

PH模型

在城轨交通车站系统中的应用

蒋阳升 著

Phase-type
and its application
in urban rail transit system



科学出版社

PH 模型在城轨交通车站系统中的应用

蒋阳升 著

科学出版社

北 京

内 容 简 介

本书主要介绍 PH 分布的相关概念、原理与方法及 PH 分布在城市轨道交通车站系统中的应用。全书共 6 章,第 1 章介绍 PH 分布及其排队模型;第 2 章介绍轨道交通车站行人数据采集及 PH 拟合;第 3 章介绍城市轨道交通车站设施节点的 PH 排队模型;第 4 章介绍城市轨道交通车站设施 PH 排队网络模型;第 5 章介绍 PH 分布及其排队模型在城市轨道交通系统中的其他应用;第 6 章介绍 PH 分布及其排队模型在其他城市交通系统中的推广应用。

本书适合高等院校城市轨道交通专业的教师和研究生阅读,也可以作为轨道交通规划与管理领域工程技术人员的参考资料。

图书在版编目(CIP)数据

PH 模型在城轨交通车站系统中的应用 / 蒋阳升著. —北京:科学出版社, 2015

ISBN 978-7-03-043539-2

I. ①P… II. ①蒋… III. ①信息系统-应用-城市铁路-车站设备-排队系统 IV. ①U239.5-39

中国版本图书馆 CIP 数据核字 (2015) 第 036104 号

责任编辑:杨 岭 于 楠 / 封面设计:墨创文化

责任校对:贺江艳 / 责任印制:余少力

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

成都创新包装印刷厂印刷

科学出版社发行 各地新华书店经销

*

2015 年 6 月 第 一 版 开本: 787*1092 1/16

2015 年 6 月第一次印刷 印张: 10 插页: 16 面

字数: 260 千字

定价: 68.00 元

前言

交通系统本质上是一类随机服务系统，因其具有高度的随机性和动态性，导致许多交通问题成为世界性的难题。大数据时代的到来，为交通系统的动态随机分析提供了前所未有的研究环境。可以预见，在不久的将来，基于大数据的交通动态随机分析将成为交通领域的重要研究分支和热点，并且该分析技术的发展将伴随着交通其他领域的深层次的技术的突破和更新，引领大数据时代交通由智能向智慧的进化。本书就是作者及其团队在国家自然科学基金青年项目等资助下对该领域的阶段性研究成果。

众所周知，以指数分布 M 为基础的经典生灭过程是分析处理随机现象的基本工具之一。指数分布 M 具有易于进行解析处理的优点，也有明显的局限性，如要求数据的变异系数等于 1（即标准差等于均值），但这种假设的不合理性是显而易见的。长期以来，随机模型分析中一直存在一个突出的矛盾：一方面，离开指数分布 M ，随机模型的定量分析就面临着实质性困难；另一方面，假定所研究的随机变量服从指数分布 M ，但在多数情况下都与客观事实大相径庭。由此可见，模型处理的方便性和描述的真实性难以兼顾。

PH 分布（即位相型分布）的出现解决了这一问题。理论上，PH 分布保持了指数分布 M 易于进行解析处理的优点，又能无限逼近任意非负随机变量（如城市轨道交通车站的客流到达间隔和服务时间），且 PH 分布的不同位相适合于描述具有多层次和变动参数的复杂随机现象。因此，PH 分布完全可取代指数分布 M （PH 分布的特例），现阶段在经典生灭过程的基础上已经发展出一种分析 PH 排队模型的拟生灭过程（简称 QBD 过程）。但如何把这一纯理论研究成果应用到具体的物理系统的随机分析中还需要探索。

作者主要依托西南交通大学优秀博士学位论文培育项目和国家自然科学基金项目（编号：51108391；课题：考虑稳定性的城市轨道交通车站交通服务设施系统优化配置理论与方法——基于排队网络视角），以城市轨道交通车站乘客服务设施系统为对象，在 Sadre 和 Smith 等研究的基础上，采用 QBD 过程首次推导了适合候车系统排队现象分析的 PH/PH^[0, C₁]/1/C 批量服务排队模型、适合售检票系统排队现象分析的 PH/PH/C₁/C 并联排队模型、适合自动扶梯系统排队现象分析的 PH/PH/C₁/C 混联排队模型和适合通道楼梯系统排队现象分析的 PH/PH(n)/C/C 状态相关混联排队模型，以及适合车站系统排队现象分析的 PH 排队网络模型。这项工作拓展了 PH 分布及排队网络分析技术在城市轨道交通领域的崭新应用，相关研究成果均已发表在国际应用数学和运筹学期刊上。

在上述研究成果的基础上，作者发现这些 PH 排队模型可以被推广应用到其他交通领域，如紧急状态下车站疏散设施配置、信号交叉口配时和高架环道入口匝道控制等。尽管这些 PH 排队模型在描述实际排队现象上还存在一定的缺陷，需要后续研究不断完善，然而其研究成果仍获得了国内外知名专家 Smith 和 Cruz 的认可，这也让作者鼓起勇气出版这部专著，以方便读者跟踪本团队的研究动态。

全书共包含6章,第1章介绍了PH分布及其排队模型,首先用以指数分布为基础的经典生灭过程引出PH分布,并在此基础上介绍了PH排队模型;第2章介绍了行人数据采集及其PH拟合,主要包括轨道交通车站行人到达间隔数据采集方法、轨道交通车站设施服务时间数据采集方法和对应的PH拟合方法;第3章介绍了城市轨道交通车站设施节点PH排队模型,主要涉及候车系统PH排队模型、售检票设施PH排队模型、自动扶梯PH排队模型和通道楼梯PH排队模型等四类,其中包括各系统的描述、系统参数的标定、排队模型的建立及对模型的验证等;第4章在第3章建立的各个节点PH排队模型的基础上,以轨道交通站点为一个整体,建立了城市轨道交通车站设施PH排队网络模型,其中涉及车站服务设施排队网络系统描述、系统参数标定、模型建立、算法设计及模型的验证;第5章是根据研究团队成员韩世凡、吴婷、胡路和朱娟秀的硕博论文整理而成,简单介绍了PH分布及其排队模型在城市轨道交通系统中的其他应用,其中涉及在车站断面客流预测、站间断面客流拟合和车站设施配置等的应用;第6章是根据研究团队成员张改和方传武的硕博论文整理而成,简单介绍了PH分布及其排队模型在其他城市交通系统中的推广应用,其中涉及在信号交叉口优化配时和高架环道入口匝道控制中的应用。

本书在写作过程中得到了许多同行学者、朋友和作者的研究生的帮助,在此一并表示感谢,有了他们的支持才能最终完成本书。还要非常感谢我的家人,有了他们在生活上的无私关怀作者才有充足的精力投入到本书的写作中。感谢国家自然科学基金委员会和西南交通大学对课题研究和本书的出版提供的资助。同时,本书在出版过程中得到了科学出版社的大力支持,特别感谢于楠编辑富有成效的工作和帮助。

由于作者专业视野和学术水平有限,本书难免有不足之处,敬请读者批评指正。

目 录

第 1 章 PH 分布及其排队模型简介	
1.1 PH 分布简介	
1.1.1 指数分布族	
1.1.2 PH 分布	
1.1.3 简化的 PH 分布	
1.2 PH 排队模型介绍	
1.2.1 生灭过程	
1.2.2 拟生灭过程	
1.2.3 PH/PH/1/C 排队模型	1
1.2.4 PH/PH/1/C 排队网络模型	1
1.3 本章小结	1
第 2 章 行人数据采集及其 PH 拟合	1
2.1 行人数据采集	1
2.2 常用的 PH 拟合方法	1
2.2.1 HErD	1
2.2.2 HErD 的极大似然估计模型	2
2.2.3 SS&IAGA-EM 求解算法	2
2.2.4 数据拟合效果展示	2
2.3 本章小结	3
第 3 章 城市轨道交通车站设施节点 PH 排队模型	3
3.1 候车系统 PH 排队模型	3
3.1.1 候车排队系统描述	3
3.1.2 候车排队系统参数标定	3
3.1.3 候车系统 PH 排队模型建立	4
3.2 售检票设施 PH 排队模型	5
3.2.1 售检票设施排队系统描述	5
3.2.2 售检票设施排队系统参数标定	5
3.2.3 售检票设施 PH 排队模型建立	5
3.3 自动扶梯 PH 排队模型	6
3.3.1 自动扶梯排队系统描述	6
3.3.2 自动扶梯排队系统参数标定	6
3.3.3 自动扶梯 PH 排队模型建立	7
3.4 通道楼梯 PH 排队模型	7

3.4.1 通道楼梯排队系统描述	79
3.4.2 通道楼梯排队系统参数标定	80
3.4.3 通道楼梯 PH 排队模型建立	90
3.5 本章小结	99
第4章 城市轨道交通车站设施 PH 排队网络模型	101
4.1 车站服务设施排队网络系统描述	101
4.2 车站服务设施排队网络系统参数标定	105
4.2.1 网络输入流的时间间隔分布的 PH 拟合及参数标定	107
4.2.2 其他排队网络参数标定	115
4.3 车站服务设施 PH 排队网络模型建立	115
4.3.1 问题分析	115
4.3.2 模型建立	116
4.3.3 算法设计	119
4.3.4 模型验证	120
4.5 本章小结	129
第5章 PH 分布及其排队模型在城市轨道交通系统中的其他应用	130
5.1 PH 分布在车站断面客流预测中的应用	130
5.1.1 研究意义	130
5.1.2 研究现状	131
5.1.3 研究思路	133
5.2 PH 分布在站间断面客流拟合中的应用	136
5.2.1 研究意义	136
5.2.2 研究现状	137
5.2.3 研究思路	138
5.3 PH 排队模型在车站设施配置中的应用	140
5.3.1 正常状态下服务设施优化配置	140
5.3.2 紧急状态下疏散设施优化配置	142
5.4 本章小结	145
第6章 PH 分布及其排队模型在其他城市交通系统中的推广应用	146
6.1 在信号配时中的应用	146
6.2 在匝道控制中的应用	148
6.3 本章小结	151
参考文献	152
彩图	155

第 1 章 PH 分布及其排队模型简介

以指数分布 M (特殊的 PH 分布) 为基础的经典生灭过程是分析处理随机现象的基本工具之一。不仅在排队论 (如采用经典生灭过程可以求解 $M/M/1$ 、 $M/M/C_1$ 、 $M/M/C$ 和 $M/M/C_1/C$ 排队模型)、库存论等学科的发展中起着重要的作用, 还广泛应用于程技术、管理科学、生命科学、生态系统、人口理论、行为科学及社会科学等领域。章将从指数分布 M 说起, 引出 PH 分布, 并在此基础上简单介绍 PH 排队模型。

1.1 PH 分布简介

1.1.1 指数分布族

众所周知, 对各种随机模型进行解析处理的主要障碍是条件概率引起的高度复杂性。例如, 计算正在接受服务的顾客的剩余服务时间, 需要对顾客已接受服务的时间有全面的了解。一方面, 这些关于历史过程的信息在很多应用中很难准确得知; 另一方面, 使有了这些信息, 条件分布的表达式也通常十分复杂。经典随机模型理论中广泛使用从指数分布 M 的随机变量, 通过指数分布 M 的无记忆性绕过这一困难。

称随机变量 X 服从参数为 λ 的指数分布 M , 当其有分布函数和密度函数

$$F(x) = 1 - e^{-\lambda x}, \quad f(x) = \lambda e^{-\lambda x}, \quad x > 0$$

X 的数学期望为

$$EX = \frac{1}{\lambda}$$

指数分布 M 最独特的性质是无后效 (无记忆) 性。设随机变量 X 表示一个顾客的服务时间, 对 $t > 0$ 和 $s > 0$, 我们有条件概率

$$P\{X > s + t \mid X > t\} = \frac{1 - F(s + t)}{1 - F(t)} = e^{-\lambda s} \quad (1-1)$$

式 (1-1) 表明, 不管顾客已经接受了多长时间服务, 其剩余服务时间都如 X 一样服从参数为 λ 的指数分布 M 。因此, 只要假定所研究的随机变量服从指数分布 M , 由条件概率产生的种种麻烦就荡然无存了。

稍稍扩大一下范围, 考虑下列指数分布族。设 $0 \leq \alpha \leq 1$, 则形如

$$F(x) = 1 - \alpha e^{-\lambda x}, \quad x \geq 0 \quad (1-2)$$

的全体分布函数称为 (修正的) 指数分布族。特别的, 若 $\alpha = 0$, 式 (1-2) 退化为集中于原点的分布, 其分布函数记为

$$U_0(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

式中, $U_0(x)$ 属于指数分布族。若 $\alpha=1$, 则式(1-2)对应于标准的指数分布 M ; 若 $0 < \alpha < 1$, 则 $F(0)=1-\alpha$, 即 $F(x)$ 在 $x=0$ 处集中着概率质量 $1-\alpha$, 并在 $(0, +\infty)$ 上有不完全概率密度函数

$$F(x) = \alpha\lambda e^{-\lambda x}, \quad x > 0 \quad (1-3)$$

这时, $F(x)$ 是一个混合型分布。它在 $x=0$ 处具有离散分布的特性, 在 $(0, +\infty)$ 上又具有连续随机变量分布的形态。指数分布族的密度函数与分布函数如图 1-1 所示。形如式(1-2)的分布有数学期望:

$$EX = \frac{\alpha}{\lambda} \quad (1-4)$$

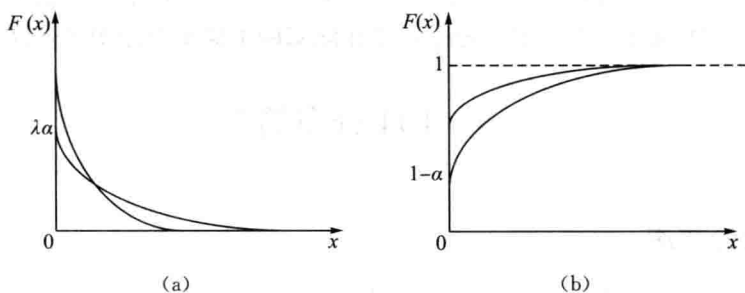


图 1-1 指数分布族的密度函数与分布函数

注意到式(1-2)可改写成

$$F(x) = (1-\alpha)U_0(x) + \alpha(1 - e^{-\lambda x}), x \geq 0 \quad (1-5)$$

这表明指数分布族中的任何分布, 都是集中于原点的分布与一个标准指数分布 M 的混合 [$F(x)$ 以概率 $1-\alpha$ 是集中于原点上的退化分布, 以概率 α 是一个标准指数分布]。因此, 如果 X 的分布函数如式(1-2)所示, 则 X 以概率 $1-\alpha$ 集中于原点, 以概率 α 服从参数 λ 的标准指数分布 M , 并具有指数分布的无后效性。

1.1.2 PH 分布

PH 分布是一个有限状态 Markov(马尔可夫)过程吸收时间的分布。Jensen(1954)首先把这一分布用于一类经济模型。Neuts(1987)成功地发展了处理这类分布的矩阵分析技巧, 其分布函数被表示成矩阵指数形式, 清晰地显示出 PH 分布类与指数分布类的深刻类比。

在给出 PH 分布的定义前, 先考虑一个在状态集 $\{1, 2, \dots, m, m+1\}$ 上的马尔可夫过程。状态 $1, 2, \dots, m$ 都是非常返的(马尔可夫过程在这些状态中不断转移), 状态 $m+1$ 是吸收的(若马尔可夫过程到达该状态则转移结束), 其状态转移如图 1-2 所示, 过程的无穷小生成元可以写成下列分块形式

$$Q = \begin{bmatrix} T & T^0 \\ \mathbf{0} & 0 \end{bmatrix}$$

n 阶方阵 $T = (T_{ij})$ 满足 $Te \leq 0$, 其中 $T_{ii} < 0$, $T_{ij} \geq 0$, $i \neq j$, $1 \leq i, j \leq m$ 。 $T^0 = (T_1^0, T_2^0, \dots, T_m^0)^T$ 是非负列向量, 满足 $Te + T^0 = 0$; $e = (1, 1, \dots, 1)^T_{m \times 1}$, 即行数

与前面相乘矩阵的列数相同且元素均为1的列向量；过程的初始概率是 (α, α_{m+1}) ，其中， $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ ， $\alpha \cdot e + \alpha_{m+1} = 1$ 。这一马尔可夫过程可以结合青蛙在池塘的荷叶上随机跳动进行理解：将状态1, 2, $\dots$, m 理解为池塘的荷叶，将状态 $m+1$ 理解为池塘外，将 α 理解为青蛙开始跳入池塘荷叶的选择概率，将 α_{m+1} 理解为青蛙还未开始就已经离开了池塘的概率，将 T 理解为青蛙在荷叶之间的转移强度，将 T^0 理解为青蛙从荷叶跳到池塘外的转移强度(可结合下文例1和例2进一步理解)。

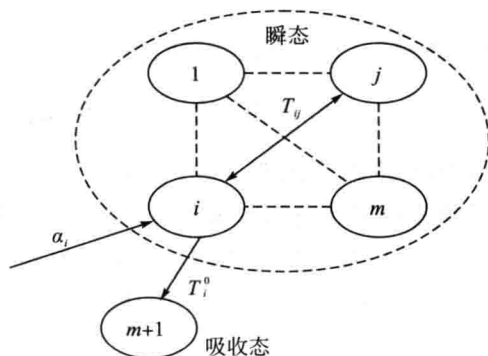


图 1-2 PH 分布状态转移图

假设一个有限状态马尔可夫过程以概率 (α, α_{m+1}) 开始进入状态集合 $\{1, 2, \dots, m, m+1\}$ 并按照 Q 矩阵进行状态之间的转移，则该马尔可夫过程进入吸收状态 $m+1$ 的时间分布定义为 PH 分布。其概率分布函数为

$$F(x) = 1 - \alpha \exp(Tx) e, x \geq 0 \quad (1-6)$$

其中，符号 (α, T) 称为 PH 分布的 m 阶表示。由定义，下列事实是显然成立的。

(1) $F(0) = 1 - \alpha e = \alpha_{m+1}$ 。若 α 是零向量，则式(1-6)退化为 $U_0(x)$ ，即此时的 PF 分布为集中于原点的一个分布；若 $0 < \alpha e < 1$ ，则 $F(x)$ 在 $x=0$ 处集中着概率质量 α_{m+1} ，并在 $(0, +\infty)$ 上有不完全概率密度函数

$$f(x) = \alpha \exp(Tx) T^0, x \geq 0 \quad (1-7)$$

(2) PH 分布的 i 阶原点矩是 $u_i = \int_0^{+\infty} x^i \alpha \exp(Tx) dx T^0 = (-1)^i i! \alpha T^{-i} e$, $i=1, 2, \dots$ ，所以 PH 分布的均值 EX 、方差 VX 和变异系数的平方 c^2 可分别表示为

$$EX = E(X) = -\alpha T^{-1} e \quad (1-8)$$

$$VX = \text{VAR}(X) = E(X^2) - E(X)^2 = 2\alpha T^{-2} e - (\alpha T^{-1} e)^2 \quad (1-9)$$

$$c^2 = \text{VAR}(X)/E(X)^2 = \alpha T^{-2} e / (-\alpha \times T^{-1} e)^2 - 1 \quad (1-10)$$

式中， $E(\cdot)$ 和 $\text{VAR}(\cdot)$ 分别为数学期望和方差运算符。

分布函数式(1-6)也称为矩阵指数分布，它是指数分布族式(1-2)从数值参数到矩阵参数的一种推广。矩阵指数函数 $\exp(T \cdot x)$ 具有与指数函数 $e^{-\lambda x}$ 完全类似的性质，详见任何一本矩阵分析方面的著作。

PH 分布除了保持指数分布 M 易于进行解析处理的优点，还有很多其他良好的性质，其中，较重要的性质是封闭性：若干个 PH 分布的运算结果通常是一个新的 PH 分布。含有 PH 分布的随机模型，其指标通常也服从 PH 分布。封闭性在模型的理论分析和实

际应用中都带来了极大的方便，具体定理及证明参见田乃硕(1990，1994，1995，2001，2002)的相关研究。除封闭性外，PH 分布还具有稠密性，也就是说，无论研究什么样的随机分布，总可选择一个适当的 PH 分布来代替它，并且使误差达到所要求的精度。指数分布 M 易于处理，但在大多数情况下与客观实际相差甚远；PH 分布在很大程度上保持了指数的优点，又可与客观实际任意接近，这正是 PH 分布的诱人之处，也是 PH 分布迅速普及的重要原因之一。以下将通过具体的实例让读者了解实际生活中的 PH 分布。

例 1(有折旧行为的部件寿命) 部件寿命按折旧程度分为三个阶段，在第 1、2、3 阶段上的使用时间分别服从均值 $\frac{1}{10}$ 、 $\frac{1}{20}$ 和 $\frac{1}{30}$ (年)的指数分布。处在寿命第 1 阶段的部件以 0.2 的概率失效，并以 0.8 的概率因老化进入第 2 阶段，处在寿命为第 2 阶段的部件以 0.4 的概率失效，以 0.6 的概率因老化进入寿命的第 3 阶段。处在第 3 阶段的部件老化后将全部失效或更新。此外，由于内在的质量原因，一部件在投入使用时处在第 1、2、3 阶段的概率分别是 0.62、0.28 和 0.09，并以 0.01 的概率根本不能使用，部件寿命折旧行为如图 1-3 所示。

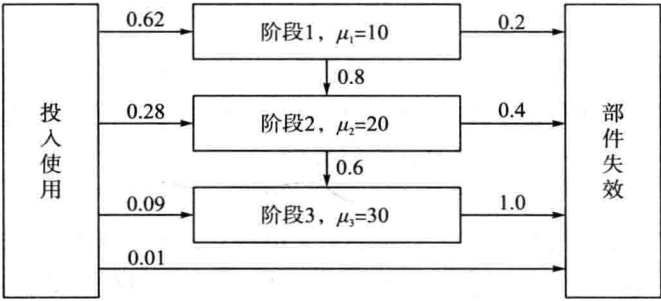


图 1-3 部件寿命的折旧行为图

以 X 表示任取一部件的使用寿命， X 服从一个 3 阶 PH 分布，有 PH 表示为 (α, T) ，其中

$$\begin{aligned} \alpha &= (\alpha_1, \alpha_2, \alpha_3) = (0.62, 0.28, 0.09), \alpha_4 = 0.01 \\ T_{11} &= -10, T_{12} = 10 \times 0.8, T_{13} = 0, T_1^0 = 10 \times 0.2 \\ T_{21} &= 0, T_{22} = -20, T_{23} = 20 \times 0.6, T_2^0 = 20 \times 0.4 \\ T_{31} &= 0, T_{32} = 0, T_{33} = -30, T_3^0 = 30 \end{aligned}$$

于是

$$T = \begin{bmatrix} -10 & 8 & 0 \\ 0 & -20 & 12 \\ 0 & 0 & -30 \end{bmatrix}, T^0 = \begin{bmatrix} 2 \\ 8 \\ 30 \end{bmatrix}$$

平均使用寿命是

$$EX = \alpha (-T^{-1}) e = 0.11932 \text{ 年}$$

例 2(柔性制造系统中产品的加工时间) 柔性制造系统中某产品的加工流程如图 1-4 所示，加工从第 1 道工序开始，并在每一道工序加工过程中自动测试各种指标，以决定

进一步加工的需要,在完成第1道工序后,有10%的半成品可送到第3道工序稍作补充加工即可出厂,另有90%半成品需送到第2道工序进一步加工,第2道工序上的半成品加工后有60%可达到成品要求,20%需退回第1道工序返工,另有20%需送到第2道工序进行补充加工后达到成品要求。产品在工序1、2、3的加工时间依次服从数学期望为6、5、2(min)的指数分布,为了简单,假定产品到达每一道工序不需要等待。

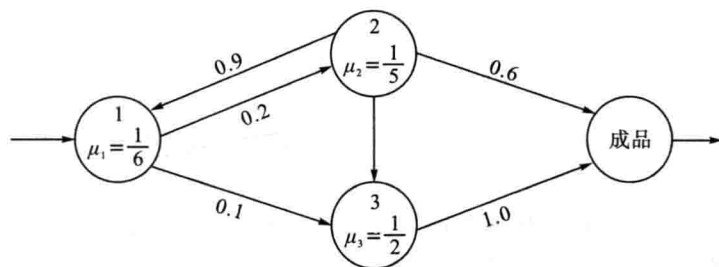


图 1-4 柔性制造系统中某产品的加工流程

以 X 表示一个产品的加工时间, X 服从 3 阶的 PH 分布, 有 PH 表示为 (α, T) , 其中,

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (1, 0, 0), \alpha_4 = 0$$

$$T_{11} = -\frac{1}{6}, T_{12} = 0.9 \times \frac{1}{6}, T_{13} = 0.1 \times \frac{1}{6}, T_1^0 = 0$$

$$T_{21} = 0.2 \times \frac{1}{5}, T_{22} = -\frac{1}{5}, T_{23} = 0.2 \times \frac{1}{5}, T_2^0 = 0.6 \times \frac{1}{5}$$

$$T_{31} = 0, T_{32} = 0, T_{33} = -\frac{1}{2}, T_3^0 = \frac{1}{2}$$

于是

$$T = \begin{bmatrix} -\frac{1}{6} & \frac{3}{20} & \frac{1}{60} \\ \frac{1}{25} & -\frac{1}{5} & \frac{1}{25} \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}, T^0 = \begin{bmatrix} 0 \\ \frac{3}{25} \\ \frac{1}{2} \end{bmatrix}$$

产品的平均加工时间是

$$EX = \alpha(-T^{-1})e = 13.4878 \text{ min}$$

1.1.3 简化的 PH 分布

尽管 PH 分布理论上可无限逼近任意非负随机变量, 但却需要确定多达 $m^2 + m$ 个参数 (m 一般不小于 2), 而在很多实际应用中, 给出的数据资料有限, 无法标定所有参数 (如城市轨道交通车站在规划设计阶段尚未建成, 无法获取详细的数据资料)。Sadr (1999) 和 Weerstra (1994) 等采用如下基于数据的均值 EX 和变异系数的平方 c^2 即可完全确定简化的 PH 拟合一般的非负随机变量取得了很好的效果:

(1) 当 $c^2 \leq 1$ 时, $m = \lceil 1/c^2 \rceil$, $\alpha = (1, 0, \dots, 0)_{1 \times m}$, 矩阵 $\mathbf{T}$ 可以表示如下

$$\mathbf{T} = \begin{bmatrix} -t_1 & t_1 & & & \\ & -t_2 & t_2 & & \\ & & \ddots & \ddots & \\ & & & -t_{m-1} & t_{m-1} \\ & & & & -t_m \end{bmatrix} \quad (1-11)$$

式中, $t_i = m/EX$, $1 \leq i \leq m-2$; $t_m = \left[1 + \sqrt{\frac{1}{2}m(mc^2 - 1)} \right] / [EX(m+2-m^2c^2)]$; $t_{m-1} = mt_m/(2t_mEX - m)$ 。当 $c^2 = 1$ 时, $m = 1$, 此时 PH 分布相当于指数分布 M , 可见指数分布 M 是简化的 PH 分布的特例; 对于 $c^2 \leq 1/30$ 的数据(包括定长分布数据, 其 c^2 为 0), 为了减少在排队分析过程中的计算复杂度, 取 $m = 30$, 此时 PH 分布相当于一个 30 阶 Erlang 分布, 能够获得较好的拟合效果。

(2) 当 $c^2 > 1$ 时, $m = 2$, $\alpha = (p, 1-p)$, $p = \frac{1}{2} [1 + \sqrt{(c^2 - 1)/(c^2 + 1)}]$, 矩阵 $\mathbf{T}$ 可以表示如下

$$\mathbf{T} = \begin{pmatrix} -2p/EX & \\ & -2(1-p)/EX \end{pmatrix} \quad (1-12)$$

在本书中, 当用该简化的 PH 分布描述到达间隔随机变量、服务时间随机变量和离去间隔随机变量时, 也分别用符号 $\langle \lambda, c_a^2 \rangle$ 、 $\langle u, c_s^2 \rangle$ 和 $\langle \omega, c_d^2 \rangle$ 来表示, 其中 $\lambda, u, \omega = 1/EX$ 分别表示到达率、服务率和离去率, c_a^2, c_s^2 和 c_d^2 分别是到达时间间隔、服务时间和离去间隔数据的变异系数的平方。该分布有以下性质, 其中性质 1-1 和性质 1-3 的证明可以参考 Sadre(2007, 1999, 2011)等人的研究文献。

性质 1-1 现有 n 股交通流, 其时间间隔分别服从 $\langle \lambda_i, c_{a,i}^2 \rangle$, $i = 1, 2, \dots, I$ 的 PH 分布, 当合并成一股交通流时, 其时间间隔服从 $\langle \lambda, c_a^2 \rangle$ 的 PH 分布, 其中

$$\lambda = \sum_{i=1}^I \lambda_i, c_a^2 = \sum_{i=1}^I \frac{\lambda_i}{\lambda} c_{a,i}^2$$

性质 1-2 非负随机变量 X 服从 $\langle u, c_s^2 \rangle$ 的 PH 分布, 其 l 阶 PH 表示为 $(\beta, \mathbf{S})$, 则 $Y = X/n$ (n 为正整数)服从 $\langle nu, c_s^2 \rangle$ 的 PH 分布, 并且有 l 阶 PH 表示 $(\beta, n\mathbf{S})$ 。

证 令 $Y = X/n$ 的分布函数为 $G(t)$, 则

$$G(t) = P(Y \leq t) = P(X/n \leq t) = P(X \leq nt) = F(nt) = 1 - \beta \exp(n\mathbf{S}t)\mathbf{e}, t \geq 0 \quad (1-13)$$

令矩阵 $\mathbf{S} = (S_{ij})_{l \times l}$, $\mathbf{S}' = n\mathbf{S}$, 则 $\mathbf{S}' = (nS_{ij})_{l \times l}$, $G(t) = 1 - \beta \exp(\mathbf{S}'t)\mathbf{e}$ 。由于 X 的 l 阶 PH 表示为 $(\beta, \mathbf{S})$, 因此, 由 PH 分布的定义可知, $\mathbf{S}$ 的元素满足 $\mathbf{S}\mathbf{e} \leq \mathbf{0}$, $S_{ii} < 0$ 和 $S_{ij} \geq 0, i \neq j$ 。于是 $\mathbf{S}'$ 的元素满足 $\mathbf{S}'\mathbf{e} = n\mathbf{S}\mathbf{e} \leq \mathbf{0}$, $nS_{ii} < 0$ 和 $nS_{ij} \geq 0, i \neq j$, 因此, $Y = X/n$ 仍服从 PH 分布, 并有 l 阶表示为 $(\beta, n\mathbf{S})$ 。

由于 X 的均值 $E(X) = 1/u$, Y 的均值 $E(Y) = E(X/n) = E(X)/n$, 因此 $E(Y) = 1/(nu)$; 由于 X 的方差 $\text{VAR}(X) = c_s^2/u^2$, Y 的方差 $\text{VAR}(Y) = \text{VAR}(X/n) = \text{VAR}(X)/n^2$, 所以 $\text{VAR}(Y) = c_s^2/(n^2u^2)$; 于是, Y 的变异系数的平方 $\text{VAR}(Y)/E(Y)^2 = c_s^2$ 。

因而 $Y=X/n$ 服从 $\langle nu, c_s^2 \rangle$ 的 PH 分布。

性质 1-3 现有一股交通流, 其时间间隔服从 $\langle \omega, c_d^2 \rangle$ 的 PH 分布。当按照概率 $p_i, i=1, 2, \dots, I$ 分离时, 各个分流的时间间隔分布服从 $\langle \omega_i, c_{d,i}^2 \rangle, i=1, 2, \dots, I$ 的 PH 分布, 其中 $\omega_i = \omega \cdot p_i, c_{d,i}^2 = p_i c_d^2 + (1-p_i)$, $i=1, 2, \dots, I$ 。

性质 1-4 非负随机变量 X 服从 $\langle \lambda, c_a^2 \rangle$ 的 PH 分布, 则其 n 重卷积 $X^{(n)}$ 仍为 PH 分布, 且相应的简化的 PH 分布为 $\langle \lambda/n, c_a^2/n \rangle$ 。

证 由 Whitt(1983)提供的文献可知, PH 分布的 n 重卷积仍为 PH 分布, 因此只需证明 $X^{(n)}$ 的简化 PH 分布为 $\langle \lambda/n, c_a^2/n \rangle$ 即可。

令 $Y=X^{(n)}$ 。由于 X 的均值 $E(X)=1/\lambda$, Y 的均值 $E(Y)=E(X^{(n)})=nE(X)$, 所以 $E(Y)=n/\lambda$; 因为 X 的方差 $\text{VAR}(X)=c_a^2/\lambda^2$, Y 的方差 $\text{VAR}(Y)=\text{VAR}(X^{(n)})=n\text{VAR}(X)$, 所以 $\text{VAR}(Y)=nc_a^2/\lambda^2$; 于是, Y 的变异系数的平方 $\text{VAR}(Y)/E(Y)^2=c_a^2/n$ 。因而 $Y=X^{(n)}$ 服从 $\langle \lambda/n, c_a^2/n \rangle$ 的简化的 PH 分布。

1.2 PH 排队模型介绍

1.2.1 生灭过程

以指数分布 M (特殊的 PH 分布) 为基础的经典生灭过程的研究和应用已经有近一百年的历史, 不仅形成了相当完整的理论体系, 而且广泛应用于自然科学、工程技术、生命科学、人口理论、行为科学及社会科学各种领域。我国学者王梓坤等(1980)在生灭过程理论研究中曾作出重要贡献, 以下将通过对经典生灭过程进行一些简单的介绍, 使读者对生灭过程有一定的了解, 由于生灭过程与随机过程息息相关, 读者也可参考相关随机过程的书籍。

非负整数集 $\{0, 1, \dots\}$ 上的马尔可夫过程 $\{X(t), t \geq 0\}$, 如果其无穷小生成元有下列三对角形式

$$\tilde{Q} = \begin{bmatrix} -\lambda_0 & \lambda_0 & & & \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & & \\ & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \\ & & \dots & \dots & \dots \\ & & & \mu_n & -(\lambda_n + \mu_n) & \lambda_n \\ & & & & \dots & \dots \end{bmatrix} \quad (1-14)$$

则称 $\{X(t), t \geq 0\}$ 是一个生灭过程。 $\{\lambda_n, n \geq 0\}$ 称为生率, $\{\mu_n, n \geq 1\}$ 称为灭率引入瞬时状态概率 $p_j(t) = P\{X(t) = j\}, j = 0, 1, \dots$ 。生灭过程的最基本特征: 在充分短的时间内, 过程几乎只能转移到相邻的状态上, 即当 $\Delta t \rightarrow 0$ 时, 有

$$P\{X(t + \Delta t) = k \mid x(t) = j\}$$

$$= \begin{cases} \lambda_j \Delta t + O(\Delta t), k = j + 1 \\ \mu_j \Delta t + O(\Delta t), k = j - 1, j \geq 1 \\ 1 - (\lambda_j + \mu_j) \Delta t + O(\Delta t), k = j \\ O(\Delta t), |k - j| \geq 2 \end{cases} \quad (1-15)$$

由式(1-15)出发, 容易建立瞬态概率 $p_j(t)$ 满足的微分方程

$$\begin{cases} p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t) \\ p'_j(t) = \lambda_{j-1} p_{j-1}(t) - (\lambda_j + \mu_j) p_j(t) + \mu_{j+1} p_{j+1}(t), j \geq 1 \end{cases} \quad (1-16)$$

式(1-16)称为生灭过程的 Kolmogorov 方程。在给定的初始条件下求解微分方程(1-16)是十分困难的。在大多数应用中, 人们更关心过程运行相当长时间以后的形态, 如果极限

$$p_j = \lim_{t \rightarrow \infty} p_j(t), j = 0, 1, \dots, L \quad (1-17)$$

存在, 并且 $p_0 + p_1 + \dots = 1$, 则称过程是正常返的, 并称 $\{p_j, j \geq 0\}$ 为过程的平稳概率分布或稳态分布。经典生灭过程的重要问题之一是如何从生成元(1-14)出发, 判断过程的正常返性并给出它的稳态分布。如果稳态分布存在, 由式(1-17)必有

$$\lim_{t \rightarrow \infty} p'_j(t) = 0, j = 0, 1, \dots, L$$

在式(1-16)的两端取极限, 给出

$$\begin{cases} \lambda_0 p_0 - \mu_1 p_1 = 0 \\ \lambda_{j-1} p_{j-1} - (\lambda_j + \mu_j) p_j + \mu_{j+1} p_{j+1} = 0, j \geq 1 \end{cases} \quad (1-18)$$

如果把稳态分布写成无穷维向量 $\pi = (p_0, p_1, \dots)$, 式(1-18)可改写成更紧凑的矩阵形式

$$\pi \tilde{Q} = 0, \pi e = 1 \quad (1-19)$$

式(1-18)和式(1-19)统称为过程的平衡方程(及正规化条件)。求解差分方程(1-18)可给出过程的平稳概率分布。

1.2.2 拟生灭过程

考虑一个二维马尔可夫过程 $\{X(t), J(t)\}$, 有状态空间

$$\Omega = \{(k, j): k \geq 0, j = 1, 2, \dots, m\}$$

状态集 $\{(k, 1), \dots, (k, m)\}$ 称为水平 $k, k \geq 0$ 。如果将状态按字典序排列后, 其生成元可写成下列分块三对角形式:

$$\tilde{Q} = \begin{pmatrix} A_0 & C_0 & & & \\ B_1 & A_1 & C_1 & & \\ & B_2 & A_2 & C_2 & \\ & & \ddots & \ddots & \ddots \\ & & & B_n & A_n & C_n \\ & & & & \ddots & \ddots & \ddots \end{pmatrix} \quad (1-20)$$

则称 $\{X(t), J(t)\}$ 是一个拟生灭过程(quasi birth and death process, 简记 QBD 过程), 其中所有子块都是 m 阶方阵, $A_k (k \geq 0)$ 有负的对角线元素和非负的非对角线元素,

并且行和非正。 B_k 和 C_k 都是非负阵,满足

$$(A_0 + C_0)e = (A_k + B_k + C_k)e = 0, k \geq 1$$

由式(1-20)与式(1-14)的比较可知, QBD 过程是经典生灭过程从一维状态空间到二维状态空间的推广。经典生灭过程的状态 $X(t) = k$, 现在被分解成 m 个子状态 $\{(k, 1), \dots, (k, m)\}$ 。这一结构顺应了刻画过程在多层次、多位相及变动参数情况下演化的要求。Evans(1967)首先在排队论分析中使用这类过程, Wallace(1969)在计算机系统研究中也遇到形如式(1-20)的过程, 并由他引入了“拟生灭过程”这一术语。20 世纪 70 年代以来, Neuts 等(1984)系统地发展了处理 QBD 过程的矩阵分析方法, 有力地推动了 QBD 过程在随机模型分析中的广泛应用。

对于形如式(1-20)的一般 QBD 过程, 尚未建立起平行于经典生灭过程的理论和处理方法。关于经典生灭过程平稳分布的 Karlin-McGregor 定理, 目前还不能推广到 QBD 过程。迄今广泛使用的只是一类特殊的 QBD 过程, 即生成元 Q 中的子块从某一个水平开始不再发生变化的情形。这时, 过程生成元形如:

$$\tilde{Q} = \begin{bmatrix} A_0 & C_0 & & & \\ B_1 & A_1 & C_1 & & \\ & \dots & \dots & \dots & \\ & & B_{F-1} & A_{F-1} & C_{F-1} \\ & & & B & A & C \\ & & & & B & A & C \\ & & & & & \dots & \dots & \dots \end{bmatrix} \quad (1-21)$$

处理 QBD 过程的方法, 完全依赖于生成元的特殊结构。为了更自然地引入这种方法, 注意到在经典生灭过程模型中与式(1-21)对应的典型例子是 $M/M/F$ 排队系统。下面使用一种特殊的技巧处理 $M/M/F$ 排队系统, 这种处理自然地导致了对式(1-21)、进行 QBD 过程解析分析的矩阵几何解方法。

在到达率 λ 和服务率 μ 的 $M/M/F$ 排队系统中, $L(t)$ 表示时刻 t 系统内顾客数, 也称为系统的队长。 $\{L(t), t \geq 0\}$ 是一个经典生灭过程, 其无穷小生成元如式(1-21)所示。

$$\tilde{Q} = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & \\ & 2\mu & -(\lambda + 2\mu) & \lambda & \\ & & \dots & \dots & \dots \\ & & & (F-1)\mu & -[\lambda + (c-1)\mu] & \lambda \\ & & & & c\mu & -(\lambda + c\mu) & \lambda \\ & & & & & c\mu & -(\lambda + c\mu) & \lambda \\ & & & & & & \dots & \dots \end{bmatrix} \quad (1-22)$$

在式(1-22)中, 自状态 $k = F$ 以后生率和灭率都不再随状态变化。过程的状态转移如图 1-5 所示。

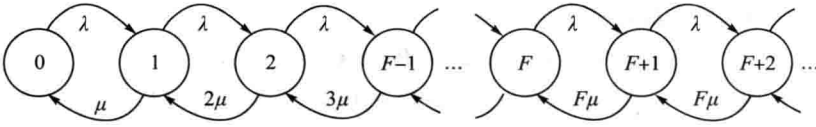


图 1-5 M/M/F 排队的状态转移

设 $\rho = \lambda(F\mu)^{-1} < 1$, 过程是正常返的。记 $\pi_k = \lim_{t \rightarrow \infty} P\{L(t) = k\}$, 当 $k > F$ 时, 平衡方程式(1-19)和生成元式(1-22)给出

$$\lambda\pi_{k-1} - (\lambda + F\mu)\pi_k + F\mu\pi_{k+1} = 0, k > 0 \quad (1-23)$$

现考虑二次方程

$$F\mu z^2 - (\lambda + F\mu)z + \lambda = 0 \quad (1-24)$$

它的两个根是 $z = \rho = \lambda(F\mu)^{-1} < 1$ 和 $z = 1$ 。 $z = \rho$ 是方程(1-24)的最小非负解。代入式(1-23)验证可知, $\pi_k = K\rho^{k-F}$, $k > F$, 满足式(1-23), 其中 K 是待定常数, 在式(1-23)中, 取 $k = F+1$, 并添加 λK 一项, 给出

$$\lambda\pi_F - \lambda K + [\lambda K - (\lambda + F\mu)K\rho + F\mu K\rho^2] = \lambda(\pi_c - K) = 0$$

于是 $K = \pi_c$ 。对 $k \geq c$ 时, 有

$$\pi_k = \pi_F \rho^{k-F}, k \geq F \quad (1-25)$$

现在, 只要需要确定 $\pi_0, \pi_1, \dots, \pi_F$ 。由平衡方程(1-19)和生成元式(1-22)推导, 并注意到 $\pi_{F+1} = \pi_c \rho$, 得到

$$\begin{cases} -\lambda\pi_0 + \mu\pi_1 = 0 \\ \lambda\pi_{k-1} - (\lambda - k\mu)\pi_k + (k+1)\mu\pi_{k+1} = 0, 1 \leq k \leq F-1 \\ \lambda\pi_{F-1} + [-\lambda + F\mu + F\pi_F\rho]\pi_F = 0 \end{cases}$$

如果引入 $F+1$ 阶方阵

$$B[\rho] = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & \\ & \ddots & \ddots & \ddots & \\ & & (F-1)\mu & -[\lambda + (F-1)\mu] & \lambda \\ & & & F\mu & -(\lambda + F\mu) + F\mu\rho \end{bmatrix} \quad (1-26)$$

并不计算出 $c_{\mu\rho} = \lambda$, 是为了在后面的讨论中加以比较。 $\pi_0, \pi_1, \dots, \pi_F$ 满足

$$(\pi_0, \dots, \pi_F)B[\rho] = \mathbf{0} \quad (1-27)$$

求解方程组(1-27)并应用表达式(1-25), 即可给出经典 M/M/F 排队系统的稳态分布

$$\pi_k = \begin{cases} K \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^k, 0 \leq k \leq F \\ K \frac{1}{F!} \left(\frac{\lambda}{\mu}\right)^F \rho^{k-F}, k > F \end{cases} \quad (1-28)$$

式中, 常数因子 K 可由正规化条件确定

$$K = \left\{ \sum_{k=0}^{F-1} \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^k + \frac{1}{F!} \left(\frac{\lambda}{\mu}\right)^F (1-\rho)^{-1} \right\}^{-1}$$

本节只对一类特殊的 QBD 过程给出了通过矩阵几何解进行分析处理的方法。正如已经指