国家自然科学基金资助出版

排队系统性能分析与Markov控制过程

Queueing System Performance Analysis and Markov Control Processes

殷保群 奚宏生 周亚平 著

中国科学技术大学出版社

国家自然科学基金资助出版

排队系统性能分析与Markov控制过程

殷保群　奚宏生　周亚平　著

中国科学技术大学出版社

2004 · 合肥

内 容 简 介

本书是基于性能势理论研究排队系统性能分析与Markov控制过程的一本专著，除第1章预备知识外，其余各章均为作者近几年来在该领域的最新研究成果．内容包括指数服务分布排队系统的性能分析、非指数服务分布排队系统的性能分析、排队网络基于理论计算的优化算法、排队网络基于样本轨道仿真的优化算法和半Markov控制过程等．

本书的读者对象为从事排队系统和Markov控制过程理论及应用研究或其他相关研究的研究生、教师以及有关的科技工作者．

图书在版编目（CIP）数据

排队系统性能分析与Markov控制过程/殷保群，奚宏生，周亚平著．一合肥：中国科学技术大学出版社，2004．5

ISBN 7-312-01652-9

I．排… II．①殷… ②奚… ③周… III．排队—系统—性能分析 IV．O226

中国版本图书馆CIP数据核字（2004）第023472号

中国科学技术大学出版社出版发行
（安徽省合肥市金寨路96号，邮编：230026
电话：0551-3602905 传真：0551-3602897
网址：http://www.press.ustc.edu.cn）
中国科学技术大学印刷厂印刷
全国新华书店经销

开本：787×1092/16 印张：12 字数：307千
2004年5月第1版 2004年5月第1次印刷
印数：1—1500册
ISBN 7-312-01652-9 / O·280 定价：25.00元

前　言

离散事件动态系统(DEDS)是系统和控制领域一门新的分支，它是由哈佛大学 Y. C. Ho 教授等人于 20 世纪 80 年代前后创立的，其名称也是由 Y. C. Ho 教授等人提出的. DEDS 是一类特殊的系统，大多是一些人造系统，一般不同于通常的连续变量动态系统(CVDS). 系统中存在着大量的离散事件，且不能由物理定律描述，而服从人为规定的一些复杂规则，因而无法用微分方程或差分方程来描述系统的演化过程. 系统的进化是由离散事件间错综复杂的相互作用决定的，且具有异步并发或随机等特征，这就决定了系统的建模与分析都不同于一般的 CVDS.

DEDS 研究的真正兴起出现在 20 世纪 80 年代初，但对它的研究最早可以追溯到对排队网络、数学规划、运筹决策及柔性生产系统等问题的研究. 虽然在最初研究这些问题时，DEDS 的概念可能还未出现，但这些问题本质上就是 DEDS 问题. 由于 DEDS 在当今一大批高新技术，诸如计算机网络、通信网络、柔性生产系统、交通管理系统及军事指挥的 C^3I 系统中，都有着极为广泛的应用，故越来越受到重视和关注.

在 DEDS 的研究中，与通常的系统一样，也存在着大量的控制和优化问题，并且在提法上与通常的控制系统也无多大差别. 它主要是通过控制系统的某些参数的变化，禁止某些事件出现，引导系统向人们期望的方向发展，以得到预期的目标. 通常控制系统理论中的许多基本概念，如稳定性、能控性、能观性以及最优控制等概念，都可以推广到 DEDS，但一般在表达方式和技术处理上会有所不同，甚至相差很大.

DEDS 的研究，根据所用模型和采用工具的不同，基本上可以分为三个不同的方向，即逻辑层次、时间层次和随机统计层次. 本书的主要研究工作是在随机统计层次模型下所做的. 在随机统计层次模型下，采用的系统模型主要是排队网络和随机过程. 一般来说，DEDS 中某个参变量的微小变化，都会导致系统以十分不同的方式运行，因此需要考虑系统中许多参变量的随机变化，从而在性能分析中需采用统计平均的方法. 可以说，从统计性能层次研究 DEDS，是 DEDS 研究的最初形式. 当初，Y. C. Ho 教授等人提出 DEDS 这个概念时，就是从该层次出发的. 在这一层次下，采用的主要方法，最典型也是最重要的方法是摄动分析(PA)方法. 最近二十多年来，经过众多学者的艰苦努力，从无穷小摄动分析(IPA)方法到非 IPA 方法，已经建立了一整套有用的技术，并被成功地应用于各种不同的系统分析与优化中. 其主要包括：建立 PA 理论(特别是 IPA 理论)，发展简单实用的 PA 算法，证明算法的收敛性，并将算法应用于性能优化等.

在随机统计模型下，采用的主要方法是摄动分析方法，它主要是由 Y. C. Ho 教授等人提出的. 它是一种将理论分析与计算机仿真技术相结合的方法，非常类似于 CVDS 中对非

线性系统按某条特定轨线线性化的方法．可以说，它是排队网络分析方法和计算机仿真技术一个创造性的结合．但它比纯粹用仿真求梯度的方法要节省大量的仿真机时，并且避免可能出现的数值计算困难与复杂性，也不需要纯解析方法所作的过多假设和推断，从而大大拓宽了应用范围．然而，由于 DEDS 适用的范围越来越广，研究的问题也越来越复杂，故传统意义上的 PA 方法，即无穷小摄动分析(IPA)方法，已表现出较大的局限性，因而近年来又相继提出了许多改进的 PA 方法．

由于 IPA 方法及一些改进的 PA 方法存在许多的缺陷和不足，特别是缺乏作为理论的普遍性，故曹希仁教授和陈翰馥院士于 1997 年在这个方向上提出了一种新的方法，我们称之为性能势方法．这种方法以 Markov 过程作为系统模型，并引入两个重要的概念：实现因子和性能势，前者类似于 IPA 方法中的实现因子，后者是对现代 Markov 过程势理论的新贡献．通过这两个基本量中的任何一个，给出了系统稳态性能关于无穷小矩阵摄动的灵敏度公式．由于实现因子和性能势可从系统的一条单一样本轨道上获得它们的精确估计，故这种方法可应用于 IPA 方法失效系统的性能分析与优化．而对 IPA 方法有效的系统，它同样可以提供精确的导数估计，故该方法为 IPA 方法和非 IPA 方法提供了一个一致的框架．

研究 DEDS 的主要目的是要在一定的条件下优化系统的性能，使之能更有效地运行，并为设计新系统或改进现有系统提供可靠的依据．这就是 DEDS 的优化问题．随着 DEDS 应用领域的日益发展，出现了许多更为复杂的人造系统，它要求决策者根据系统的实时演化提出相应的策略，使系统的整体性能达到最佳，这可视为 DEDS 的控制问题．

对于排队网络的性能分析与优化问题的研究，已有近百年的历史．最初可以追溯到对电话话务理论的研究，以后排队网络理论又逐渐应用于陆空交通、机器管理、水库设计和可靠性理论等领域．20 世纪 70 年代以来，随着计算机通信技术的广泛应用与发展，网络规模的不断扩大，设备的不断拓展，出现了越来越复杂的排队系统．对于这些复杂的排队系统，排队论的已有结果难以全面地进行深入、细致的分析，尤其是在很多情况下难以得到系统主要指标的解析解．因此，计算机仿真方法也就成为分析这些复杂系统的重要手段，这一方法能够为各种类型的实际问题提供数值解．另外，理论研究中有时要作一些假设，这些假设需要鉴定是否合乎实际．同时，理论研究的一些结论也应通过实际系统来检验，由于通常的实际系统都可以用计算机仿真来实现，这样就可以利用它来代替某些费用昂贵的试验，并且能够更快地积累数据．

如何合理地设计与控制排队网络，使得它既能满足顾客的需要，又能使系统的花费最为经济，这是排队网络的实际优化问题．随机因素在排队网络中起着根本性的作用．顾客到来的时刻一般是无法事先规定的，顾客要求服务的时间也是随不同顾客而异，并且是与服务者有关的一个随机量．因此，排队网络的优化本质上是一个随机优化问题．由于我们所考虑的排队网络基本上都是 Markov 型的，故排队网络的优化问题本质上可看作是 Markov 决策问题，也即在本书中被称为 Markov 控制过程的性能优化问题．

Markov 控制过程(MCP，又称为 Markov 决策过程或 Markov 决策规划等)是研究一类随机序贯决策问题的理论．所谓随机序贯决策问题，是指在一系列相继的或连续的时刻点

(决策点)上，决策者根据观察到的状态，从可用的若干决策中选择一个，并在一系列的决策点上执行所选择的一系列决策后，使系统在某种意义下(一般称为准则)获得最优或次最优的运行效果．系统在任一决策点处的状态是随机的，决策者在每一个决策点所采取的决策会影响到下一个决策时刻系统的运行，并以此影响将来．如果描述系统状态演化的是一个 Markov 过程，就是所谓的 Markov 控制过程，它是动态规划方法与 Markov 过程相结合的产物．MCP 既是随机运筹学的一门分支，也是应用概率论的一门分支，同时也可作为 Markov 系统最优控制的理论．它属于随机系统最优控制的范畴，更是和 DEDS 理论密不可分．实际上，它是随机 DEDS 惟一的动态控制方法．

本书作者有幸得到两项国家自然科学基金(69974037, 60274012)和一项安徽省自然科学基金(01042308)的支持，对排队系统、Markov 过程及半 Markov 过程的性能分析与优化问题进行系统、深入的研究，并取得了一系列的成果，与合作者一起完成已发表或待发表的论文达 50 余篇．主要工作包括以下几个方面：(1) 基于性能势理论，成功地解决了具有指数服务分布排队系统的性能灵敏度分析问题，并研究了传统 IPA 方法与现有方法中各基本量之间的关系；基于单样本轨道，给出了系统性能导数的串行及并行仿真估计算法．(2) 讨论了一类具有非指数服务分布，即相型服务分布排队系统的性能灵敏度分析问题，并进一步研究了两类具有一般分布的排队系统的性能灵敏度分析问题．(3) 基于性能势理论，成功地建立了 Markov 控制过程的性能优化理论，并将其应用到排队网络的性能优化中，同时给出了基于样本轨道仿真的串行及并行优化算法．(4) 运用等价 Markov 过程方法，成功地解决了半 Markov 过程的性能灵敏度分析问题，同时建立了半 Markov 控制过程的性能优化理论．

本书共 6 章，第 1 章介绍一些基本的数学知识，以及随机 DEDS 性能分析的两种基本方法．第 2 章到第 6 章均为作者近年来在随机 DEDS 领域的最新研究成果．第 2 章与第 3 章基于性能势方法，研究了各种类型排队系统的性能分析问题，并讨论了传统方法与现有方法中各基本量之间的关系．第 4 章与第 5 章研究了有限 Markov 决策问题，并将其结果成功地应用于排队网络的性能优化中，同时讨论了排队网络基于样本轨道仿真的优化算法．第 6 章研究了可数 Markov 决策问题、有限以及可数半 Markov 决策问题，建立了相应的性能优化理论．此外，本书后面还提供了 2 个附录．附录 1 较详细地讨论了非负矩阵的一些性质，特别是所谓的 Perron-Frobenius 定理及其推论，它们在相型分布以及有限 Markov 过程的研究中起着重要作用．附录 2 介绍了可数矩阵的一些基本概念和结论，它们在可数 Markov 控制过程的研究中起到了一定的作用．

在我们的前期研究工作中，孙德敏教授、曹希仁教授、杨孝先教授等给予了很多的指导和帮助．在本书的撰写过程中，作者所在实验室的许多研究生也给予了较大的支持和帮助，在这里特别提到的有唐昊博士、代桂平博士、李衍杰博士、康宇博士、高旭东硕士等等．中国科学技术大学出版社的领导和编辑也给予了大力支持．作者在此对他们表示衷心

的感谢. 作者也非常感激国家自然科学基金委以及安徽省科委对我们研究工作的资助.

由于作者水平有限，书中错误在所难免，恳请专家以及广大读者批评指正.

作　者

2003 年 10 月

目 录

第 1 章　预 备 知 识

在这一章里，我们将首先介绍在本书中要用到的一些基础知识，包括随机过程和排队模型方面的相关概念和结论，然后简单地介绍一下本书所涉及领域的两种常用的研究方法，包括无穷小摄动分析方法和 Markov 过程无穷小矩阵摄动方法及其相关概念和结果.

1.1　随 机 过 程

本节我们将复习一下随机过程的有关知识，包括概率论、Poisson 过程、Markov 链、Markov 过程、再生过程和 Markov 更新过程等，着重于这些过程的状态特征描述及稳态性能方面的概念和结果．本节材料可参看文献[1~4].

1.1.1　概率论

我们将以测度论的观点来简单地介绍一下概率论的一些基本概念和结论.

定义 1.1.1　设 Ω 是一个非空基本集，$\mathcal{F}$ 是 Ω 的一个子集族，且满足条件：

(1) 若 $E_1, E_2 \in \mathcal{F}$，则 $E_1 - E_2 \in \mathcal{F}$；

(2) 若 $E_n \in \mathcal{F}$，$n = 1, 2, \cdots$，则 $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$，

则称 $\mathcal{F}$ 为 Ω 上的一个 σ-环．若还有 $\Omega \in \mathcal{F}$，则称 $\mathcal{F}$ 为 Ω 上的一个 σ-域或 σ-代数.

设 $\mathcal{G}$ 是 Ω 的一个子集族，包含 $\mathcal{G}$ 的最小 σ-域称为由 $\mathcal{G}$ 生成的 σ-域．由实数空间 $\mathcal{R}$ 上的所有开区间生成的 σ-域称为 Borel σ-域或 Borel 集类，记为 $\mathcal{B}$.

定义 1.1.2　设 $\mathcal{F}$ 为 Ω 上的一个 σ-域，μ 是定义在 $\mathcal{F}$ 上的一个非负的，允许取 $+\infty$，但不恒取 $+\infty$ 的集函数，若它满足可列可加性：当 $E_n \in \mathcal{F}$，$n = 1, 2, \cdots$，$E_i \cap E_j = \varnothing$ $(i \neq j)$ 时，有

$$\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n),$$

则称 μ 是 $\mathcal{F}$ 上的一个测度，$(\Omega, \mathcal{F}, \mu)$ 称为一个测度空间，$\mathcal{F}$ 中的集合称为可测集，$\mu(E)$ 称为 E 的测度.

定理 1.1.1 设 μ 是 σ-域 $\mathcal{F}$ 上的一个测度，则有

(1) 规定性：$\mu(\varnothing) = 0$；

(2) 有限可加性：若 $E_1, \cdots, E_n \in \mathcal{F}$，$E_i \cap E_j = \varnothing$ $(i \neq j)$，则

$$\mu(\bigcup_{k=1}^{n} E_k) = \sum_{k=1}^{n} \mu(E_k);$$

(3) 单调性：若 $E_1, E_2 \in \mathcal{F}$，$E_1 \subset E_2$，则 $\mu(E_1) \le \mu(E_2)$.

定义 1.1.3 设 $(\Omega, \mathcal{F}, \mu)$ 为一个测度空间，若 $\mu(\Omega) = 1$，并记 $P(\cdot) = \mu(\cdot)$，称 P 为 σ-域 $\mathcal{F}$ 上的一个概率测度，并称 $(\Omega, \mathcal{F}, P)$ 为一个概率空间，而 $\mathcal{F}$ 中的任一元素称为一个事件. 换句话说，定义在 σ-域 $\mathcal{F}$ 上的一个集函数称为一个概率测度，若它满足条件：

(1) 对任意的 $A \in \mathcal{F}$，$P(A) \ge 0$；

(2) $P(\Omega) = 1$，$P(\varnothing) = 0$；

(3) 若 $A_n \in \mathcal{F}$，$n = 1, 2, \cdots$，$A_i \cap A_j = \varnothing$ $(i \neq j)$，则有

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

设 $(\Omega, \mathcal{F}, P)$ 为一个概率空间，则由定理 1.1.1 的（3）可以知道，对任意的 $A \in \mathcal{F}$，$P(A) \le 1$. 对 $A \in \mathcal{F}$，若 $P(A) = 1$，则称 A 为 P 的支集.

定义 1.1.4 设 $(\Omega, \mathcal{F}, P)$ 为一个概率空间，$B \in \mathcal{F}$，$P(B) > 0$，定义

$$P(A|B) = \frac{P(AB)}{P(B)}, \quad A \in \mathcal{F}, \tag{1.1.1}$$

则 $P_B(\cdot) = P(\cdot|B)$ 为定义在 $\mathcal{F}$ 上的一个概率测度，称为给定事件 B 的条件概率，而称 $(\Omega, \mathcal{F}, P_B)$ 为给定事件 B 的条件概率空间.

条件概率除了具有概率的有关性质外，还具有以下一些特殊性质：

(1) 设 $(\Omega, \mathcal{F}, P_A)$ 为一条件概率空间，$B \in \mathcal{F}$，$P_A(B) > 0$，则有

$$P_A(C|B) = P(C|AB) = P_{AB}(C); \tag{1.1.2}$$

(2) (乘法公式) 设 $A_n \in \mathcal{F}$，$n = 1, 2, \cdots, k$，且 $P(A_1A_2\cdots A_{k-1}) > 0$，则有

$$P(A_1A_2\cdots A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)P(A_k|A_1A_2\cdots A_{k-1}); \tag{1.1.3}$$

(3) (全概率公式) 设 $A\in\mathcal{F}$，$B_n\in\mathcal{F}$，$n=1,2,\cdots$，$B_i\cap B_j=\varnothing$ $(i\neq j)$，$P(B_n)>0$，$\bigcup_{n=1}^{\infty}B_n\supset A$，则有

$$P(A)=\sum_{n=1}^{\infty}P(B_n)P(A|B_n)\text{；} \tag{1.1.4}$$

(4) (Bayes 公式) 设 $A\in\mathcal{F}$，$B_n\in\mathcal{F}$，$n=1,2,\cdots$，$B_i\cap B_j=\varnothing$ $(i\neq j)$，$P(A)>0$，$P(B_n)>0$，$\bigcup_{n=1}^{\infty}B_n\supset A$，则有

$$P(B_n|A)=\frac{P(B_n)P(A|B_n)}{\sum_{i=1}^{\infty}P(B_i)P(A|B_i)}，\quad n=1,2,\cdots. \tag{1.1.5}$$

定义 1.1.5 设 $(\Omega,\mathcal{F},P)$ 为一个概率空间，$A_n\in\mathcal{F}$，$n=1,2,\cdots,k$，若对任意的 n （$1\le n\le k$）及任意的 $1\le i_1<i_2<\cdots<i_n\le k$，都有

$$P(A_{i_1}A_{i_2}\cdots A_{i_n})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_n})\text{，} \tag{1.1.6}$$

则称事件 $A_1,A_2,\cdots,A_n$ 是独立的.

定义 1.1.6 设 $(\Omega,\mathcal{F},P)$ 为一个概率空间，$X:\Omega\to\mathcal{R}$，若对任意的 $t\in\mathcal{R}$，均有 $\{X\le t\}=\{\omega\in\Omega:X(\omega)\le t\}\in\mathcal{F}$，则称 X 为 $(\Omega,\mathcal{F},P)$ 上的一个随机变量.

设 X 为概率空间 $(\Omega,\mathcal{F},P)$ 上的一个随机变量，称

$$F(t)=P\{X\le t\} \tag{1.1.7}$$

为随机变量 X 的分布函数. 易见，$F(t)$ 不减，且 $F(-\infty)=0$，$F(+\infty)=1$. 若存在一个非负函数 $f(t)$，使对任意的 $t\in\mathcal{R}$，均有

$$F(t)=\int_{-\infty}^{t}f(t)\mathrm{d}t\text{，} \tag{1.1.8}$$

则称 $f(t)$ 为随机变量 X 的概率密度函数，并称 X 为连续型随机变量. 随机变量 X 的数学期望或均值定义为

$$E[X]=\int_{-\infty}^{+\infty}tF(\mathrm{d}t)=\int_{-\infty}^{+\infty}tf(t)\mathrm{d}t\text{，} \tag{1.1.9}$$

它的方差定义为

$$D[X]=E[(X-E[X])^2]=E[X^2]-(E[X])^2\text{，} \tag{1.1.10}$$

风险率函数定义为

$$h(t)=\frac{f(t)}{1-F(t)}. \tag{1.1.11}$$

如果 $F(t)$ 在 $\mathcal{R}$ 上绝对连续，则 $f(t)=F'(t)$，从而有

$$f(t)=h(t)\mathrm{e}^{-\int_{-\infty}^{t}h(t)\mathrm{d}t}. \tag{1.1.12}$$

如果 X 表示一个事件的寿命，则 $h(t)$ 表示该事件活到时刻 t，而将在时间段 $[t,t+\Delta t)$ 内终结的比率.

下面将简单介绍一下在本书中要用到的有关概率论的一些基本事实和结论. 首先，我们来考虑所谓的逆变换方法，该方法可被用来从一个在 $[0,1)$ 上服从均匀分布的随机变量，生成一个具有已知分布函数 $F(t)$ 的随机变量，它在排队网络的仿真中有重要应用.

已知随机变量 X 的分布函数为 $F(t)$，注意到 $\xi=F(t)$ 服从 $[0,1)$ 上的均匀分布，为了生成随机变量 X，我们先产生一个随机变量 ξ，它服从 $[0,1)$ 上的均匀分布，然后令

$$t=F^{-1}(\xi)=\sup\{t:F(t)\le\xi\}, \tag{1.1.13}$$

这样生成的随机变量就一定具有分布函数 $F(t)$. 这是因为

$$P\{X\le t\}=P\{F^{-1}(\xi)\le t\}=P\{\xi\le F(t)\}=F(t).$$

下面我们来考虑随机变量序列的收敛性. 设 $(\Omega,\mathcal{F},P)$ 为一个概率空间，X_n，$n=1,2,\cdots$ 及 X 为定义在其上的随机变量，我们有下列四种收敛概念：

(1) 以概率 1 收敛（几乎肯定收敛）：若 $P\{\lim\limits_{n\to\infty}X_n=X\}=1$，或等价地，对任意的 $\varepsilon>0$，有 $\lim\limits_{n\to\infty}P\{|X_k-X|>\varepsilon, \text{对某个} k\ge n\}=0$，则称随机序列 $\{X_n\}$ 以概率 1 收敛到 X，记为 $\lim\limits_{n\to\infty}X_n=X$（w.p. 1)，也称随机序列 $\{X_n\}$ 几乎肯定收敛到 X，记为 $\lim\limits_{n\to\infty}X_n=X$ (a.s.).

(2) 以概率收敛：若对任意的 $\varepsilon>0$，有 $\lim\limits_{n\to\infty}P\{|X_n-X|\ge\varepsilon\}=0$，则称随机序列 $\{X_n\}$ 以概率收敛到 X，记为 $\lim\limits_{n\to\infty}X_n=X$ (P).

(3) r 阶收敛：设对 $r>0$，$E[|X_n|^r]<\infty$，$E[|X|^r]<\infty$，若 $\lim\limits_{n\to\infty}E[|X_n-X|^r]=0$，则称随机序列 $\{X_n\}$ r 阶收敛到 X，记为 $\lim\limits_{n\to\infty}X_n=X$ (r). 特别当 $r=1,2$ 时，分别称为均值收敛和均方收敛.

(4) 以分布收敛和弱收敛：设 $F_n(t)$ 和 $F(t)$ 分别是 X_n 和 X 的分布函数，$n=1,2,\cdots$，若 $\lim\limits_{n\to\infty}F_n(t)=F(t)$，则称随机序列 $\{X_n\}$ 以分布收敛到 X，记为 $\lim\limits_{n\to\infty}X_n=X$ (d)，也称分布函数序列 $\{F_n(t)\}$ 弱收敛到 $F(t)$.

可以证明：$\lim\limits_{n\to\infty} X_n = X$ (a.s.) $\Rightarrow$ $\lim\limits_{n\to\infty} X_n = X$ (P) $\Rightarrow$ $\lim\limits_{n\to\infty} X_n = X$ (d)；$\lim\limits_{n\to\infty} X_n = X$ (r) $\Rightarrow \lim\limits_{n\to\infty} X_n = X$ (P)．

定理 1.1.2（单调收敛定理） 设$\{X_n\}$为一个单调不降的非负随机变量序列，如果$\lim\limits_{n\to\infty} X_n = X$ (a.s.)，则有

$$\lim_{n\to\infty} E[X_n] = E[X].$$

定理 1.1.3（Lebesgue 控制收敛定理） 设$\{X_n\}$为一随机变量序列，若存在一个具有有限均值的随机变量Y，使对任意的n，有$|X_n| \le Y$，且$\lim\limits_{n\to\infty} X_n = X$ (P)，则

$$\lim_{n\to\infty} E[X_n] = E[X].$$

定义 1.1.7 设$\{X_n\}$为定义在概率空间$(\Omega, \mathcal{F}, P)$上存在均值的随机变量序列，若$\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=1}^{n}\{X_k - E[X_k]\} = 0$ (P)，则称$\{X_n\}$服从（弱）大数定律．

定理 1.1.4 设$\{X_n\}$为定义在概率空间$(\Omega, \mathcal{F}, P)$上的随机变量序列，若对任意的自然数n，$D[\sum\limits_{k=1}^{n} X_k]$存在，且$\lim\limits_{n\to\infty}\frac{1}{n^2}D[\sum\limits_{k=1}^{n} X_k] = 0$，则$\{X_n\}$服从大数定律．

定理 1.1.5（辛钦大数定律） 设$\{X_n\}$为相互独立同分布的随机变量序列，则$\{X_n\}$服从大数定律的充分必要条件是X_1具有有限的均值．

定义 1.1.8 设$\{X_n\}$为定义在概率空间$(\Omega, \mathcal{F}, P)$上存在均值的随机变量序列，若$\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=1}^{n}\{X_k - E[X_k]\} = 0$ (a.s.)，则称$\{X_n\}$服从强大数定律．

定理 1.1.6 设$\{X_n\}$为相互独立的随机变量序列，若$\sum\limits_{n=1}^{\infty}\frac{D[X_n]}{n^2} < \infty$，则$\{X_n\}$服从强大数定律．

定理 1.1.7 设$\{X_n\}$为相互独立同分布的随机变量序列，则$\{X_n\}$服从强大数定律的充

分必要条件是 X_1 具有有限的均值.

设 $(\Omega,\mathcal{F},P)$ 为一个概率空间，X_t，$t\in T$ 为定义在 $(\Omega,\mathcal{F},P)$ 上的一族随机变量，我们称 $X=\{X_t;t\in T\}$ 为定义在 $(\Omega,\mathcal{F},P)$ 上的一个随机过程. 这里，$T\subset\mathcal{R}$ 称为随机过程 X 的参数集. 如果 $T=\mathcal{Z}^+=\{0,1,2,\cdots\}$，则称 X 为离散时间的随机过程，一般用 n 表示离散的时间参数；如果 $T=[a,b]$，或 $\mathcal{R}$，或 $\mathcal{R}^+=[0,+\infty)$，则称 X 为连续时间的随机过程，一般用 t 表示连续的时间参数. 令 $\boldsymbol{\Phi}$ 为随机过程 X 的所有可能取值组成的集合，我们称 $\boldsymbol{\Phi}$ 为 X 的状态空间，$\boldsymbol{\Phi}$ 中的元素称为状态. $\boldsymbol{\Phi}$ 可以是一个有限集，或可数无限集，或不可数无限集. 我们也称 X 为定义在状态空间 $\boldsymbol{\Phi}$ 上的一个随机过程.

1.1.2 Poisson 过程

定义在概率空间 $(\Omega,\mathcal{F},P)$ 上的一个随机过程 $N=\{N_t;t\geq 0\}$，如果它是非减的，纯跳跃的，右连续的，且 $N_0=0$，则称其为一个到达过程. 应用最广泛的一类到达过程是 Poisson 过程.

定义 1.1.9 状态空间 $\boldsymbol{\Phi}=\{0,1,2,\cdots\}$ 上的一个到达过程 $N=\{N_t;t\geq 0\}$ 称为 Poisson 过程，如果它满足下列条件：

(1) 对任意的 $s,t\geq 0$，$N_{t+s}-N_t$ 与 $\{N_u;u\leq t\}$ 独立；

(2) 对任意的 $s,t\geq 0$，$N_{t+s}-N_t$ 的分布与 t 无关.

我们有下列引理，它描述了 Poisson 过程的基本特征，通过它可给出 Poisson 过程的多种定义.

引理 1.1.1 如果 $N=\{N_t;t\geq 0\}$ 是一个 Poisson 过程，则

(1) 存在一个常数 $\lambda\geq 0$，使对所有的 $t\geq 0$，$P\{N_t=0\}=\mathrm{e}^{\lambda t}$；

(2) $\lim\limits_{t\to 0}\dfrac{P\{N_t\geq 2\}}{t}=0$；

(3) $\lim\limits_{t\to 0}\dfrac{P\{N_t=1\}}{t}=\lambda$.

我们称引理 1.1.1 中的常数 λ 为 Poisson 过程 $N=\{N_t;t\geq 0\}$ 的到达率.

定理 1.1.8 如果 $N=\{N_t;t\geq 0\}$ 是一个到达率为 λ 的 Poisson 过程，则对任意的 $t\geq 0$，有

$$P\{N_t=k\}=\frac{\mathrm{e}^{-\lambda t}(\lambda t)^k}{k!},\quad k=0,1,\cdots. \tag{1.1.14}$$

设 $N=\{N_t;t\geq 0\}$ 是一个到达率为 λ 的 Poisson 过程，$0=T_0<T_1<T_2<\cdots$ 是相继的到达时刻，则我们有下列引理：

引理 1.1.2　如果 $N=\{N_t;t\ge 0\}$ 是一个到达率为 λ 的 Poisson 过程，则对任意的 $n\ge 0$，有

$$P\{T_{n+1}-T_n\le t|T_0,\cdots,T_n\}=1-\mathrm{e}^{-\lambda t},\ t\ge 0. \tag{1.1.15}$$

由引理 1.1.2 可知，一个到达率为 λ 的 Poisson 过程，其相继到达时间间隔 $T_{n+1}-T_n$，$n=0,1,\cdots$ 相互独立，且具有共同的指数分布 $1-\mathrm{e}^{-\lambda t}$．由这个引理，可给出 Poisson 过程的另一个等价定义：

定理 1.1.9　设 $0=T_0<T_1<T_2<\cdots$ 是一个到达过程 $N=\{N_t;t\ge 0\}$ 的相继到达时刻，则 N 是一个 Poisson 过程的充分必要条件为到达时间间隔 $T_{n+1}-T_n$，$n=0,1,\cdots$ 相互独立，且具有共同的指数分布 $1-\mathrm{e}^{-\lambda t}$．

Poisson 过程还有许多其他等价的定义，请参看文献[2].

1.1.3 Markov 链

如果一个随机过程在已知当前状态的情况下，它未来的行为与它过去的历史无关，则我们一般称该随机过程具有 Markov 性．Markov 链和下一小节将要介绍的 Markov 过程都是这样的随机过程.

为叙述方便，以下如无特殊声明，均设状态空间 $\Phi=\{1,2,\cdots\}$ 为一个可数集.

定义 1.1.10　状态空间 Φ 上的一个随机过程 $X=\{X_n;n\ge 0\}$，如果它对任意的 $j\in\Phi$ 及 $n\ge 0$，满足

$$P\{X_{n+1}=j|X_0,\cdots,X_n\}=P\{X_{n+1}=j|X_n\}, \tag{1.1.16}$$

则称 X 为一个具有状态空间 Φ 的 Markov 链.

对 Markov 链 $X=\{X_n;n\ge 0\}$，条件概率 $P\{X_{n+1}=j|X_n=i\}$，$i,j\in\Phi$ 称为 Markov 链 X 的转移概率，若它与 n 无关，记 $P\{X_{n+1}=j|X_n=i\}=P_{ij}$，则称其为时间齐次的，简称齐次的．本书中出现的 Markov 链均是齐次的．矩阵 $P=[P_{ij}]$ 称为 Markov 链 X 的转移矩阵．显然，对任意的 $i,j\in\Phi$，$P_{ij}\ge 0$，且

$$\sum_{j\in\Phi}P_{ij}=1.$$

满足此条件的矩阵一般称为一个 Markov 矩阵或随机矩阵.

对 $i,j\in\Phi$，及 $k\ge 0$，令 $P_{ij}^{(k)}$ 为矩阵 P^k 的第 (i,j) 个元素，则易证

$$P\{X_{n+k}=j|X_n=i\}=P_{ij}^{(k)}\ . \tag{1.1.17}$$

因此，矩阵 P^k 称为 Markov 链 X 的 k-步转移矩阵．这里，特别 $P^0=I$ 为单位矩阵．显然，对任意的 $n,m\ge 0$，有