_\mu^1=R+\sqrt{\frac{RC}{\bar{\mu}}}$ 被舍弃。

$$\Pi_{1\mu}^{in}=\begin{cases}\Pi_{\mu}^{1}, & \text{若}\Pi_{\mu}^{1}>\max\{\Pi_{k\mu}^{2},2\le k\le n\}\ \&\ \dfrac{R}{C}>\dfrac{1}{\mu_n}\\ \Pi_{k\mu}^{2}, & \text{若}\Pi_{k\mu}^{2}>\max\{\Pi_{\mu}^{1},\Pi_{j\mu}^{2},2\le j,k\le n,j\ne k\}\ \&\ \dfrac{1}{\mu_{k-1}}<\dfrac{R}{C}\le\dfrac{1}{\mu_k}\\ 0, & \text{若}\dfrac{R}{C}\le\dfrac{1}{\mu_1}\end{cases}\quad(7\text{-}10)$$

假定 $R=20$，$C=4$，$p_1=p_2=p_3=1/3$，$\mu_1=4$，$\mu_2=2$，$\mu_3=1$ 时，图 7-1（a）展示的是平均利润率与价格之间的变化关系。在这些参数取值之下，$R-C/\mu_3=16$ 且 $R-C/\mu_2=18$。而且，注意到即使当 $\mu=\mu_2$ 或 $\mu=\mu_3$ 时，服务员仍值得降低价格以吸引更多的顾客。在图 7-1（b）中，将 μ_1 从 4 增加到 16，服务员如果选择一个更高的价格则将使利润更大。但在此价格下，当 $\mu=\mu_3$ 时，将没有顾客选择进入。在图 7-1（c）中，将 μ_1 再次从 16 增加到 64，服务员会选择一个比图 7-1（b）中还要高的价格。但在此价格下，当 $\mu=\mu_2$ 或 $\mu=\mu_3$ 时，将没有顾客选择进入。

7.3.2.3　服务速率信息公开 —— $k+1$ 重价格

假定服务中对服务率的前 $k(1\le k\le n-2)$个实现值采取不同的定价，而对其后的 n-k 个实现价值采取单一定价。因此，服务员可采取 k+1 重定价。在这种混合定价策略下，即使在服务率较小的情况下，也可以鼓励顾客进入系统。

对于μ的前 k 个实现值，服务员可获取顾客的所有剩余效用，因此，根据信息公开且 n 重价格情况下的结果，得到利润率的第一部分，记为 $\Pi_{(k+1)\mu}^{1}$，其为

$$\Pi_{(k+1)\mu}^{1}=\begin{cases}\displaystyle\sum_{l=1}^{k}p_l\left(\sqrt{R\mu_l}-\sqrt{C}\right)^2, & \text{若}\dfrac{R}{C}>\dfrac{1}{\mu_k}\\ \displaystyle\sum_{l=1}^{j-1}p_l\left(\sqrt{R\mu_l}-\sqrt{C}\right)^2, & \text{若}\dfrac{1}{\mu_{j-1}}<\dfrac{R}{C}\le\dfrac{1}{\mu_j},2\le j\le k\\ 0, & \text{若}\dfrac{R}{C}\le\dfrac{1}{\mu_1}\end{cases}\quad(7\text{-}11)$$

对于 μ 的其他 n-k 个实现值，根据信息公开且单一价格的分析过程，可类似分析服务员的定价策略和利润率。首先假定 $T<R-C/\mu_n$，因此 $\lambda_{k+1},\lambda_{k+2},\cdots,\lambda_n>0$。给定 $\mu=\mu_l$，顾客到达率为 $\lambda_1=\mu_l-C/(R-T)(k+1\le l\le n)$。给定 T，顾客的平均到达率为

$$\lambda=\sum_{l=k+1}^{n}p_l\lambda_l=\sum_{l=k+1}^{n}p_l\left(\mu_l-\frac{C}{R-T}\right)\quad(7\text{-}12)$$

因为平均利润率为 $\Pi_{\mu}^{2}(T)=\lambda T$，由一阶最优条件 $\Pi_{\mu}^{2'}(T)=0$ 可得到利润最大化价格

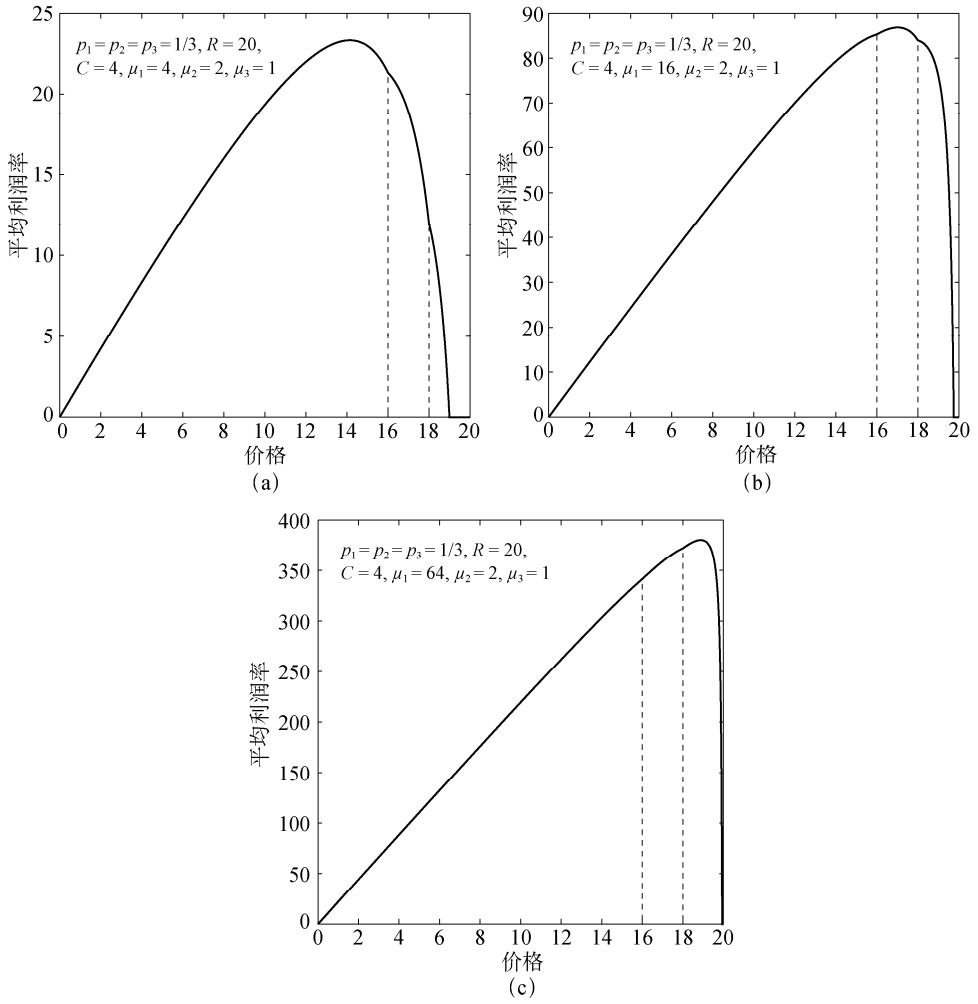

图 7-1　平均利润率与价格之间的变化关系

$$T_\mu^2 = R - \sqrt{\frac{\sum_{l=k+1}^{n} p_l RC}{\sum_{l=k+1}^{n} p_l \mu_l}}$$

以及相应的到达率为

$$\lambda_\mu^2 = \sum_{l=k+1}^{n} p_l \left(\mu_l - \sqrt{\frac{\sum_{l=k+1}^{n} p_l \mu_l C}{\sum_{l=k+1}^{n} p_l R}} \right)$$

因此，利润率的第二部分，记为 $\prod_\mu^2$，为

$$\prod_{\mu}^{2}=\lambda_{\mu}^{2}T_{\mu}^{2}=\left(\sqrt{R\sum_{l=k+1}^{n}p_l\mu_l}-\sqrt{C\sum_{l=k+1}^{n}p_l}\right)^2 \tag{7-13}$$

另一方面，假定 $R-C/\mu_j \le T < R-C/\mu_{j-1}(k+2\le j\le n)$，因此 $\lambda_{k+1},\lambda_{k+2},\cdots,\lambda_{j-1}>0$ 且 $\lambda_j,\lambda_{j+1},\cdots,\lambda_n=0$。当 $\mu=\mu_{j-1}$ 时，服务员将选择利润最大化价格，记为 $T_{j\mu}^2$，为 $T_{j\mu}^2=R-\sqrt{RC/\mu_{j-1}}$，因此利润率的第二部分，记为 $\prod_{j\mu}^2$，其为

$$\prod_{j\mu}^{2}=\left(R-\sqrt{\frac{RC}{\mu_{j-1}}}\right)\sum_{l=k+1}^{j-1}p_l\left(\mu_l-\sqrt{\frac{C\mu_{j-1}}{R}}\right),\ k+2\le j\le n \tag{7-14}$$

从而，总平均利润率 $\prod_{(k+1)\mu}^{in}$ 归纳如下

$$\prod_{(k+1)\mu}^{in}=\begin{cases}\sum_{l=1}^{k}p_l\left(\sqrt{R\mu_l}-\sqrt{C}\right)^2+\prod_{\mu}^{2}, & 若\dfrac{R}{C}>\dfrac{1}{\mu_n}\\ \sum_{l=1}^{k}p_l\left(\sqrt{R\mu_l}-\sqrt{C}\right)^2+\prod_{j\mu}^{2}, & 若\dfrac{1}{\mu_{j-1}}<\dfrac{R}{C}\le\dfrac{1}{\mu_j},k+2\le j\le n\\ \sum_{l=1}^{i-1}p_l\left(\sqrt{R\mu_l}-\sqrt{C}\right)^2, & 若\dfrac{1}{\mu_{i-1}}<\dfrac{R}{C}\le\dfrac{1}{\mu_i},2\le i\le k+1\\ 0, & 若\dfrac{R}{C}\le\dfrac{1}{\mu_1}\end{cases} \tag{7-15}$$

假定服务率的所有实现值等概率出现，图 7-2 比较了 $n=3,4,5$ 时 $\prod_{(k+1)\mu}^{in}(1\le k\le n\text{-}2)$的大小。我们观察到不论 n 和 k 的取值如何，$\prod_{(k+1)\mu}^{in}$ 都随 R 递增，但随 n 递减。而且，对于 n 的同一取值，$\prod_{(k+1)\mu}^{in}$ 随 k 递增，原因在于当 k 较大时，服务

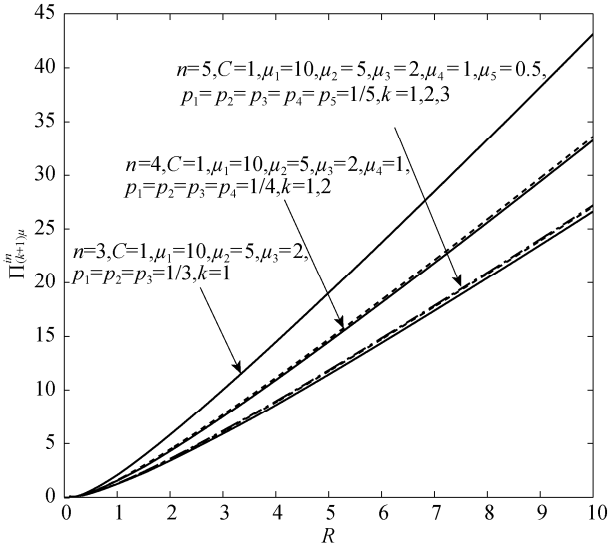

图 7-2　$\prod_{(k+1)\mu}^{in}(1\leq k\leq n\text{-}2,\ n=3,4,5)$

员可获取更多顾客的剩余效用。然而，对于固定的 n，观察到对不同的 k 值，利润之间的差距并不明显，尤其是当 μ 的 n 个实现值比较接近时更是如此。

7.3.2.4　服务速率信息公 —— n-k+1 重价格

假定服务员对服务率的后 n-k $(2\leq k\leq n\text{-}1)$个实现值采取不同的定价，而对前 k 个实现值采取单一定价。因此，服务员可采取 n-k +1 重定价。

对于 μ 的后 n-k 个实现值，服务员可获取顾客的所有剩余效用，因此根据信息公开且 n 重价格情况下的结果，得到利润率的第二部分，记为 $\prod_{(n\text{-}k+1)\mu}^{2}$，为

$$\prod_{(n-k+1)\mu}^{2}=\begin{cases}\sum_{l=k+1}^{n}p_l\left(\sqrt{R\mu_l}-\sqrt{C}\right)^2, & 若\dfrac{R}{C}>\dfrac{1}{\mu_n}\\ \sum_{l=k+1}^{j-1}p_l\left(\sqrt{R\mu_l}-\sqrt{C}\right)^2, & 若\dfrac{1}{\mu_{j-1}}<\dfrac{R}{C}\leq\dfrac{1}{\mu_j},\ k+2\leq j\leq n\\ 0, & 若\dfrac{R}{C}\leq\dfrac{1}{\mu_{k+1}}\end{cases}\quad(7\text{-}16)$$

对于前 k 个实现值，首先假设 $T<R-C/\mu_k$，所以 $\lambda_1,\lambda_2,\cdots,\lambda_k>0$。给定 $\mu=\mu_l(1\leq l\leq k)$，顾客到达率为 $\lambda_l=\mu_l-C/(R-T)$。因此，给定价格 T，顾客的平均到达率为

$$\lambda=\sum_{l=1}^{k}p_l\lambda_l=\sum_{l=1}^{k}p_l\left(\mu_l-\frac{C}{R-T}\right)\qquad(7\text{-}17)$$

因为平均利润率为 $\prod_{\mu}^{1}(T)=\lambda T$，由一阶最优条件 $\prod_{\mu}^{1'}(T)=0$ 可得到利润最大化价格，记为 $T_{(n\text{-}k+1)\mu}^{1}$，其为

$$T_{(n-k+1)\mu}^{1}=R-\sqrt{\frac{\sum_{l=1}^{k}p_lRC}{\sum_{l=1}^{k}p_l\mu_l}}$$

以及相应的到达率为

$$\lambda_{(n-k+1)\mu}^{1}=\sum_{l=1}^{k}p_l\left(\mu_l-\sqrt{\frac{\sum_{l=1}^{k}p_l\mu_lC}{\sum_{l=1}^{k}p_lR}}\right)$$

因此，利润率的第一部分，记为 $\prod_{(n\text{-}k+1)\mu}^{1}$，其为

$$\prod_{(n-k+1)\mu}^{1}=\lambda_{(n-k+1)\mu}^{1}T_{(n-k+1)\mu}^{1}=\left(\sqrt{R\sum_{l=1}^{k}p_l\mu_l}-\sqrt{C\sum_{l=1}^{k}p_l}\right)^2\qquad(7\text{-}18)$$

另一方面，假定 $R-C/\mu_j\leq T<R-C/\mu_{j-1}(2\leq j\leq k)$，因此 $\lambda_1,\lambda_2,\cdots,\lambda_{j-1}>0$ 且 λ_j，

$\lambda_{j+1}, \cdots, \lambda_k = 0$。当 $\mu = \mu_{j-1}$ 时，服务员将选择利润最大化价格 $T_{j\mu}^1$，且 $T_{j\mu}^1 = R - \sqrt{RC/\mu_{j-1}}$。因此利润率的第一部分，记为 $\prod_{j\mu}^1$，其为

$$\prod_{j\mu}^1 = \left(R - \sqrt{\frac{RC}{\mu_{j-1}}}\right)\sum_{l=1}^{j-1} p_l \left(\mu_l - \sqrt{\frac{C\mu_{j-1}}{R}}\right),\ 2 \le j \le k \tag{7-19}$$

从而，总平均利润率 $\prod_{(n-k+1)\mu}^{in}$ 归纳如下

$$\prod_{(n-k+1)\mu}^{in} = \begin{cases} \sum_{l=k+1}^{n} p_l \left(\sqrt{R\mu_l} - \sqrt{C}\right)^2 + \prod_{(n-k+1)\mu}^1, & \text{若} \dfrac{R}{C} > \dfrac{1}{\mu_n} \\ \sum_{l=k+1}^{j-1} p_l \left(\sqrt{R\mu_l} - \sqrt{C}\right)^2 + \prod_{(n-k+1)\mu}^1, & \text{若} \dfrac{1}{\mu_{j-1}} < \dfrac{R}{C} \le \dfrac{1}{\mu_j}, k+2 \le j \le n \\ \prod_{(n-k+1)\mu}^1, & \text{若} \dfrac{1}{\mu_k} < \dfrac{R}{C} \le \dfrac{1}{\mu_{k+1}} \\ \prod_{i\mu}^1, & \text{若} \dfrac{1}{\mu_{i-1}} < \dfrac{R}{C} \le \dfrac{1}{\mu_i}, 2 \le i \le k \\ 0, & \text{若} \dfrac{R}{C} \le \dfrac{1}{\mu_1} \end{cases} \tag{7-20}$$

假定服务率的所有实现值等概率出现，图 7-3 比较了 $n = 3,4,5$ 时 $\prod_{(n-k+1)\mu}^{in}$ $(2 \le k \le n-1)$的大小。图 7-2 与图 7-3 的主要区别为对于 n 的同一取值，$\prod_{(n-k+1)\mu}^{in}$ 随 k 递减，这是由于当 k 较小时，服务员可以获取更多的顾客剩余效用。

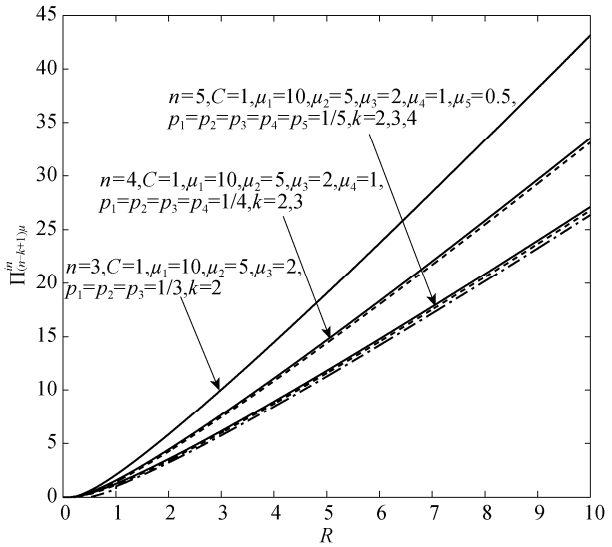

图 7-3　$\prod_{(n-k+1)\mu}^{in}$ $(2 \le k \le n-1,\ n = 3,4,5)$

图 7-4 则比较了 $n=3,4,5$ 情况下其他三类利润率。不论 n 的取值如何以及服务率分布信息是否公开，这三类利润率均随 R 递增。对于任意的 n，$\Pi_{n\mu}^{in} \geq \Pi_{1\mu}^{in} \geq \Pi_{\mu}^{un}$，且利润率均随 n 递减。有趣的是，信息封闭和信息公开条件下的利润率之间的差距随 n 递增。这说明信息量对服务员利润的影响是随服务率实现值个数的增多而不断增大的。而且，如果服务员无法有效控制 n 值，对顾客封闭信息是对其最不利的选择。

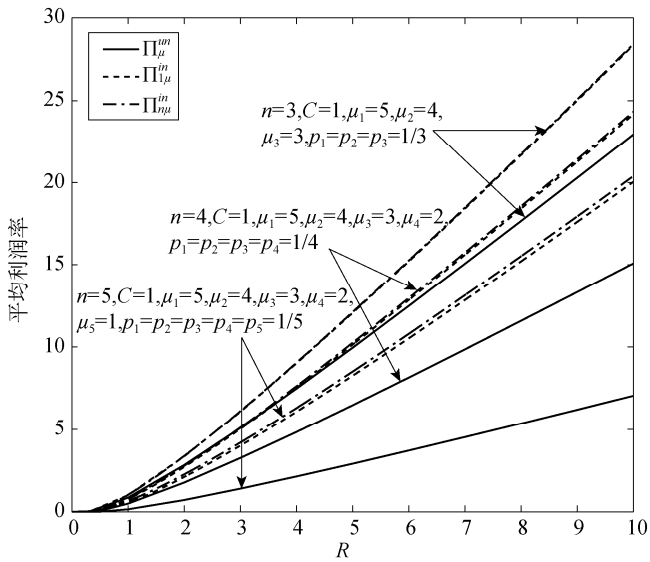

图 7-4　$\Pi_{\mu}^{un}, \Pi_{1\mu}^{in}, \Pi_{n\mu}^{in}$ $(n=3,4,5)$

7.3.3　可变服务速率对利润的影响

下面，我们将数值观察服务率的不确定性对服务员利润的影响。考虑不确定性下的两种情况，且在两种情况下固定 $\bar{\mu}=\sum_{k=1}^{n} p_k \mu_k$。

- 第一种情况下，除了 μ_k 之外，固定所有的 $\mu_j(1\leq j, k\leq n, j\neq k)$。假定存在一系列的关于 $p_j\,(1\leq j\leq n)$和 μ_k 的关系方程 $\{f_j(p_1, p_2, \cdots, p_n, \mu_k)=0, 1\leq j\leq n-2\}$。通过同时变换 $p_j\,(1\leq j\leq n)$和 μ_k，$\bar{\mu}$保持不变但 $\mathrm{Var}(\mu)$发生变化。在图 7-5（a）中，假定 $n=3$，$R=20$，$C=1$，$\bar{\mu}=2$，$\mu_1=3$，$\mu_2=2$ 以及 $p_1=2p_2$。

- 第二种情况下，固定所有满足 $0\leq p_k\leq 1$ 和 $\sum_{k=1}^{n} p_k=1$ 条件的概率 $p_k(1\leq k\leq n)$，且假定存在一系列关于 $\mu_k(1\leq k\leq n)$的方程 $\{f_j(\mu_1, \mu_2, \cdots, \mu_n)=0,$

$1 \le j \le n\text{-}2\}$[①]。通过同时变换 μ_k，$\bar{\mu}$保持不变但 Var(μ)发生变化。在图 7-5（b）中，假定 $n=3$，$R=20$，$C=1$，$\bar{\mu}=2$，$p_1=p_2=1/6$，$p_3=2/3$ 及 $\mu_1=4/3\mu_2$。

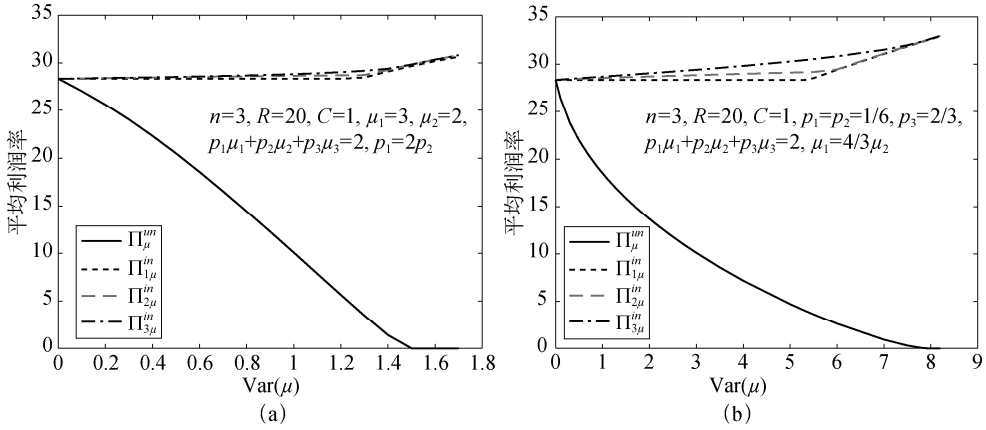

图 7-5　服务率的不确定性对服务员利润的影响

从图 7-5 中，观察到$\prod_{\mu}^{un}$随 Var(μ)递减，且$\prod_{\mu}^{un} \le \prod_{1\mu}^{in} \le \prod_{2\mu}^{in} \le \prod_{3\mu}^{in}$始终成立。利润率$\prod_{\mu}^{un}$在两种情况下的唯一区别在于图 7-5（a）中，它为凹减函数，而在图 7-5（b）中，它为凸减函数。然而，由于$\prod_{2\mu}^{in}(k=1)$[②]和$\prod_{3\mu}^{in}$为 Var(μ)的非减函数，因此它们的曲线均为递增曲线。至于$\prod_{1\mu}^{in}$，先是保持不变，然后再递增，直至非常逼近$\prod_{3\mu}^{in}$。

7.4　可变等待损耗

假定单位时间等待损耗的分布律为 $P\{C=C_k\}=p_k$ $(1 \le k \le n)$，$\sum_{k=1}^{n} p_k=1$ 且 $C_1 < C_2 < \cdots < C_n$，因此 C 的均值可记为 $\bar{C}=\sum_{k=1}^{n} p_k C_k$。

① 在第一种情况中，除了以下约束条件和定义式：

$$\begin{cases} \sum_{k=1}^{n} p_k = 1 \\ \sum_{k=1}^{n} p_k \mu_k = \bar{\mu} \\ \sum_{k=1}^{n} p_k (\bar{\mu} - \mu_k)^2 = \mathrm{Var}(\mu) \end{cases}$$

我们假定其他方程个数为 n-2 个，以保证在 Var(μ)给定的条件下，$p_j(1 \le j \le n)$和 μ_k这 n+1 个未知量能确定下来；在第二种情况中，除了以上的后面两个定义式之外，仍然假定其他方程个数为 n-2 个，以保证在 Var(μ)给定的条件下，$\mu_k(1 \le k \le n)$这 n 个未知量能确定下来。

② 由于 $\prod_{2\mu}^{in}$ (k+1=2)和 $\prod_{2\mu}^{in}$ (n-k+1=2)的差距并不明显，因此在此只考虑在 k+1 重价格策略下的服务员的利润率。对于后面将要讨论的 $\prod_{2C}^{in}$ (k =1)和 $\prod_{2R}^{in}$ (k =1)仍类似处理。

7.4.1 等待损耗信息封闭

给定价格 T，顾客的均衡到达率为 $\lambda = \mu - \bar{C}/(R-T)$。因此，服务员的最大利润率可由一阶最优条件 $\prod_C'(T) = \lambda + \lambda_T' T = 0$ 得到，即

$$\prod_C^{un} = \begin{cases} \left(\sqrt{R\mu} - \sqrt{\bar{C}}\right)^2, & 若 R\mu > \bar{C} \\ 0, & 若 R\mu \le \bar{C} \end{cases} \tag{7-21}$$

7.4.2 等待损耗信息公开

7.4.2.1 等待损耗信息公开 —— n 重价格

假定单位等待损耗的分布信息对顾客是公开的，且服务员可根据 C 的 n 个实现值实行 n 重定价。类似于式（7-5），可得到服务员利润率为

$$\prod_{nC}^{in} = \begin{cases} \sum_{k=1}^{n} p_k \left(\sqrt{R\mu} - \sqrt{C_k}\right)^2, & 若 R\mu > C_n \\ \sum_{k=1}^{j-1} p_k \left(\sqrt{R\mu} - \sqrt{C_k}\right)^2, & 若 C_{j-1} < R\mu \le C_j,\ 2 \le j \le n \\ 0, & 若 R\mu \le C_1 \end{cases} \tag{7-22}$$

7.4.2.2 等待损耗信息公开 —— 单一价格

对于 $T < R - C_n/\mu$，有 $\lambda_1, \lambda_2, \cdots, \lambda_n > 0$。给定价格 T，顾客的均衡到达率为 $\lambda_k = \mu - C_k/(R-T)$。因此，服务员的平均利润率 $\prod_C$ 为

$$\prod_C = T\sum_{k=1}^{n} p_k \lambda_k = T\sum_{k=1}^{n} p_k T\left(\mu - \sqrt{\frac{C_k}{R-T}}\right) = T\mu - \frac{\bar{C}T}{R-T} \tag{7-23}$$

这与信息封闭情况下的表达式相同，且平均利润率为

$$\prod_C^1 = \left(\sqrt{R\mu} - \sqrt{\bar{C}}\right)^2 \tag{7-24}$$

对于 $R - C_k/\mu \le T < R - C_{k-1}/\mu$ $(2 \le k \le n)$，有 $\lambda_1, \cdots, \lambda_{k-1} > 0$ 且 $\lambda_k, \cdots, \lambda_n = 0$。在此范围内，利润最大化价格为 $T_{kC}^2 = R - \sqrt{RC_{k-1}/\mu}$，利润率为

$$\prod_{kC}^2 = \left(R - \sqrt{\frac{RC_{k-1}}{\mu}}\right)\left(\sqrt{\mu}\sum_{j=1}^{k-1} p_j \left(\sqrt{\mu} - \frac{C_j}{\sqrt{RC_{k-1}}}\right)\right), 2 \le k \le n \tag{7-25}$$

比较式（7-24）和式（7-25）的大小，以确定在 R 和 μ 给定条件下的最大利润率。因此，

$$\prod_{1C}^{in}=\begin{cases}\prod_{C}^{1}, & 若R\mu>C_n\\ \prod_{kC}^{2}, & 若C_{k-1}<R\mu\le C_k,\ 2\le k\le n\\ 0, & 若R\mu\le C_1\end{cases} \tag{7-26}$$

当 $R=20$，$\mu=2$，$p_1=p_2=p_3=1/3$，$C_1=1$，$C_2=2$，$C_3=6$ 时，图 7-6（a）表明服务员值得通过降低价格的方式来吸引更多的顾客进入系统。在图 7-6（b）中，分别增大 C_2 和 C_3，使得 $C_2=4$，$C_3=16$，观察到服务员如果选择一个更高的价格将使得它的利润更大。但在此价格下，当 $C=C_3$ 时，将没有顾客进入系统。在图 7-6（c）中，继续增大 C_2 和 C_3，使得 $C_2=16$，$C_3=32$，服务员会选择一个比图 7-6（b）中还高的价格。但在此价格下，当 $C=C_2$ 和 $C=C_3$ 时，将没有顾客进入系统。

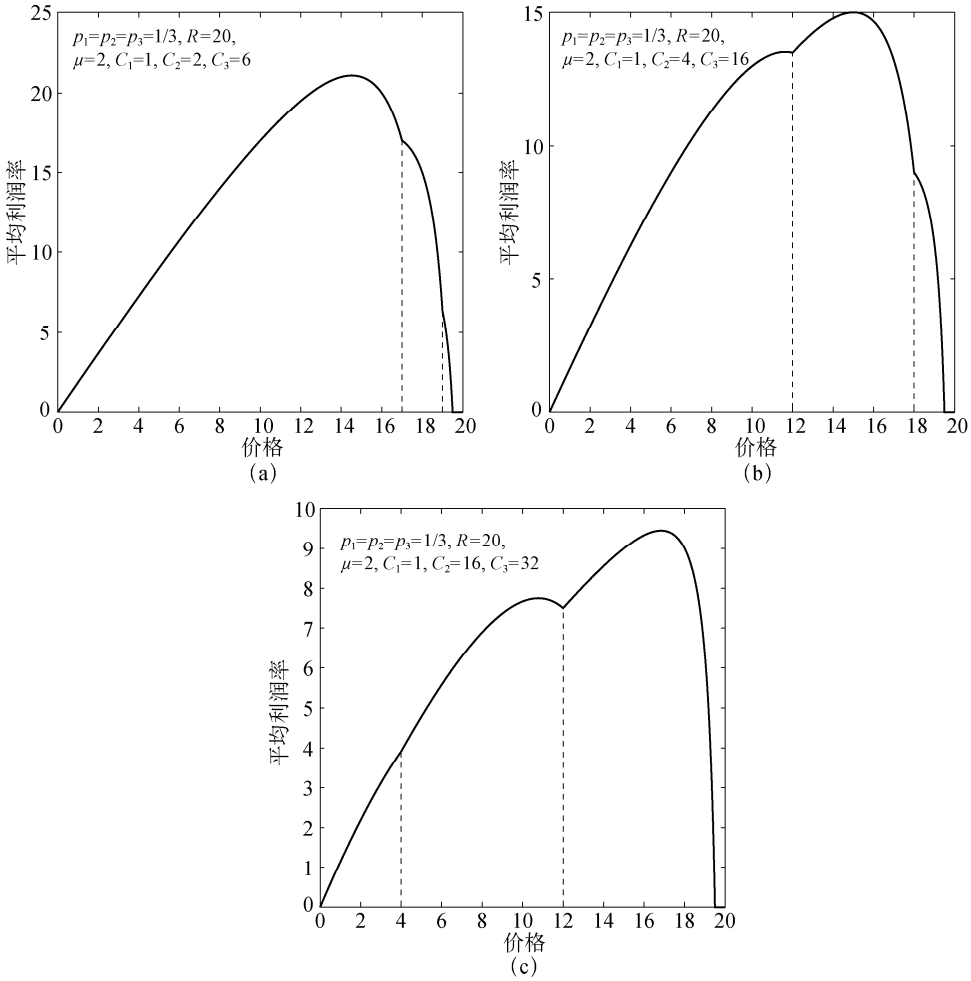

图 7-6　平均利润率与价格之间的变化关系

7.4.2.3 等待损耗信息公开——$k+1$ 重价格 对于 C 的前 k $(1\leq k\leq n-2)$个实现值, 根据信息公开且 n 重价格情况下的结果, 得到服务员利润率的第一部分, 记为 $\prod_{(k+1)C}^{1}$，为

$$\prod_{(k+1)C}^{1}=\begin{cases}\sum_{l=1}^{k}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2, & \text{若}R\mu>C_k\\ \sum_{l=1}^{j-1}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2, & \text{若}C_{j-1}<R\mu\leq C_j,\ 2\leq j\leq k\\ 0, & \text{若}R\mu>C_k\end{cases}\quad(7\text{-}27)$$

对于其他的 $n-k$ 个实现值，首先假定 $T<R-C_n/\mu$，因此 $\lambda_{k+1},\lambda_{k+2},\cdots,\lambda_n>0$。给定 $C=C_l$ $(k+1\leq l\leq n)$，顾客到达率为 $\lambda_l=\mu-C_l/(R-T)$。因此，在给定 T 的情况下，顾客的平均到达率为

$$\lambda=\sum_{l=k+1}^{n}p_l\lambda_l=\sum_{l=k+1}^{n}p_l\left(\mu-\frac{C_l}{R-T}\right)\quad(7\text{-}28)$$

因为平均利润率为 $\prod_C^2(T)=\lambda T$，由一阶最优条件 $\prod_C^{2\prime}(T)=0$ 得到利润最大化价格为

$$T_C^2=R-\sqrt{\frac{\sum_{l=k+1}^{n}p_lC_lR}{\sum_{l=k+1}^{n}p_l\mu}}$$

以及相应的到达率为

$$\lambda_C^2=\sum_{l=k+1}^{n}p_l\left(\mu-\frac{C_l\sqrt{\sum_{l=k+1}^{n}p_l\mu}}{\sum_{l=k+1}^{n}p_lC_lR}\right)$$

因此，利润率的第二部分，记为 $\prod_C^2$，为

$$\prod_C^2=\left(\sqrt{R\mu\sum_{l=k+1}^{n}p_l}-\sqrt{\sum_{l=k+1}^{n}p_lC_l}\right)^2\quad(7\text{-}29)$$

假定 $R-C_j/\mu\leq T<R-C_{j-1}/\mu(k+2\leq j\leq n)$，因此 $\lambda_{k+1},\lambda_{k+2},\cdots,\lambda_{j-1}>0$ 且 $\lambda_j,\lambda_{j+1},\cdots,\lambda_n=0$。当 $C=C_{j-1}$ 时，服务员将选择利润最大化价格 T_{jC}^2，且 $T_{jC}^2=R-\sqrt{RC_{j-1}/\mu}$。因此，利润率的第二部分，记为 $\prod_{jC}^2$，其为

$$\prod_{jC}^2=\left(R-\sqrt{\frac{RC_{j-1}}{\mu}}\right)\sqrt{\mu}\sum_{l=k+1}^{j-1}p_l\left(\sqrt{\mu}-\frac{C_l}{\sqrt{RC_{j-1}}}\right),\ k+2\leq j\leq n\quad(7\text{-}30)$$

从而，总平均利润率 $\prod_{(k+1)C}^{in}$ 归纳如下：

$$\prod_{(k+1)C}^{in}=\begin{cases}\sum_{l=1}^{k}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2+\prod_{C}^{2}, & \text{若}R\mu>C_n\\ \sum_{l=1}^{k}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2+\prod_{jC}^{2}, & \text{若}C_{j-1}<R\mu\le C_j,\ k+2\le j\le n\\ \sum_{l=1}^{i-1}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2, & \text{若}C_{i-1}<R\mu\le C_i,\ 2\le i\le k+1\\ 0, & \text{若}R\mu\le C_1\end{cases}\tag{7-31}$$

假定等待损耗的所有实现值等概率出现，图 7-7 比较了 n = 3,4,5 时 $\prod_{(k+1)C}^{in}$ ($1\le k\le n$-2)的大小。类似于图 7-2，图 7-7 表明 $\prod_{(k+1)C}^{in}$ 也随 n 递减。这说明顾客的剩余效用随 n 递减，从而直接影响到了服务员的最大利润。而且，对于 n 的同一取值，$\prod_{(k+1)C}^{in}$ 仍随 k 递增。

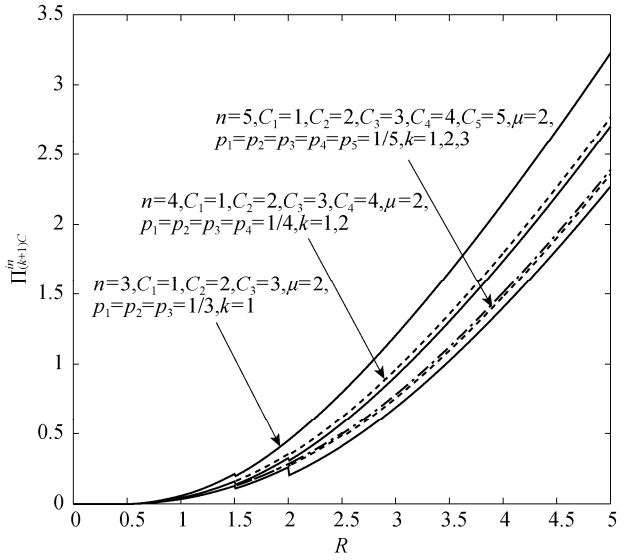

图 7-7　$\prod_{(k+1)C}^{in}$ ($1\le k\le n$-2，n = 3,4,5)

7.4.2.4　等待损耗信息公开 —— n-k+1 重价格

对于 C 的后 n-k ($2\le k\le n$-1)个实现值, 根据信息公开且 n 重价格情况下的结果，可得到服务员利润率的第二部分，记为 $\prod_{(n-k+1)C}^{2}$，其为

$$\prod_{(n-k+1)C}^{2}=\begin{cases}\sum_{l=k+1}^{n}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2, & \text{若}R\mu>C_n\\ \sum_{l=k+1}^{j-1}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2, & \text{若}C_{j-1}<R\mu\le C_j,\ k+2\le j\le n\\ 0, & \text{若}R\mu\le C_{k+1}\end{cases}\quad(7\text{-}32)$$

对于 C 的前 k 个实现值，首先假定 $T<R-C_k/\mu$，因此 $\lambda_1,\lambda_2,\cdots,\lambda_k>0$。给定 $C=C_l\,(1\le l\le k)$，顾客到达率为 $\lambda_l=\mu-C_l/(R-T)$。给定 T，顾客的平均到达率为

$$\lambda=\sum_{l=1}^{k}p_l\lambda_l=\sum_{l=1}^{k}p_l\left(\mu-\frac{C_l}{R-T}\right)\quad(7\text{-}33)$$

因为平均利润率为 $\prod_C^1(T)=\lambda T$，由一阶最优条件 $\prod_C^{1\prime}(T)=0$ 得到利润最大化价格为

$$T_{(n-k+1)C}^{1}=R-\sqrt{\frac{\sum_{l=1}^{k}p_lC_lR}{\sum_{l=1}^{k}p_l\mu}}$$

以及相应的到达率为

$$\lambda_{(n-k+1)C}^{1}=\sum_{l=1}^{k}p_l\left(\mu-\frac{C_l\sqrt{\sum_{l=1}^{k}p_l\mu}}{\sum_{l=1}^{k}p_lC_lR}\right)$$

因此，利润率的第一部分为

$$\prod_{(n-k+1)C}^{1}=\left(\sqrt{R\mu\sum_{l=1}^{k}p_l}-\sqrt{\sum_{l=1}^{k}p_lC_l}\right)^2\quad(7\text{-}34)$$

假定 $R-C_j/\mu\le T<R-C_{j-1}/\mu\,(2\le j\le n)$，因此 $\lambda_1,\lambda_2,\cdots,\lambda_{j-1}>0$ 且 $\lambda_j,\lambda_{j+1},\cdots,\lambda_k=0$。当 $C=C_{j-1}$ 时，服务员将选择利润最大化价格 T_{jC}^1，且 $T_{jC}^1=R-\sqrt{RC_{j-1}/\mu}$。因此，利润率的第一部分，记为 $\prod_{jC}^1$，其为

$$\prod_{jC}^{1}=\left(R-\sqrt{\frac{RC_{j-1}}{\mu}}\right)\sqrt{\mu}\sum_{l=1}^{j-1}p_l\left(\sqrt{\mu}-\frac{C_l}{\sqrt{RC_{j-1}}}\right),\ 2\le j\le k\quad(7\text{-}35)$$

从而，总平均利润率 $\prod_{(n-k+1)C}^{in}$ 归纳如下：

$$\prod_{(n-k+1)C}^{in}=\begin{cases}\sum_{l=k+1}^{n}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2+\prod_{(n-k+1)C}^{1}, & \text{若}R\mu>C_n\\ \sum_{l=k+1}^{j-1}p_l\left(\sqrt{R\mu}-\sqrt{C_l}\right)^2+\prod_{(n-k+1)C}^{1}, & \text{若}C_{j-1}<R\mu\le C_j,\ k+2\le j\le n\\ \prod_{(n-k+1)C}^{1}, & \text{若}C_k<R\mu\le C_{k+1}\\ \prod_{iC}^{1}, & \text{若}C_{i-1}<R\mu\le C_i,\ 2\le i\le k\\ 0, & \text{若}R\mu\le C_1\end{cases}\tag{7-36}$$

假定等待损耗的所有实现值等概率出现, 图 7-8 比较了 $n=3, 4, 5$ 时 $\prod_{(n-k+1)C}^{in}$ $(2\le k\le n-1)$的大小。与图 7-7 不同的是，图 7-8 表明 $\prod_{(n-k+1)C}^{in}$ 随 k 递减。而与图 7-7 类似的是，$\prod_{(n-k+1)C}^{in}$ 仍随 n 递减。

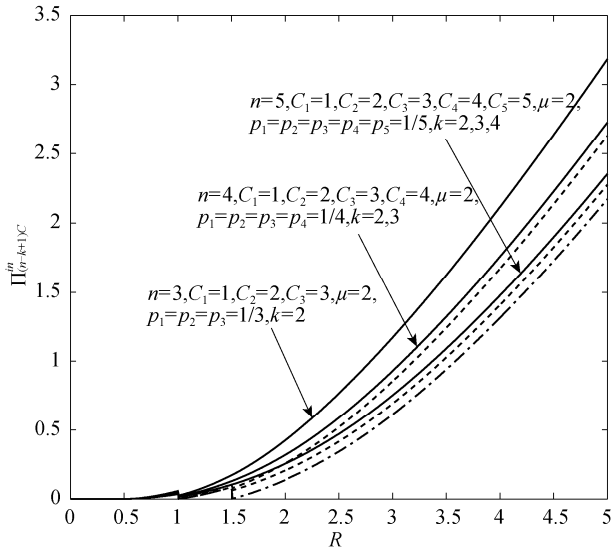

图 7-8　$\prod_{(n-k+1)C}^{in}$ $(2\le k\le n-1,\ n=3,4,5)$

图 7-9 比较了 $n=3,4,5$ 情况下其他三类利润率。可见这三类利润率都随 R 递增，且对于每个 n，$\prod_{nC}^{in}\ge\prod_{1C}^{in}\ge\prod_{C}^{un}$。然而，利润率均随 n 递减，且 $\prod_{nC}^{in}$ 与 $\prod_{1C}^{in}$ 或 $\prod_{C}^{un}$ 之间的差距随 n 递增。因此，类似于图 7-4，这说明信息量对服务员利润的影响是随着等待损耗可能实现值个数的增多而不断增大的。

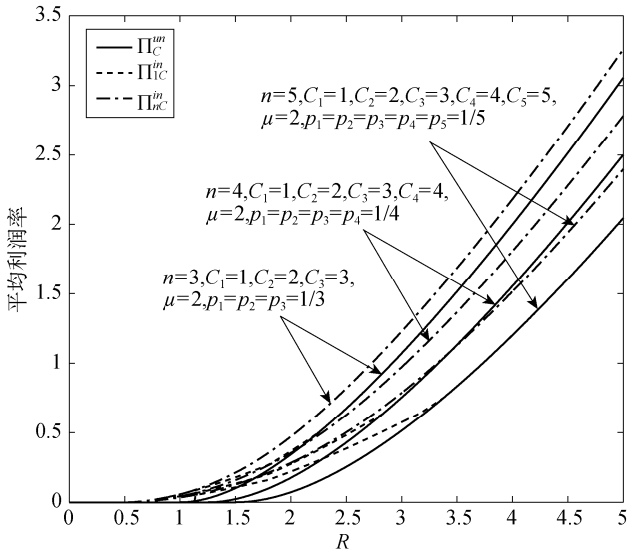

图 7-9　$\Pi_C^{un}, \Pi_{1C}^{in}, \Pi_{nC}^{in}$　(n = 3,4,5)

7.4.3　可变等待损耗对利润的影响

下面，考虑等待损耗的不确定性对服务员利润的影响。考虑不确定性下的两种情况，且在两种情况下固定 $\bar{C} = \sum_{k=1}^{n} p_k C_k$。

- 第一种情况下，除了 C_k 之外，固定所有的 $C_j (1 \leq j,\ k \leq n,\ j \neq k)$。假定存在一系列关于 $p_j (1 \leq j \leq n)$和 C_k 的方程 $\{f_j(p_1, p_2, \cdots, p_n, C_k) = 0, 1 \leq j \leq n\text{-}2\}$，通过同时变换 $p_j (1 \leq j \leq n)$和 C_k，$\bar{C}$ 保持不变而 Var(C)发生变化。在图 7-10（a）中，假定 $n = 3$，$R = 20$，$\mu = 2$，$\bar{C} = 2$，$C_1 = 1$，$C_2 = 2$ 以及 $p_1 = 2p_2$。
- 第二种情况下，固定所有满足 $0 \leq p_k \leq 1$ 和 $\sum_{k=1}^{n} p_k = 1$的概率 $p_k (1 \leq k \leq n)$，且假定存在一系列关于 $C_k (1 \leq k \leq n)$的方程 $\{f_j(C_1, C_2, \cdots, C_n) = 0, 1 \leq j \leq n\text{-}2\}$。通过同时变换 $C_k (1 \leq k \leq n)$，$\bar{C}$ 保持不变而 Var(C)发生变化。在图 7-10（b）中，假定 $n = 3$，$R = 20$，$\mu = 2$，$\bar{C} = 2$，$p_1 = 2/3$，$p_2 = p_3 = 1/6$ 以及 $C_2 = 2C_1$。

从图 7-10 中可见，由于 Π_C^{un} 与 Var(C)无关，Π_C^{un} 不随 Var(C)的增大而变化。然而，$\Pi_{2C}^{in} (k = 1)$ 和 Π_{3C}^{in} 始终随 Var(C)递增，Π_{1C}^{in} 先不变后递增。而且，在第一种情况下，Π_{1C}^{in} 几乎平行于 Π_{3C}^{in}，而在第二种情况下，Π_{1C}^{in} 非常接近于 Π_{3C}^{in}。

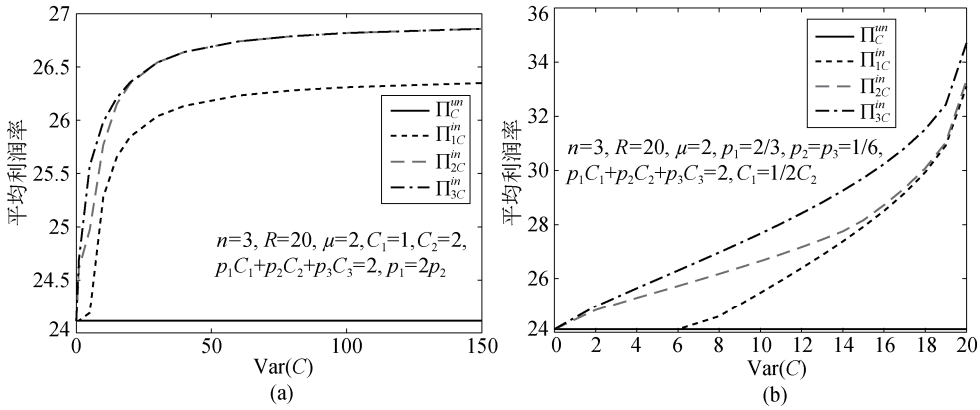

图 7-10　等待损耗的不确定性对服务员利润的影响

7.5　可变服务收益

假定服务收益的分布律为 $P\{R=R_k\}=p_k(1\le k\le n)$。$\sum_{k=1}^{n}p_k=1$ 且 $R_1>R_2>\cdots>R_n$。记 R 的均值为 $\overline{R}=\sum_{k=1}^{n}p_kR_k$。

7.5.1　服务收益信息封闭

给定 T，顾客的均衡到达率为 $\lambda=\mu-C/\left(\overline{R}-T\right)$。因此，服务员的最大利润率为

$$\Pi_R^{un}=\begin{cases}\left(\sqrt{\overline{R}\mu}-\sqrt{C}\right)^2, & 若\overline{R}\mu>C\\ 0, & 若\overline{R}\mu\le C\end{cases}\tag{7-37}$$

7.5.2　服务收益信息公开

7.5.2.1　服务收益信息公开 —— n 重价格

定单位等待损耗的分布信息对顾客是公开的，且服务员可根据 R 的 n 个实现值实行 n 重定价。因此，可得到

$$\Pi_{nR}^{in}=\begin{cases}\displaystyle\sum_{k=1}^{n}p_k\left(\sqrt{R_k\mu}-\sqrt{C}\right)^2, 若R_n>\frac{C}{\mu}\\ \displaystyle\sum_{k=1}^{j-1}p_k\left(\sqrt{R_k\mu}-\sqrt{C}\right)^2, 若R_j\le\frac{C}{\mu}<R_{j-1},2\le j\le n\\ 0, \qquad 若R_1\le\frac{C}{\mu}\end{cases}\tag{7-38}$$

7.5.2.2 服务收益信息公开——单一价格

对于 $T<R_n-C/\mu$，有 $\lambda_1,\lambda_2,\cdots,\lambda_n>0$。给定 T，顾客的均衡到达率为 $\lambda_k=\mu-C/(R_k-T)$。因此服务员的平均利润率 $\prod_R$ 为

$$\prod_R=T\sum_{k=1}^{n}p_k\lambda_k=T\sum_{k=1}^{n}p_k\left(\mu-\frac{C}{R_k-T}\right)=T\mu-CT\sum_{k=1}^{n}\frac{p_k}{R_k-T} \tag{7-39}$$

一阶最优条件为

$$\mu-C\sum_{k=1}^{n}\frac{p_k}{R_k-T}-CT\sum_{k=1}^{n}\frac{p_k}{(R_k-T)^2}=0 \tag{7-40}$$

求解该方程，得到利润最大化价格 T_R^1 以及平均利润率为 $\prod_R^1=\lambda_R^1T_R^1$，其中 $\lambda_R^1=\sum_{k=1}^{n}p_k\left(\mu-C/\left(R_k-T_R^1\right)\right)$。

对于 $R_k-C/\mu\le T<R_{k-1}-C/\mu\,(2\le k\le n)$，有 $\lambda_1,\cdots,\lambda_{k-1}>0$ 且 $\lambda_k,\cdots,\lambda_n=0$。在此范围内，利润最大化价格为 $T_{kR}^2=R_{k-1}-\sqrt{R_{k-1}C/\mu}$，利润率为

$$\prod_{kR}^2=\left(R_{k-1}-\sqrt{\frac{R_{k-1}C}{\mu}}\right)\left(\sum_{j=1}^{k-1}p_j\left(\mu-\frac{C}{R_j-\left(R_{k-1}-\sqrt{\frac{R_{k-1}C}{\mu}}\right)}\right)\right),\ 2\le k\le n \tag{7-41}$$

比较式（7-39）和式（7-41），以确定C和μ给定条件下利润率的最大可能值。因此，

$$\prod_{1R}^{in}=\begin{cases}\prod_R^1, & 若R_n>\dfrac{C}{\mu}\\ \prod_{kR}^2, & 若R_k<\dfrac{C}{\mu}\le R_{k-1},\ 2\le k\le n\\ 0, & 若R_1\le\dfrac{C}{\mu}\end{cases} \tag{7-42}$$

当 $C=1,\mu=2$，$p_1=p_2=p_3=1/3$，R_1=20，R_2=18，R_3=16时，图 7-11（a）表明服务员值得降价以吸引顾客。在图 7-11（b）中，增大 R_1 和 R_2，使 $R_1=30$，$R_2=20$，减小 R_3，使 R_3=10，此时若选择一个更高的价格将使其利润更大。但当 $R=R_3$ 时，将无顾客进入系统。在图 7-11（c）中，继续增大 R_1，使 $R_1=60$，服务员会选择一个比图 7-11（b）中还高的价格。但当 R=R_2 或 R=R_3 时，将无顾客进入系统。

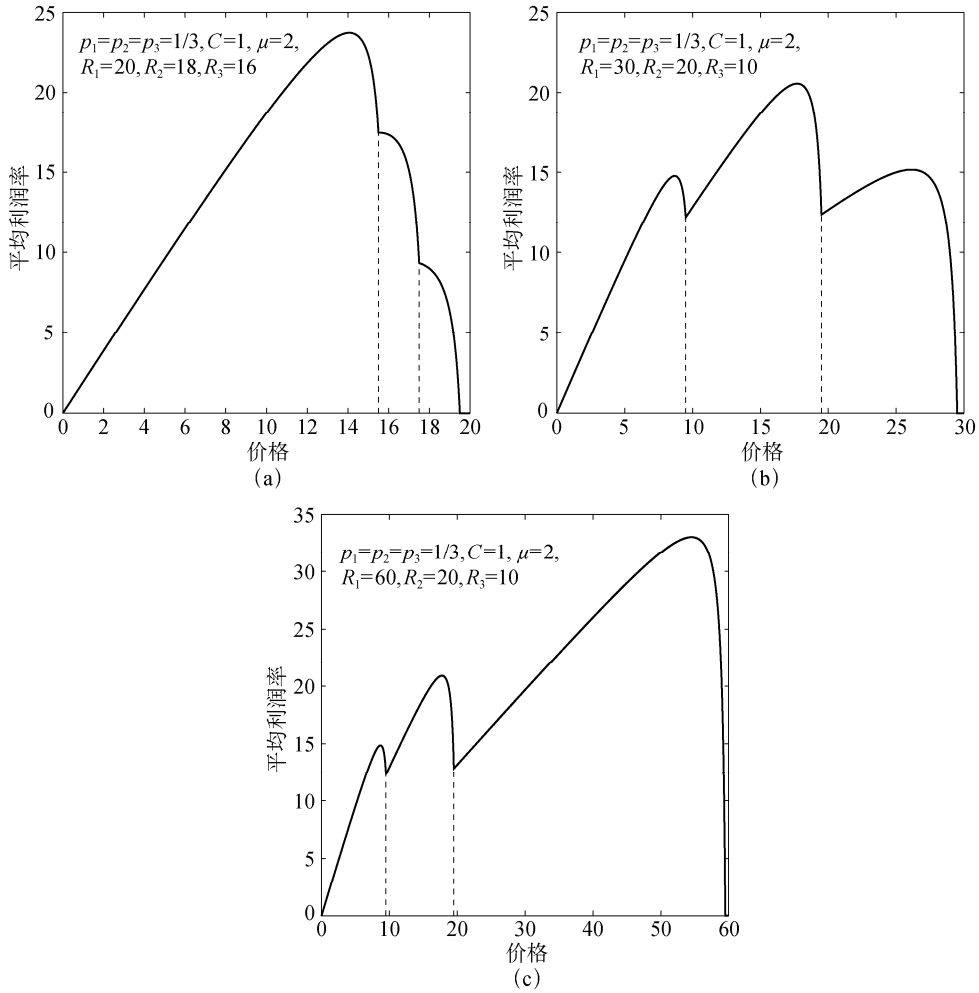

图 7-11　平均利润率与价格之间的变化关系

7.5.2.3　服务收益信息公开 —— k+1 重价格

对于 R 的前 $k(1\le k\le n-2)$ 个实现值，可以得到服务员利润率的第一部分，记为 $\prod_{(k+1)R}^{1}$，其为

$$\prod_{(k+1)R}^{1}=\begin{cases}\sum_{l=1}^{k}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2, & 若R_k>\dfrac{C}{\mu}\\ \sum_{l=1}^{j-1}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2, & 若R_j<\dfrac{C}{\mu}\le R_{j-1},\ 2\le j\le k\\ 0, & 若R_1\le\dfrac{C}{\mu}\end{cases}\tag{7-43}$$

对于其他的 $n-k$ 个实现值，首先假定 $T<R_n-C/\mu$，因此 $\lambda_{k+1},\lambda_{k+2},\cdots,\lambda_n>0$。给定 $R=R_l(k+1\le l\le n)$，顾客到达率为 $\lambda_l=\mu-C/(R_l-T)$。因此，平均利润率为

$$\Pi_R^2(T)=T\sum_{l=k+1}^{n}p_l\lambda_l=T\sum_{L=k+1}^{n}P_l\left(\mu-\frac{C}{R_l-T}\right)=T\mu\sum_{l=k+1}^{n}p_l-CT\sum_{l=k+1}^{n}\frac{p_l}{R_l-T} \tag{7-44}$$

一阶最优条件为

$$\mu\sum_{l=k+1}^{n}p_l-C\sum_{l=k+1}^{n}\frac{p_l}{R_l-T}-CT\sum_{l=k+1}^{n}\frac{p_l}{\left(R_l-T\right)^2}=0 \tag{7-45}$$

求解该方程，得到利润最大化价格 T_R^2 以及平均利润率的第二部分为 $\Pi_R^2=\lambda_R^2T_R^2$，其中 $\lambda_R^2=\sum_{l=k+1}^{n}p_l\left(\mu-C/\left(R_l-T_R^2\right)\right)$。

然后假定 $R_j\text{-}C/\mu\le T<R_{j\text{-}1}\text{-}C/\mu$ (k+2$\le j\le n$)，因此 $\lambda_{k+1},\lambda_{k+2},\cdots,\lambda_{j-1}>0$ 且 $\lambda_j,\lambda_{j+1},\cdots,\lambda_n=0$。当 $R=R_{j-1}$ 时，服务员选择利润最大化价格 T_{jR}^2，且 $T_{jR}^2=R_{j-1}-\sqrt{R_{j-1}C/\mu}$。因此，利润率的第二部分，记为 Π_{jR}^2，其为

$$\Pi_{jR}^2=\left(R_{j-1}-\sqrt{\frac{R_{j-1}C}{\mu}}\right)\sum_{l=k+1}^{j-1}p_l\left(\mu-\frac{C}{R_l-\left(R_{j-1}-\sqrt{\frac{R_{j-1}C}{\mu}}\right)}\right),\ k+2\le j\le n \tag{7-46}$$

从而，总平均利润率 $\Pi_{(k+1)R}^{in}$ 归纳如下

$$\Pi_{(k+1)R}^{in}=\begin{cases}\sum_{l=1}^{k}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2+\Pi_R^2, & 若R_n>\dfrac{C}{\mu}\\ \sum_{l=1}^{k}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2+\Pi_{jR}^2, & 若R_j<\dfrac{C}{\mu}\le R_{j-1},\ k+2\le j\le n\\ \sum_{l=1}^{i-1}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2, & 若R_i<\dfrac{C}{\mu}\le R_{i-1},\ 2\le i\le k+1\\ 0, & 若R_1\le\dfrac{C}{\mu}\end{cases} \tag{7-47}$$

假定服务收益的所有实现值等概率出现，图 7-12 比较了 $n=3,4,5$ 时 $\Pi_{(k+1)R}^{in}(1\le k\le n-2)$ 的大小。不同于图 7-2 与图 7-7，图 7-12 表明 $\Pi_{(k+1)R}^{in}$ 不仅随 C 递减，而且随 n 递减。对于 n 的同一取值，$\Pi_{(k+1)R}^{in}$ 仍随 k 递增。

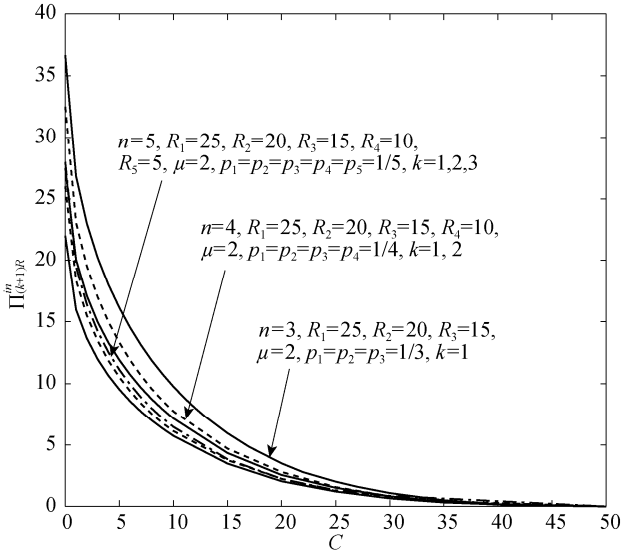

图 7-12　$\Pi^{in}_{(k+1)R}(1\le k\le n-2, n=3, 4, 5)$

7.5.2.4　服务收益信息公开——n-k+1 重价格

对于 R 的后 $n-k(2\le k\le n-1)$ 个实现值，可以得到服务员利润率的第二部分，记为 $\Pi^{2}_{(n-k+1)R}$，其为

$$\Pi^{2}_{(n-k+1)R}=\begin{cases}\displaystyle\sum_{l=k+1}^{n}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2, & 若R_n>\dfrac{C}{\mu}\\ \displaystyle\sum_{l=k+1}^{j-1}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2, & 若R_j<\dfrac{C}{\mu}\le R_{j-1}, k+2\le j\le n\\ 0, & 若R_{k+1}\le\dfrac{C}{\mu}\end{cases}\quad(7\text{-}48)$$

对于 R 的前 k 个实现值，首先假定 $T<R_k\text{-}C/\mu$，因此 $\lambda_1, \lambda_2,\cdots,\lambda_k>0$。给定 $R=R_l(1\le l\le k)$，顾客到达率为 $\lambda_l=\mu-C/(R_l-T)$。因此，给定价格 T，平均利润率为

$$\Pi^{1}_{R}(T)=T\sum_{l=1}^{k}p_l\lambda_l=T\sum_{l=1}^{k}p_l\left(\mu-\frac{C}{R_l-T}\right)=T\mu\sum_{l=1}^{k}p_l-CT\sum_{l=1}^{k}\frac{p_l}{R_l-T}\quad(7\text{-}49)$$

一阶最优条件为

$$\mu\sum_{l=1}^{k}p_l-C\sum_{l=1}^{k}\frac{p_l}{R_l-T}-CT\sum_{l=1}^{k}\frac{p_l}{(R_l-T)^2}=0\quad(7\text{-}50)$$

求解该方程，得到利润最大化价格 $T^{1}_{(n-k+1)R}$，并由此得到平均利润率的第一部分

$\prod_{(n-k+1)R}^{1}=\lambda_{(n-k+1)R}^{1}T_{(n-k+1)R}^{1}$，其中 $\lambda_{(n-k+1)R}^{1}=\sum_{l=1}^{k}p_l\left(\mu-C/\left(R_l-T_{(n-k+1)R}^{1}\right)\right)$。

然后假定 $R_j-C/\mu\leq T<R_{j-1}-C/\mu\,(2\leq j\leq k)$，因此 $\lambda_1,\lambda_2,\cdots,\lambda_{j-1}>0$ 且 $\lambda_j,\lambda_{j+1},\cdots,\lambda_k=0$。当 $R=R_{j-1}$ 时，服务员将选择利润最大化价格 T_{jR}^{1}，且 $T_{jR}^{1}=R_{j-1}-\sqrt{R_{j-1}C/\mu}$。因此，利润率的第一部分，记为 $\prod_{jR}^{1}$，其为

$$\prod_{jR}^{1}=\left(R_{j-1}-\sqrt{\frac{R_{j-1}C}{\mu}}\right)\sum_{l=1}^{j-1}p_l\left(\mu-\frac{C}{R_l-\left(R_{j-1}-\sqrt{\frac{R_{j-1}C}{\mu}}\right)}\right),\ 2\leq j\leq k \quad (7\text{-}51)$$

从而，总平均利润率 $\prod_{(n-k+1)R}^{in}$ 归纳如下：

$$\prod_{(n-k+1)R}^{in}=\begin{cases}\sum_{l=k+1}^{n}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2+\prod_{(n-k+1)R}^{1}, & 若R_n>\dfrac{C}{\mu}\\ \sum_{l=k+1}^{j-1}p_l\left(\sqrt{R_l\mu}-\sqrt{C}\right)^2+\prod_{(n-k+1)R}^{1}, & 若R_j<\dfrac{C}{\mu}\leq R_{j-1},\ k+2\leq j\leq n\\ \prod_{(n-k+1)R}^{1}, & 若R_{k+1}<\dfrac{C}{\mu}\leq R_k\\ \prod_{iR}^{1}, & 若R_i<\dfrac{C}{\mu}\leq R_{i-1},2\leq i\leq k\\ 0, & 若R_1\leq\dfrac{C}{\mu}\end{cases} \quad (7\text{-}52)$$

假定服务收益所有实现值等概率出现，图 7-13 比较了 $n=3,4,5$ 时 $\prod_{(n-k+1)R}^{in}$ $(2\leq k\leq n-1)$ 的大小。不同于图 7-3 与图 7-8，图 7-13 表明 $\prod_{(n-k+1)R}^{in}$ 随 C 和 n 均递减。而和图 7-3 和图 7-8 类似，对于 n 的同一取值，$\prod_{(n-k+1)R}^{in}$ 仍随 k 递减。

图 7-14 比较了 $n=3,4,5$ 情况下其他三类利润率。可见这三类利润率都随 C 递减。对于每个 n，$\prod_{1R}^{in}$ 先大于后小于 $\prod_{R}^{un}$，最后等于 $\prod_{nR}^{in}$。然而，利润率均随 n 递减。

7.5.3 可变服务收益对利润的影响

下面考虑服务收益的不确定性对服务员利润的影响。考虑不确定性下的两种情况，且在两种情况下固定 $\bar{R}=\sum_{k=1}^{n}p_kR_k$。

- 第一种情况下，除了 R_k 之外，固定所有的 $R_j(1\leq j,\ k\leq n,\ j\neq k)$。假定存在一系列关于 $p_j(1\leq j\leq n)$ 和 R_k 的方程 $\{f_j(p_1,p_2,\cdots,p_n,R_k)=0,1\leq j\leq n-2\}$。通过同时变换 $p_j(1\leq j\leq n)$ 和 R_k，$\bar{R}$ 保持不变而 Var (R) 发生变化。在图 7-15（a）中，

假定$n=3,\ C=1,\ \mu=2,\ \bar{R}=20,\ R_1=30,\ R_2=20$以及$p_1=2p_2$。

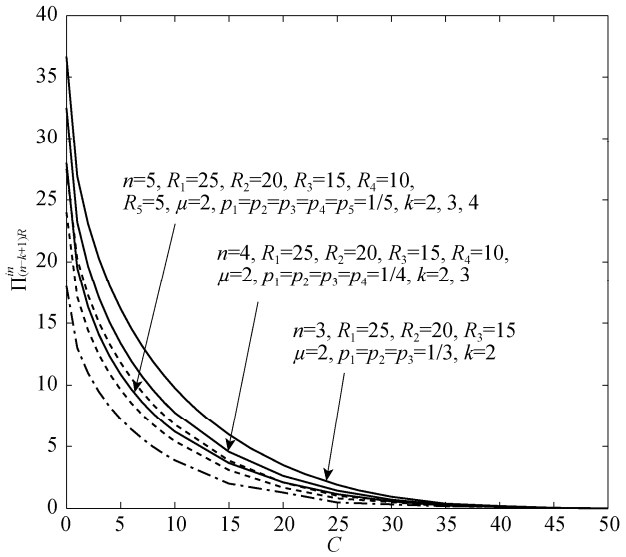

图 7-13　$\Pi^{in}_{(n-k+1)R}(2\le k\le n-1,\ n=3,4,5)$

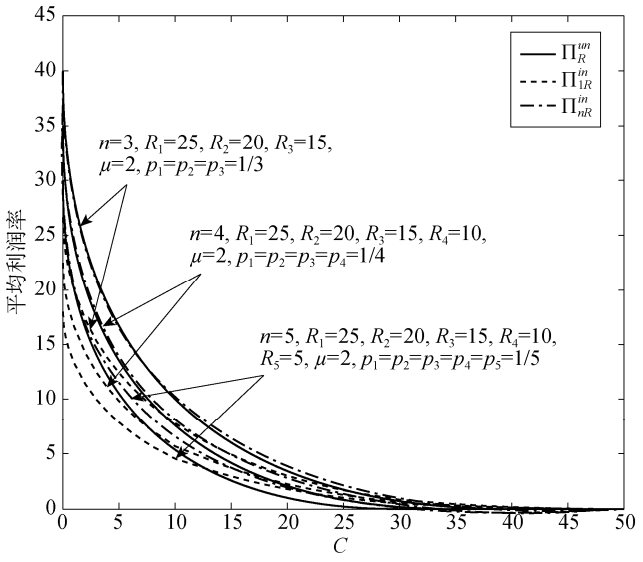

图 7-14　Π^{un}_{R}, Π^{in}_{1R}, $\Pi^{in}_{nR}(n=3,4,5)$

- 第二种情况下，固定所有满足$0\le p_k\le 1$和$\sum_{k=1}^{n}p_k=1$的概率$p_k$$(1\le k\le n)$，且假定存在一系列关于$R_k(1\le k\le n)$的方程$\{f_j(R_1,R_2,\cdots,R_n)=0,1\le j\le n-2\}$。通过同时变换$R_k(1\le k\le n)$，$\bar{R}$保持不变而Var (R)发生变化。

在图 7-15（b）中，假定 $n=3,\ C=1,\ \mu=2,\ \overline{R}=20,\ p_1=2/3,\ p_2=p_3=1/6$ 以及 $R_1=2R_2$。

从图 7-15 中，观察到 $\prod_R^{un}$ 保持不变而 $\prod_{3R}^{in}$ 始终递增。然而，$\prod_{1R}^{in}$ 先减后增且小于 $\prod_R^{un}$，$\prod_{2R}^{in}(k=1)$ 也是先减后增最后等于 $\prod_{3R}^{in}$。有趣的是，图 7-15（a）表明 $\prod_{2R}^{in}(k=1)$ 的变化趋势为：锐减 → 缓增 → 激增。而且，随着 Var (R) 的增大，$\prod_{2R}^{in}(k=1)$ 最先为最小的利润率，而后顺次超过 $\prod_{1R}^{in}$ 和 $\prod_R^{un}$，最终成为与 $\prod_{3R}^{in}$ 相等的最大的利润率。

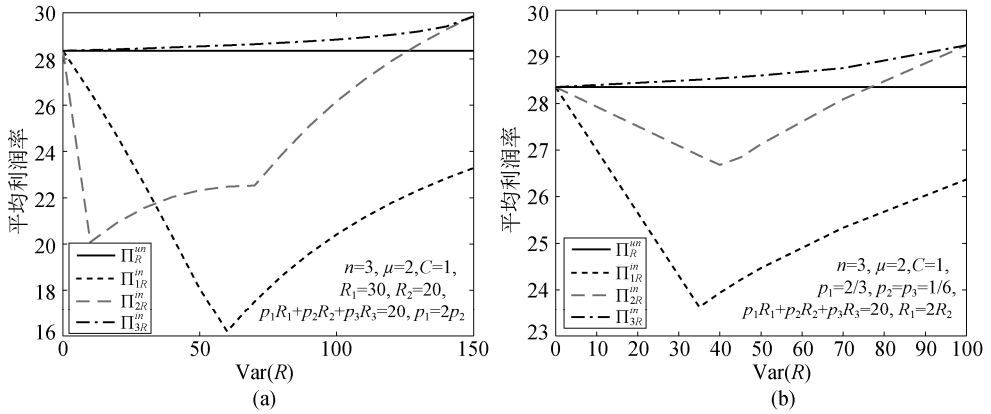

图 7-15　服务收益的不确定性对服务员利润的影响

7.6　比较结果总结

通过以上数值比较，总结出参数的不确定性对服务员利润的主要影响如下。

（1）对于固定的 n（$n>2$），服务员无论在何种情况下，都更加愿意公开服务率和单位等待损耗的分布信息，但只有当单位等待损耗大于某个阈值时，服务员才愿意公开服务收益的分布信息。

（2）对于固定的 n（$n>2$），在分布信息公开且采用混合价格策略的条件下，服务员的利润率对 k 的取值并不敏感。因此，服务员不必为了 k 值的选取而过于费心。

（3）对于固定的 n（$n>2$），在分布信息公开且采用单一价格策略的条件下，如果可变服务参数的实现值之间差距不大，服务员愿意以低价格来吸引更多顾客，否则，服务员宁愿设定较高价格。

（4）对于变化的 n（$n>2$），当可变服务参数的分布律为均匀分布或递增分布时，n 值的增大对服务员利润具有负作用。因此，如果 n 为服务员的内生量，

其将设法减少可变服务参数的可能实现值个数。

（5）对于变化的 n（$n>2$），无论对于哪类价格策略，随着 n 的增大，信息公开相比信息封闭来说，应该都是服务员更好的选择。

7.7　相关文献评述

对于系统信息的价值，目前为止已有不少的文献对其进行了研究，这里只列举一些最具代表性的成果。Hassin[33]发现阻止服务员封闭队长信息不一定是社会最优的，但是，如果服务员愿意公开队长信息，而管理者却鼓励其封闭信息一定不是最优的。Guo 和 Zipkin[37]讨论了三类系统信息水平，证明了信息精度越高，社会福利越大。随后，Guo 和 Zipkin[102]将上述模型一般化，采用了位相型服务时间分布。Guo 等[95]分析了具有部分服务时间分布信息的排队系统中顾客的均衡和最优止步行为。关于系统信息的价值，读者还可以参考其他相关参考文献[24,100,103–106]。

对于系统信息对主体行为的影响，Arnott 等[107,108]分析了在具有服务率和需求波动的排队系统中，系统信息对顾客止步行为的影响，且发现在合理的价格管控下，信息精度的提高有助于增加社会福利，否则将适得其反。Guo 和 Zipkin[109]研究了系统信息对顾客和服务员的影响，发现其对服务员利润的影响取决于顾客对延迟的敏感度，且较高精度的信息一般情况下对顾客是有利的。关于系统信息对主体行为的影响，读者还可以参考其他相关参考文献[35,43,59,64,110]。

除此之外，对于服务参数的不确定性对主体行为的影响，Hassin[111]以及孙微、李世勇[101]分别研究了三类系统参数服从两点分布和服从一般离散时间分布的具有多层次系统信息及多类定价策略的队长不可视排队系统。他们主要分析了服务参数的不确定性对顾客止步行为和服务员定价策略选择的影响以及服务员利润关于服务参数的不确定性的灵敏度。

参 考 文 献

[1] A. Erlang. The theory of probability and telephone conversations. Nyt Tidsskrift Matematik B, 1909, 20:33–39

[2] D. Kendall. Stochastic processes occurring in the theory of queues and their analysis of by the method of imbedded Markov chains. Annals of Mathematical Statistics, 1953, 24:207–212

[3] T. Meisling. Discrete time queueing theory. Operations Research, 1958, 6(1):96–105

[4] J. Hunter. Mathematical techniques of applied probability. New York: Academic Press, 1983

[5] 田乃硕, 徐秀丽, 马占友. 离散时间排队论. 北京: 科学出版社, 2007

[6] Y. Levy, U. Yechiali. Utilization of idle time in an M/G/1 queueing system. Management Science,1975, 22:202–211

[7] H. Takagi. Queueing analysis, vol.1 vacation and priority systems. Amsterdam: North-Holland, 1991, 1–203

[8] N. Tian, G. Zhang. Vacation queueing models—theory and applications. New York: Springer Verlag, 2006

[9] D. Servi, G. Finn. M/M/1 queue with working vacations (M/M/1/WV). Performance Evalua tion, 2002, 50(1):41–52

[10] N. Tian, J. Li, Z. Zhang. Matrix analysis method and working vacation queues-A survey. International Journal of Information and Management Sciences, 2009, 20(4):603–633

[11] 李继红. 工作休假排队系统的稳态理论及应用. 秦皇岛: 燕山大学, 2009

[12] G. Owen. Game Theory. New York: Academic Press, 1982

[13] 保罗·萨谬尔森等著. 微观经济学. 北京: 华夏出版社, 1982

[14] K. Goldberg, A. Goldman, M. Newman. The probability of an equilibrium point. Journal of Research of the National Bureau of Standards–B. Mathematical Sciences, 1968, 72B (2):93–101

[15] M. Dresher. Probability of a pure equilibrium point in *n*-person games. Journal of Combinatorial Theory, 1970, 8:134–145

[16] S. Takahashi. The number of pure Nash equilibria in a random game with nondecreasing best responses. Games and Economic Behavior, 2008, 63:328–340

[17] 孙微. 基于博弈论的排队经济学模型及策略分析. 秦皇岛: 燕山大学, 2010

[18] P. Naor. The regulation of queue size by levying tolls. Econometrica, 1969, 37:15–24

[19] 张峰. 排队服务系统中策略性顾客的经济博弈策略分析. 北京: 北京交通大学, 2014

[20] N. Knudsen. Individual and social optimization in a multi-server queue with a general cost-benefit structure. Econometrica, 1972, 40:515–528

[21] U. Yechiali. On optimal balking rules and toll charges in the GI/M/1 queue. Operations Research,1971, 19:349–370

[22] P. Afèche, H. Mendelson. Priority auctions vs. uniform pricing in queueing systems with a generalized delay cost structure. Management Science, 2004, 50:869–882

[23] H. Chen, M. Frank. State dependent pricing with a queue. IIE Transactions, 2001, 33(10):847–860

[24] R. Hassin. On the advantage of being the first server. Management Science, 1996, 42(4):618–623

[25] H. Chen, Y. Wan. Price competition of make-to-order firms. IIE Transactions, 2003, 35(9):817–832

[26] X. Su. Intertemporal pricing with strategic customer behavior. Management Science, 2007, 53 (5):726–741

[27] H. Chen, M. Frank. Monopoly pricing when customers queue. IIE Transactions, 2004, 36 (6):569–581

[28] D. Christ, B. Avi-Itzhak. Strategic equilibrium for a pair of competing servers with convex cost and balking. Management Science, 2002, 48(6):813–820

[29] H. Chen, Y. Wan. Capacity competition of make-to-order firms. Operations Research Letters, 2005,33(2):187–194

[30] R. Hassin, J. Puerto, F. Fernández. The use of relative priorities in optimizing the performance of a queueing system. European Journal of Operational Research, 2009, 193:476–483

[31] P. Guo, R. Hassin. On the advantage of leadership in service pricing competition. Operations Research Letters, 2013, 41(4):397–402

[32] N. Edelson, K. Hildebrand. Congestion tolls for Poisson queueing processes. Econometrica, 1975,43:81–92

[33] R. Hassin. Consumer information in markets with random products quality: The case of queues and balking. Econometrica, 1986, 54:1185–1195

[34] W. Stein, A. Rapoport, D. Seale, et al. Batch queues with choice of arrivals: Equilibrium analysis and experimental study. Games and Economic Behavior, 2007, 59:345–363

[35] A. Burnetas, A. Economou. Equilibrium customer strategies in a single server Markovian queue with setup times. Queueing Systems, 2007, 56:213–228

[36] R. Hassin, M. Haviv. Equilibrium behavior in queueing systems: to queue or not to queue. Boston:Kluwer, 2003

[37] P. Guo, P. Zipkin. Analysis and comparison of queues with different levels of delay information.Management Science, 2007, 53(6):962–970

[38] P. Guo, R. Hassin. Strategic behavior and social optimization in Markovian vacation queues. Operations Research, 2011, 59(4):986–997

[39] P. Guo, R. Hassin. Strategic behavior and social optimization inMarkovian vacation queues: the case of heterogeneous customers. European Journal of Operational Research, 2012, 222 (2):278–286

[40] K. Balachandran. Purchasing priorities in queues. Management Science, 1972, 18:319–326

[41] R. Hassin. Decentralized regulation of a queue. Management Science, 1995, 41:163–173

[42] 张维迎. 博弈论与信息经济学. 上海: 上海人民出版社, 2012

[43] W. Sun, P. Guo, N. Tian. Equilibrium threshold strategies in observable queueing systems with setup/closedown times. Central European Journal of Operations Research, 2010, 18 (3):241–268

[44] W. Sun, Y. Wang, N. Tian. Pricing and setup/closedown policies in unobservable queues with strategic customers. 4OR: A Quarterly Journal of Operations Research, 2012, 10(3):287–311

[45] B. Doshi. Queueing Systems with Vacations–a Survey. Queueing Systems, 1986, 1(1):29–66

[46] G. K. Reddy, R. Nadarajan, R. Arumuganathan. Analysis of a bulk queue with N-policy multiple vacations and setup times. Computers and Operations Research, 1998, 25 (11): 957–967

[47] G. Choudhury. On a batch arrival poisson queue with a random setup time and vacation period. Computers and Operations Research, 1998, 25:1013–1026

[48] G. Choudhury. An M^x/G/1 queueing system with a setup period and a vacation period. Queueing Systems, 2000, 36:23–38

[49] W. Bischof. Analysis of M/G/1 queues with setup times and vacations under six different service disciplines. Queueing Systems, 2001, 39:265–301

[50] J. Ke. The optimal control of an M/G/1 queueing system with server startup and two vacation types. Applied Mathematical Modelling, 2003, 27:437–450

[51] J. Artalejo, A. Economou, M. Lopez-Herrero. Analysis of a multiserver queue with setup times. Queueing Systems, 2005, 52:53–76

[52] R. Arumuganathan, S. Jeyakumar. Analysis of a bulk queue with multiple vacations and closedown times. International Journal of Information and Management Sciences, 2004, 15 (1):45–60

[53] R. Arumuganathan, S. Jeyakumar. Steady state analysis of a bulk queue with multiple vacations, setup times with N-policy and closedown times. Applied Mathematical Modelling, 2005, 29:972–986

[54] J. Ke. On M/G/1 system under NT policies with breakdowns, startup and closedown. Applied Mathematical Modelling, 2006, 30:49–66

[55] J. Ke. Batch arrival queues under vacation policies with server breakdowns and startup/ close-down times. Applied Mathematical Modelling, 2007, 31:1282–1292

[56] W. Sun, S. Li, N. Tian. Equilibrium mixed strategies of customers in an unobservable queue with multiple vacations. Quality Technology and Quantitative Management, 2013, 10 (4):389–421

[57] W. Sun, S. Li, C. E. Equilibrium and optimal balking strategies of customers in Markovian queues with multiple vacations and *N*-policy. Applied Mathematical Modelling, 2016, 40 (1):284–301

[58] M. Neuts. Matrix-geometric Solutions in Stochastic Models. Baltimore: Johns Hopkins University Press, 1981

[59] A. Economou, S. Kanta. Equilibrium balking strategies in the observable single-server queue with breakdowns and repairs. Operations Research Letters, 2008, 36

[60] I. Adan, A. Economou, S. Kapodistria. Synchronized reneging in queueing systems with vacations.Queueing Systems, 2009, 62(1):1–33

[61] W. Liu, Y. Ma, J. Li. Equilibrium threshold strategies in observable queueing systems under single vacation policy. Applied Mathematical Modelling, 2012, 36(12):6186–6202

[62] J.Wang, F. Zhang. Equilibrium analysis of the observable queues with balking and delayed repairs.Applied Mathematics and Computation, 2011, 218(6):2716–2729

[63] Y. Ma,W. Liu, J. Li. Equilibrium balking behavior in the Geo/Geo/1 queueing system with multiple vacations. Applied Mathematical Modelling, 2013, 37(6):3861–3878

[64] A. Economou, A. Gómez-Corral, S. Kanta. Optimal balking strategies in single-server queues with general service and vacation times. Performance Evaluation, 2011, 68:967–982

[65] P. Guo, Q. Li. Strategic behavior and social optimization in partially-observable Markovian vacation queues. Operations Research Letters, 2013, 41:277–284

[66] Y. Dimitrakopoulos, A. Burnetas. Customer equilibrium and optimal strategies in an M/M/1 queue with dynamic service control. Working paper, 2011

[67] W. Sun, S. Li, Q. Li. Equilibrium balking strategies of customers in Markovian queues with twostage working vacations. Applied Mathematics and Computation, 2014, 248:195–214

[68] W. Sun, S. Li, N. Tian. Equilibrium and optimal balking strategies of customers in unobservable queues with double adaptive working vacations. Quality Technology and Quantitative Management, 2016

[69] W. Sun, S. Li, W. Liu. Comparisons of exhaustive and nonexhaustive M/M/1/N queues with working vacation and threshold policy. Working paper, 2016

[70] W. Sun, S. Li. Equilibrium and optimal behavior of customers in Markovian queues with multiple working vacations. Top, 2014, 22(2):694–715

[71] W. Liu, X. Xu, N. Tian. Stochastic decompositions in the M/M/1 queue with working vacations. Operations Research Letters, 2007, 35(5):595–600

[72] D.Wu, H. Takagi. M/G/1 Queue with Multiple Working Vacations. Performance Evaluation, 2006,63:654–681

[73] J. Li, N. Tian, Z. Zhang, et al. Analysis of the M/G/1 queue with exponentially working vacations-amatrix analytic approach. Queueing Systems, 2009, 61(2-3):139–166

[74] J. Li, N. Tian. Performance analysis of a GI/M/1 queue with single working vacation. Applied Mathematics and Computation, 2011, 217(10):4960–4971

[75] F. Zhang, J. Wang, B. Liu. Equilibrium balking strategies in Markovian queues with working vacations. Applied Mathematical Modelling, 2013, 37(16-17):8264–8282

[76] F. Wang, J. Wang, F. Zhang. Equilibrium customer strategies in the Geo/Geo/1 queue with single working vacation. Discrete Dynamics in Nature and Society, 2014, 9:Article ID 309489

[77] D. Guha, V. Goswami, A. Banik. Equilibrium balking strategies in renewal input batch arrival queues with multiple and single working vacation. Performance Evaluation, 2015, 94:1–24

[78] W. Sun, S. Li, C. E, et al. Comparisons of site-clearing and non-site-clearing Markovian queues with server failures and unreliable repairer. Working paper, 2016

[79] W. Sun, S. Li. Comparisons of customer behavior in Markovian queues with vacation policies and geometric abandonments. Working paper, 2016

[80] E. Altman, U. Yechiali. Analysis of customers' impatience in queues with server vacations. Queueing Systems, 2006, 52(4):261–279

[81] P. Laxmi, K. Jyothsna. Impatient customer queue with Bernoulli schedule vacation interruption. Computers and Operations Research, 2015, 56:1–7

[82] D. Yue, W. Yue, G. Zhao. Analysis of an M/M/1 queue with vacations and impatience timers which depend on the server's states. Journal of Industrial and Management Optimization, 2015,12(2):653–666

[83] S. Kapodistria. The M/M/1 queue with synchronized abandonments. Queueing systems, 2011, 68(1):79–109

[84] S. Dimou, A. Economou, D. Fakinos. The single server vacation queueing model with geometric abandonments. Journal of Statistical Planning and Inference, 2011, 141:2863–2877

[85] U. Yechiali. Queues with system disasters and impatient customers when system is down. Queuein Systems, 2007, 56:195–202

[86] A. Economou, S. Kapodistria. Synchronized abandonments in a single server unreliable queue. European Journal of Operational Research, 2010, 203:143–155

[87] S. Dimou, A. Economou. The single server queue with catastrophes and geometric reneging. Methodology and Computing in Applied Probability, 2013, 5(3):595–621

[88] O. Boudali, A. Economou. Optimal and equilibrium balking strategies in the single server Markovian queue with catastrophes. European Journal of Operational Research, 2012, 218 (3):708–715

[89] O. Boudali, A. Economou. The effect of catastrophes on the strategic customer behavior in queueing systems. Naval Research Logistics, 2013, 60(7):571–587

[90] R. Hassin, M. Haviv. Equilibrium strategies for queues with impatient customers. Operations Research Letters, 1995, 17:41–45

[91] M. Haviv, Y. Ritov. Homogeneous customers renege from invisible queues at random times under deteriorating waiting conditions. Queueing Systems, 2001, 38:495–508

[92] A. Mandelbaum, N. Shimkin. A model for rational abandonments from invisible queues. Queueing Systems, 2000, 36:141–173

[93] E. Zohar, A. Mandelbaum, N. Shimkin. Adaptive behavior of impatient customers in Tele-Queues:theory and empirical support. Management Science, 2002, 48(4):566–583

[94] N. Shimkin, A. Mandelbaum. Rational abandonment from tele-queues: Nonlinear waiting cost with heterogeneous preferences. Queueing Systems, 2004, 47:117–146

[95] P. Guo,W. Sun, Y.Wang. Equilibrium and optimal strategies to join a queue with partial information on service times. European Journal of Operational Research, 2011, 214 (2):284–297

[96] J. Kapur, H. Kesavan. Entropy Optimization Principles with Applications. Boston: Academic Press, 1992

[97] C. Shannon. A mathematical theory of communication. The Bell System Technical Journal, 1948,27:623–656

[98] E. Jaynes. Information theory and statistical mechanics. Physical Review, 1957, 106:620–630

[99] G. Perakis, G. Roels. Regret in the Newsvendor model with partial information. Operations Research,2008, 56(1):188–203

[100] S. Veeraraghavan, L. Debo. Herding in queues with waiting costs: rationality and regret. Manufacturing & Service Operations Management, 2011, 13(3):329–346

[101] W. Sun, S. Li. Effect of information, uncertainty and parameter variability on profits in a queue with various pricing strategies. International Journal of Systems Science, 2014, 45 (8):1781–1798

[102] P. Guo, P. Zipkin. The effects of information on a queue with balking and phase-type service times.Naval Research of Logistics, 2008, 55(5):406–411

[103] C. Larsen. Investigating sensitivity and the impact of information on pricing decisions in an M/M/1/1 queueing model. International Journal of Production Economics, 1998, 56-57: 365–377

[104] R. Hassin, M. Haviv. Equilibrium strategies and the value of information in a two line queueing system with threshold jockeying. Communications in Statistics–Stochastic Models, 1994, 10(2):415–435

[105] L. Debo, L. Toktay, L. V. Wassenhove. Queuing for Expert Services. Management Science, 2008, 54 (8):1497–1512

[106] S. Veeraraghavan, L. Debo. Joining longer queues: information externalities in queue choice. Manufacturing & Service Operations Management, 2009, 11(4):543–562

[107] R. Arnott, A. de Palma, R. Lindsey. Information and usage of free-access congestible facilities with stochastic capacity and demand. International Economic Review, 1996, 37(1):181–203

[108] R. Arnott, A. de Palma, R. Lindsey. Information and time-of-usage decisions in the bottleneck model with stochastic capacity and demand. European Economic Review, 1999, 43 (3): 525–548

[109] P. Guo, P. Zipkin. The effects of the availability of waiting-time information on a balking queue.European Journal of Operational Research, 2009, 198(1):199–209

[110] E. Altman, R. Marquez, U. Yechiali. Admission and routing control with partial information and limited buffers. International Journal of Systems Science, 2003, 34(10-11):615–626

[111] R. Hassin. Information and uncertainty in a queueing system. Probability in the Engineering and Informational Sciences, 2007, 21(3):361–380